Doctoral Thesis (Abridged)
博士論文 (要約)

Endocrine network essential for energy homeostasis in the two-spotted cricket, *Gryllus bimaculatus*

（フタホシコオロギにおけるエネルギー恒常性に重要な内分泌系ネットワーク）

周 已筠

Doctoral Thesis (Abridged)

博士論文 (要約)

Endocrine network essential for energy homeostasis in the two-spotted cricket, *Gryllus bimaculatus*

（フタホシコオロギにおけるエネルギー恒常性に重要な内分泌系ネットワーク）

周 已筠

Graduate School of Frontier Sciences

The University of Tokyo

Table of contents

Acknowledgements

I would like to express my deepest gratitude to Professor Shinji Nagata for his constant guidance and cheerful encouragement throughout this study and during my doctor course. Not only the supervision and suggestion on the research, but also trigger my interest for it.

High tribute shall be paid for Dr. Haruka Endo and Dr. Keisuke Fukumura, whose profound knowledge and experience are my guiding stars on the way forward. I am also grateful to all the members in the laboratory of Molecular recognition for constant support and encouragement both in research and communication. Also, special acknowledgements should be paid to Professor Takeda Makio from Kobe University for kindly providing the antiserum of CCAP.

Last but not the least, I would like to thank my parents a for supporting me spiritually throughout doctor course and my life in general.

This work was partly supported by JSPS KAKENHI JP201822078.

Abbreviations

7-TM: Seven-transmembrane

AcG: Accessory gland

ACP: AKH/Corazonin-related peptide

AG: Abdominal ganglia

AKH: Adipokinetic hormone

APC: AKH-producing cell

Ast-C: Allatostatin-C

BLAST: The Basic local alignment search tool

Br: Brain

Burs-α: Bursicon-α

CA: Corpora allata

cAMP: Cyclic Adenosine monophosphate

CC: Corpora cardiaca

CCAP: Crustacean cardioactive peptide

CCAPR: CCAP receptor

CHCA: α-Cyano-4-hydroxycinnamic acid

CNS: Central nervous system

CoG: Cluster of orthologous groups of proteins

CRZ or Crz: Corazonin

DAG: Diacyl glycerol

DCM: Dichloromethane

DH: Diuretic hormone

DNA: Deoxyribonucleic acid

dsRNA: Double-stranded RNA

EDT: 3,4-Ethoxylene dioxythiophene

EH: Eclosion hormone

FB: Fat body

FG: Foregut

G-6-DH: Glucose-6-phosphate dehydrogenase

Glut1: Glucose transporter

GO: Gene ontology

Gp: Glycogen phosphatase

GPCR: G-protein coupled receptor

HG: Hindgut

HL: Hemolymph

HPLC: High performance liquid chromatography

IPC: Insulin-producing cell

IRS: Insulin receptor substrate

KEGG: Kyoto Encyclopedia of genes genomes

MALDI-TOF: Matrix-assisted laser desorption ionization-time of flight

MG: Midgut

MIP: Myoinhibitory peptide

MS: Mass spectrometry

MT: Malpighian tubules

Mu: Muscle

n.s.: Not significant

NCBI: National center for biotechnology information

NCC: Nervi corporis cardiaci

NGS: Next-generation sequencer

Op: Optic lobe

PBS: Phosphate-buffered saline

PCR: Polymerase chain reaction

PDF: Pigment dispersing factor

PI: Pars intercerebralis

Real time qRT-PCR: Real time quantitative RT-PCR

RNA: Ribonucleic acid

RNAi: RNA interference

RT-PCR: Reverse transcription-PCR

SG: Subesophageal ganglion

sNPF: Short neuropeptide F

TAG: Triacyl glycerol

Te: Testis

TFA: Trifluoroacetic acid

TG: Thoracic ganglia

TMSBr: Trimethylsilyl bromide

TPP: Trehalose phosphate phosphatase

TPS: Trehalose phosphate synthetase

Tre: Trehalase

Tret1: Trehalose transporter

VNS: Ventral nervous system

Introduction

Energy homeostasis and its regulation by endocrine factors

Insects, as heterotrophic organisms, have to obtain energy from outside sources. The dietary energy is allocated for all living processes *in vivo* (Fig. 0-1 A). The concept of energy homeostasis was elaborated by Walter Cannon (Cannon, 1929) and Claude Bernard (Holmes, 1986) with an essential role for survival, which promises a coordinated regulation on food intake and energy expenditure to maintain a fairly stable condition. In related researches, energy homeostasis has been frequently illuminated as the mechanisms by which all living creatures satisfy their energy demands, but also manipulates food intake when expenditure changes (Chapelot and Charlot, 2019). However, a linear relationship is not adequate for describing energy expenditure and food intake. In mammals, overfeeding leads to a 7 to 18% increase in total energy expenditure, allowing ~25% of excess energy to be dissipated excess amount of food intake causes a 7 to 18% increase in total energy expenditure, allowing approximately 25% of over-loaded energy to be consumed (Klein and Goran, 1993), indicating the dynamic correlation between energy intake and expenditure for maintaining energy homeostasis.

To maintain energy homeostasis, animals rely on various endogenous signals, such as circulating metabolites, neuropeptides and neuromodulators (Fig. 0-1 B). Flexible physiology and behavior enable animals to adapt against extendedly varied challenges. The nervous system and endocrine control organize effectively in these adaptations and guarantee coordination between the external and the internal environments, in addition to the regulations of appropriate balanced nutrients and various living processes. Endocrine factors have been demonstrated as regulators in energy homeostasis in various insects. Different from synaptic transmission, endocrine factors deliver their signals slowly to their target tissues or cells, because neuropeptides and peptide hormones are synthesized and released into hemolymph. Second messages mediate intracellular signals after the combination endocrine factors (ligands) and their corresponding receptors, which are G-protein coupled receptors (GPCRs) in most cases. The activated intracellular signals with specific spatial distributions promise the specificity of neuropeptide signals. The variety of endocrine factors and the activated downstream signals enable them to involve the regulations in almost all the living processes, especially the energy homeostasis that is composed of complex collaboration between different organs and different living processes. Hence, multiple neuropeptides are recognized to contribute to an extensive range of temporal and spatial harmonization (Kim et al., 2017).

To date, energy homeostasis has been imagined as a fixed programmed value, but only studies about adaptive responses to external stimulus are understood. In a review of energy homeostasis, Chapelot and Charlot elaborate that energy homeostasis can be considered as a stochastic value with some possibility to be predicted with an expected accuracy if enough determinants are known (Chapelot and Charlot, 2019). However, accumulation of these determinants and mathematically calculating the index will be daunting. Indeed, multiple metabolites, hormones, neurotransmitters and neurons, and new molecules with regulatory roles in energy homeostasis have been constantly discovered in terms of their functions in food intake and energy expenditure (Gavrieli and Mantzoros, 2016).

Communications or cross-talk between neuropeptides

Gastrointestinal tract collects dietary information and CNS activates adaptive feeding behavior and metabolic flux. The coordinated regulation of the gastrointestinal tract and brain is elaborated as brain-gut axis, which are composed of bidirectional passes from both the CNS and the intestine. The gastrointestinal tracts of animals have enteroendocrine cells (EECs), which perceive the internal intestinal information and produce neuropeptide and peptide hormones to secrete for control of various physiological events including intestinal contraction, nutrient homeostasis and appetite (Miguel-Aliaga et al., 2018).

In mammals, Glucagon-like peptide-1, pancreatic polypeptide, peptide tyrosine-tyrosine and ghrelin are experimentally proved to be intestinal peptides that contribute to energy metabolism (Maric et al., 2014, Murphy et al., 2006). Moreover, ghrelin secreted from intestine suppresses insulin secretion. By contrast, cholecystokinin can enhance the secretion of insulin and other hormones (Lo et al., 2011); (Rehfeld et al., 1980). Hence, the nutritional states are regulated by intestinal peptides relaying the modulation of the secretion of insulin in the beta-cells and other neuropeptides in CNS. Such endocrinal situations are similar to those in invertebrates. To date, in the fruit fly, *Drosophila melanogaster*, the midgut enteroendocrine cells express 11 neuropeptides, including Allatostatin-A, Allatostatin-B, Allatostatin-C, Bursicon-α, CCHamide-1 and -2, Diuretic hormone 31, Neuropeptide F, Orcokinin-B, Short neuropeptide F and Tachykinin, although their distributions are varied in larvae and adult (Nassel and Zandawala, 2019). Besides, *D. melanogaster* intestine expresses 23 neuropeptide receptors, whose ligands are all express in the brain. These neuropeptide signals in intestine transfer the ‘information’ of the fluctuation of the digested compounds or metabolites, and sequentially activate downstream

signals for adaptive transitions. For example, DH31-producing cells in intestine can sense the dietary proteins and amino acids (Park et al., 2016). CCHamide-2 from intestine is thought to stimulate brain insulin-producing cells to regulate food intake (Ren et al., 2015). Not only the brain-gut axis, communication cross organs occurs between brain and fat body. For example, Bursicon-α-producing neurons are activated by dietary sucrose, and Bursicon-α can activate Lgr2 receptors in the brain neurons, which downregulate Adipokinetic hormone (AKH) signaling to involve in regulation of energy homeostasis (Scopelliti et al., 2019). Indeed, unidirectional signal activities are discovered in other organs. For example, Allatotropin (AT) stimulates juvenile hormone biosynthesis in the corpora allata in *Bombyx mori* (Yamanaka et al., 2008). Not only organ-cross communications, an antagonistic effect of sNPF in modulating feeding motivation was proved within intestine of *B. mori* (Matsumoto et al., 2019).

The study on neuropeptides communication reveals the high possibility that the neuropeptides have been practically collaborated each other or are coordinately regulated.

Endocrine network for understanding neuropeptides in energy homeostasis

The concept of networked regulation has been discussed for a long time. In mammals, networks involved in the regulation of appetite, thermogenesis and energy homeostasis are concluded in several specific regions in the brain which generate cognitive and autonomic systems of the brain. These networks are modulated by peripheral factors including hormones from adipose and intestinal tissues, addressing the brain via the circulation and neuronal signals to transmit the information about energy balance and nutritional state to CNS. Actually, the endocrine network which is constructed by the neuropeptides and peptide hormones related to food intake has been reported recently with abundant connections between nodes (Fig. 0-2). However, as the network is generated from conclusions, which are collected from different studies, the validity in individual levels cannot be promised.

For endocrinal regulation of energy homeostasis, it is almost a common sense that a single neuropeptide is unlikely to be the only player capable of making decisions in entire metabolic states. Instead, endocrine network can mediate the cues from multiple factors to make a ‘metabolic’ decision (Banerjee et al., 2017). The orchestration of multiple neuropeptides promises energy homeostasis. For example, sNPF and TK receptors expressed in olfactory sensory neurons modifies the sensitivity to specific odorant inputs in the hungry fly, eventually increasing food odor firing recurrence and activating food-searching behavior (Root et al., 2011); (Ko et al., 2015). Meanwhile, NPF enhances the sensitivity of the receptor against sweet compounds, by contrast,

sNPF reduces the sensitivity of the receptor against bitter compounds (Inagaki et al., 2014); (Yu et al., 2016); (Sayin et al., 2018). The systemically integrated neuropeptide information results in the adaptive feeding behavior to maintain energy homeostasis. Thus, it is necessary to understand if a specific neuropeptide serves multiple disparate functions to orchestrate systemically unified functions.

Collectively, a comprehensive understanding of neuropeptides involved in energy homeostasis is imperative for insects, which could be realized at an individual level.

Advantages of the study using *G. bimaculatus*

The two-spotted cricket *Gryllus bimaculatus* (Fig. 0-3), which is used as a research target for investigating mechanisms of development and regeneration from early 1990s. Developmental mode of *G. bimaculatus* is the typical of arthropods compared to that of *D. melanogaster* (Hadley and Mito, 2017). In contrast to mammals, one predominantly different structure in insects is the circulatory system, indicating that insects have the distinctive nature of delivering system of oxygen and metabolites to the CNS (Maier and Orion, 1969).

All stages' animals of *G. bimaculatus* are easily maintained and prepared in the laboratory, which might be the biggest reason to be widely used for studying insect physiology and neurobiology so far; for example, Allatostatin B (MIP) was originally identified in *G. bimaculatus* as early as 1990s (Lorenz et al., 1995). Until now, *G. bimaculatus* has become suitable as an appropriate model insect for addressing the molecular mechanisms of feeding behavior and more applicable for observing behavior because of its high locomotor activity (Hadley and Mito, 2017). In addition, high efficiency of RNA interference (RNAi) in cricket is convenient for design of the RNAi-based experiments on behavioral assays through loss of function.

Objectives of this study

In this study, I sought to describe a prototype of endocrine network that is essential for energy homeostasis at an individual level in *G. bimaculatus*. To address the endocrine network, following questions would be answered through this study.

1) Is endocrine network essential for maintaining energy homeostasis?

Binary communications, such as cross-talk between two neuropeptides, have been studied for a long time. However, multivariant communications are poorly understood. Importance of endocrine network should be

confirmed at an individual level in insects.

2) What is the arrangement of endocrine network in maintaining energy homeostasis? Are there any important regulators that are extensively related to other neuropeptides?

   The involvement of multiple neuropeptides determines the complexity of endocrine network in maintaining energy homeostasis. Although proof of communications between neuropeptides are plentiful, especially for Insulin/AKH signals, the frequently co-researched endocrine factors have not been considered as an extensive regulator among neuropeptides. Are there any possibilities for several neuropeptides to dominate the regulation of energy homeostasis?

3) How can 'information' from peripheral or circulating metabolites be delivered to CNS to maintain energy homeostasis, especially with the existence of hemolymph-brain barrier (HBB)?

   The existence of brain-gut axis in insects somehow proves the possibility of some neuropeptides passing through the HBB and functioning as neurotransmitters. However, except those neuropeptides, how do most neuropeptides 'send' information to brain? Is there any tissue in insects that resembles hypothalamus in mammals?

4) What relationship is present between circulating metabolites and endocrine network in maintaining energy homeostasis in insects?

   As stated above, energy from diets is metabolized and turns into different structures of metabolites, such as carbohydrates and lipids. The regulation of metabolites in circulation by neuropeptides needs to be elucidated in the endocrine network.

Challenges of this study

As stated above, more and more neuropeptides have been proved as regulators of biological and physiological processes. In this study, other than conventional neuropeptide research and characterization of a specific neuropeptide in a specific function, I focused on the overview of communications between neuropeptides across cerebral and peripheral tissues based on the hypothesis of endocrine network. However, since endocrine network-related literatures are extremely limited, exploring feasible methods is the most challenging part in this study. Highly conserved endocrine network has been studied in the reproductive system of D. melanogaster (Meiselman et al., 2017). However, in the above study, the appearance of endocrine factors is limited to Ecdysis-

triggering hormone (ETH), ecdysone (20E) and Juvenile hormone (JH), whereas more endocrine factors are proved to be involved in the reproductivity (Arakane et al., 2008). The investigation of endocrine network needs to include neuropeptides as much as possible for vividly revealing the dynamic regulation of energy homeostasis. Thus, the massive work of synthesizing neuropeptides and figuring out their regulatory roles in energy homeostasis are tough. Due to the experimental limitation and massive amounts of research targets, I decided to construct an endocrine network based on key-regulators and their associated signals (Fig. 0-4). Key regulators can be determined after screening of neuropeptides. Also, its representativeness provides the possibility of study on the endocrine network through manipulation of the key regulators.

Another challenging issue is to understand the comprehensive interplay among neuropeptides from the obtained results of each neuropeptide's role in energy homeostasis to generate the endocrine network. Due to the above limitation, research on the detail mechanism of each neuropeptide involving in energy homeostasis is difficult in this study. Instead, I presented here their function and communication between neuropeptides by transcriptional regulation and conditioned peptidome analyses.

Components of this study

Four chapters are enclosed in this study for revealing the prototype of endocrine network (Fig. 0-5): (1) A virtual network of neuropeptides in energy homeostatic regulation is described by bibliometric mapping. Followingly, all neuropeptides in *G. bimaculatus* were identified and evaluated their regulatory function responding to nutrient, feeding motivation, food intake and excretion. AKH and AKH/Corazonin-related peptide (ACP) were screened as key regulators. Adaptive modulation by transcriptional alteration of the selected GPCRs in fat body and corresponding ligands in brain were evaluated. (2) RNA-sequencing analysis of CC was performed for screening GPCRs in the CC and CC-synthesizing neuropeptides. Also, the whole transcriptome of CC was annotated. (3) Elucidating the counteractive regulation of energy homeostasis by crustacean cardioactive peptide (CCAP) and Myosuppressin (MS), which is revealed by their alterations in food intake, excretion, carbohydrate and lipid level in hemolymph. (4) As main hemolymph carbohydrates in cricket, trehalose and glucose levels were evaluated for their influences on energy homeostasis. Abnormal carbohydrate concentration was utilized to evaluate endocrine network activity. Collectively, I described the prototype of endocrine network that is essential for energy homeostasis in the circumstance of diet insufficient, in which neuropeptides are considered as nodes

and communication between them are considered as linkages.

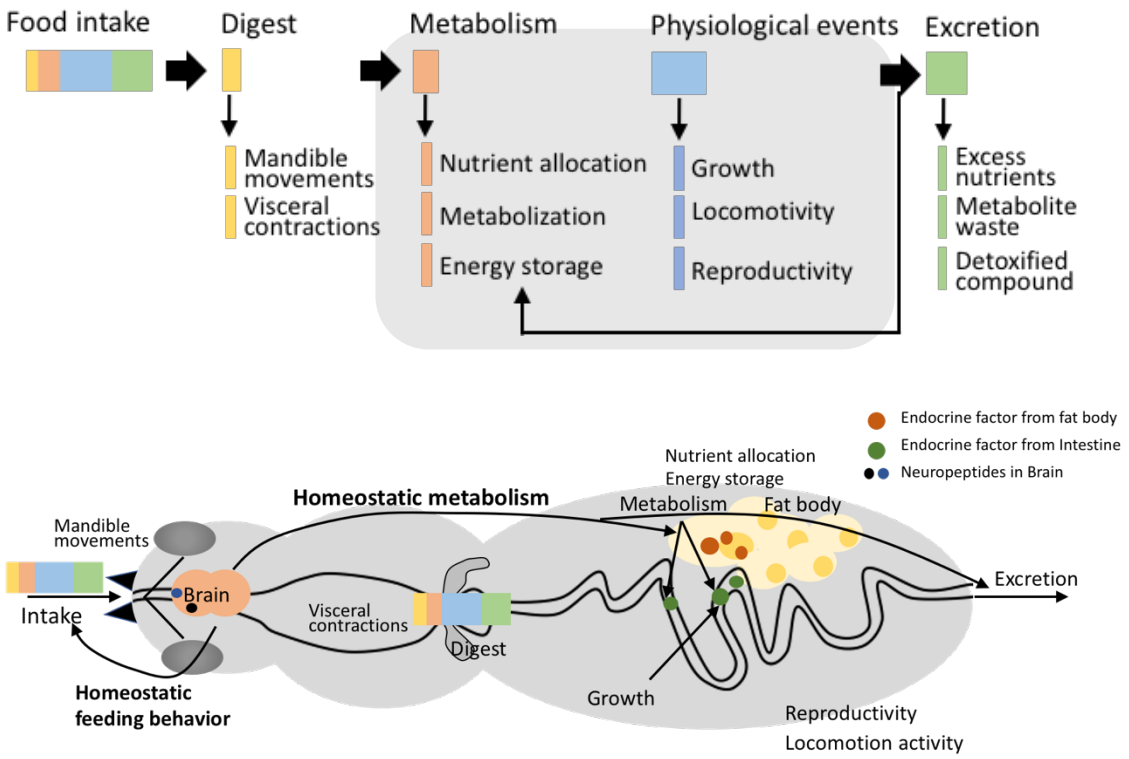

Figure 0-1 Physiological flow of energy and its maintaining mechanism by endocrine network

(A) Dietary nutrient flow in insect body (Nagata and Zhou, 2020). Dietary source changes by several feeding processes from ingestion to excretion. All steps include further detailed physiological processes. Harmonization of all processes accomplishes to obtain nutrients effectively and to maintain energy homeostasis.

(B) Schematic endocrine network of how to maintain energy homeostasis along with feeding behavior in insects. The stored energy in fat body and free energy in intestine give endocrinal signals, which target at central nervous system and activate homeostatic feeding behavior and metabolic flux. The metabolic flux and signals in brain control the balance between energy storage and excretion of excess energy.

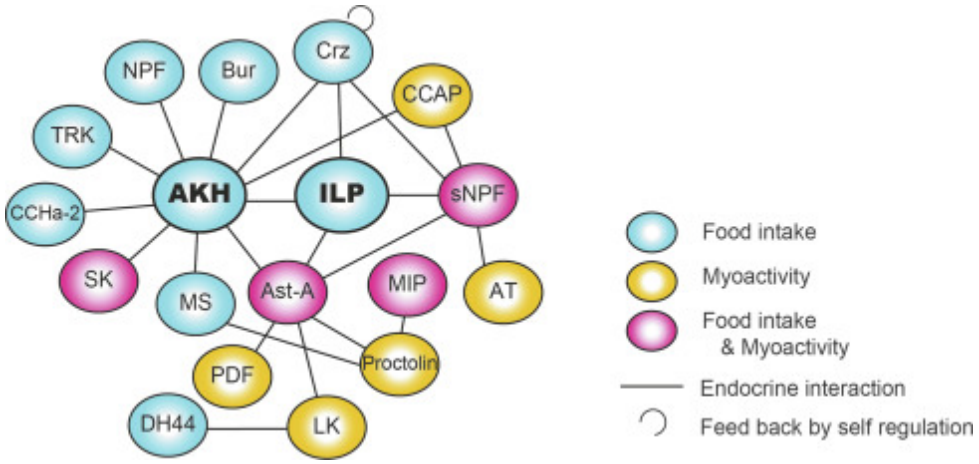

Figure 0-2. Schematic simple network graph generated by knowledges of neuropeptides and peptide hormones related to food intake (Nagata and Zhou, 2020).

All nodes are corresponding to neuropeptides and peptide hormones, which are selected by the previous literatures related to food intake in insects. –AKH, adipokinetic hormone; Ast-A, allatostatin-A; AT, allatotropin; Bur, bursicon; CCAP, crustacean cardio-active peptide; Crz, corazonin; DH44, diuretic hormone 44; ILP, insulin-like peptide; MIP, Myoinhibitory peptide, Ast-B, allatostatin-B; MS, myosuppressin; NPF, neuropeptide F; PDF, pigment dispersing factor; sNPF, short neuropeptide F; SK, sulfakinin; TRK, tachykinin-related peptide.

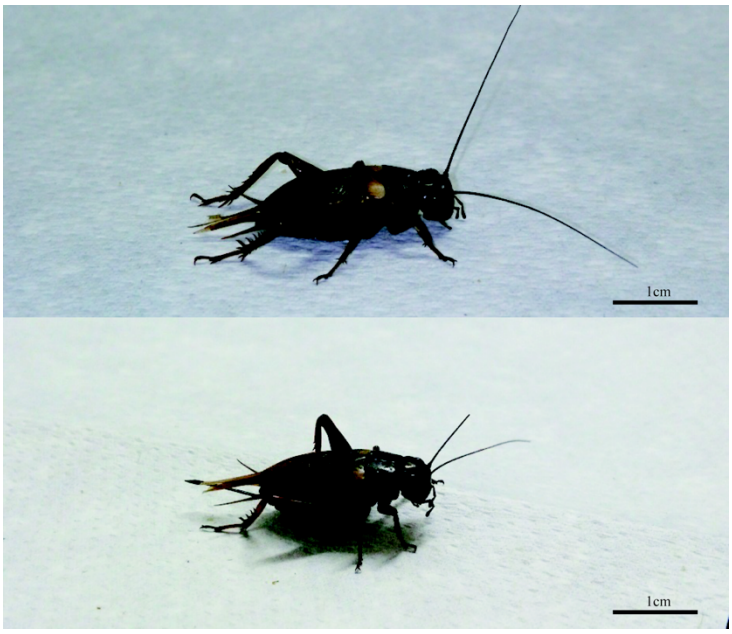

Figure. 0-3. The two-spotted crickets, *G. bimaculatus*; male (upper) and female (below).

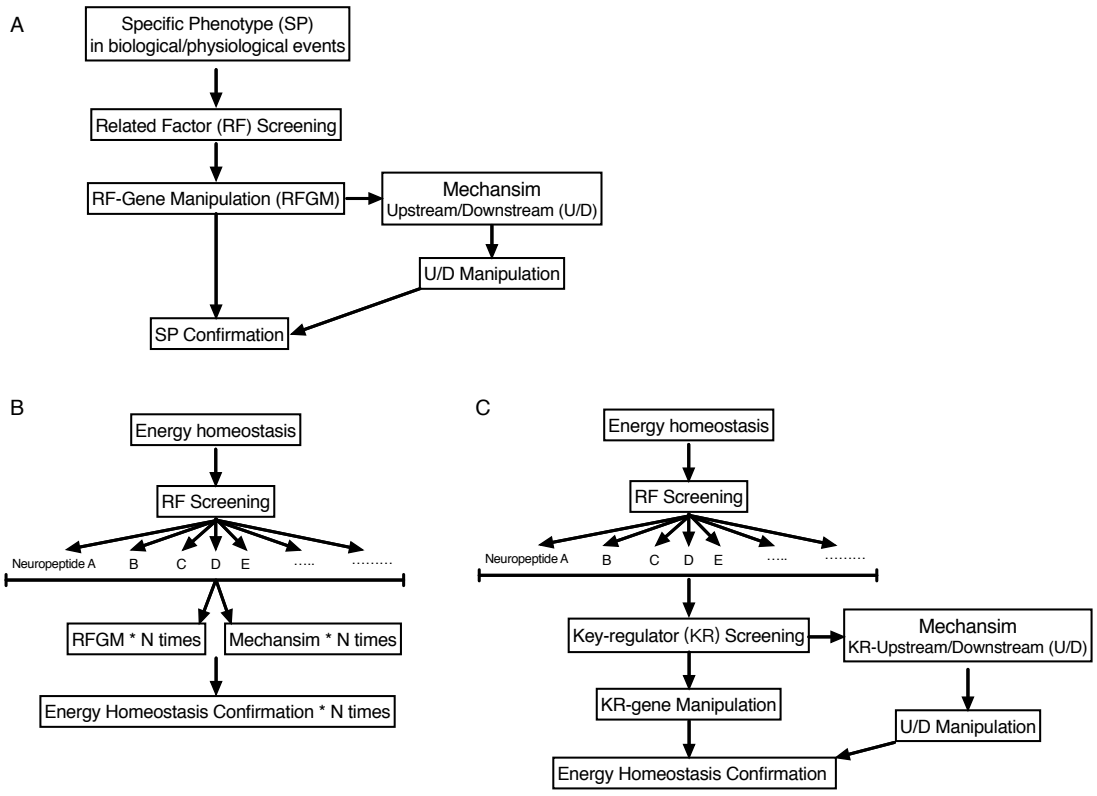

Figure 0-4. Schemes of conventional neuropeptide studies and in this study.

(A) Conventional neuropeptide studies. A specific phenotype is observed, and the endocrinal factors are screened for its regulatory role. Gene manipulation is used both for phenotype confirmation and for understanding its mechanism. The upstream/downstream signals are elucidated and their regulation on the specific phenotype are studied as well. (B) The summarized study scheme based on the classic neuropeptides studies. (C) The practical study scheme in this study. To avoid the unreachable experiments, key regulators screening is employed and manipulated.

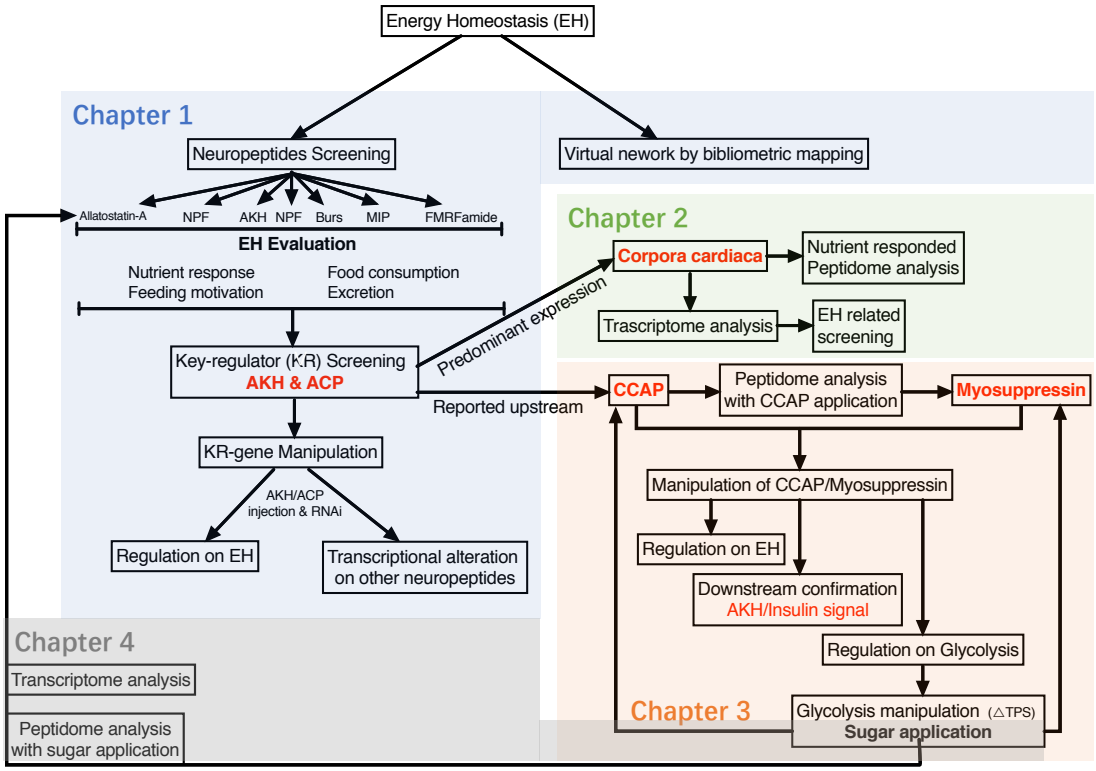

Figure 0-5. Schematic experimental flow in this study.

Red words represent important factors in each chapter.

AKH: Adipokinetic hormone; CCAP: Crustacean cardioactive peptide.

Supplemental figures

A

```
   1 GCCCCCGCCGCCGCCGCTCCCGCCGCGCCCGCGCCCTCCGATCACGCCCCACGCCTCCAC
   1 A  P  A  A  A  A  P  A  A  P  A  P  S  D  H  A  P  R  L  H
  61 AAGAGGCAAGCGGACCCCGCAGAGCTGGAGCGCCTCGCAGCGGAGCCCAAACGCAAGCGG
  21 K  R  Q  A  D  P  A  E  L  E  R  L  A  A  E  P  K  R  K  R
 121 CCCTTCTGCAACGCCTTCACCGGGTGCGGCAAGAAGCGCGCGGACGAGAGCCTGGGCACG
  41 P  F  C  N  A  F  T  G  C  G  K  K  R  A  D  E  S  L  G  T
 181 CTGGTGGAGCTGAACTCGGAGCCCGCCGTGGCGGAGCTCAGCCGGCAGATCCTGTCCGA
  61 L  V  E  L  N  S  E  P  A  V  A  E  L  S  R  Q  I  L  S  E
 241 GCCAAGCTGTGGGAGGCCATCCAGGAGGCACGCGCGGAGCTCCTGCGCCGCCGCCAGCAG
  81 A  K  L  W  E  A  I  Q  E  A  R  A  E  L  L  R  R  R  Q  Q
 301 CAGCTGCAGACGAACCGCATCGCGGCCGACGGCCCGCTGCCGCTGCCGCTCACCAGCTTC
 101 Q  L  Q  T  N  R  I  A  A  D  G  P  L  P  L  P  L  T  S  F
 361 CGCAAGCGGCGCGCCGCCCTCGCCGCCCCCGAGGCCGCCCCCGCCGCCGACGCCGCGCC
 121 R  K  R  R  A  A  L  A  A  P  E  A  A  P  A  A  D  A  A  P
 421 GCCC
 141 A
```

B

```
   1 TGCAGATGGCGCGGGAGGCCCGACTTGTATTCGAGAGGACATTCAAGTGCTTGTAAACGG
  61 TTGTGGAAGATTTAGAATTGACGTGTCAATACACATACCTGCAAGTGGTGTTGGATATTC
 121 ATTATTTCTGGTAGTAGATTTTTTTTCCTGCTGGAATTGAAGAAATGCAATTTTCATCCA
   1                     M  R  R  Q  Q  Q  R  P  P  S  S  A  T  S
 181 TACATTACATCACAAAAAATGCGACGACAGCAGCAGCGGCCTCCGTCGTCGGCGACGTCG
  15 A  A  A  S  A  L  A  T  A  T  P  C  S  S  G  R  W  G  T  A
 241 GCGGCGGCTTCGGCGTTGGCGACGGCCACACCATGCTCAAGTGGACGGTGGGGCACCGCA
  35 R  A  A  A  A  A  D  T  D  A  E  T  A  A  A  A  A  T  A  A
 301 AGGGCGGCGGCAGCAGCCGACACGGACGCCGAGACGGCGGCGGCGGCGGCAACGGCGGCG
  55 T  P  L  L  D  D  A  A  T  V  P  A  P  D  L     D  S  R  P
 361 ACGCCGCTGCTGGACGACGCCGCGACGGTGCCGGCGCCCGACCTCAACGACTCCCGGCCC
  75 D  R  E  A  Q  N  V  T  F     S  F  Y  F  Y  Q  T  E  Q  F
 421 GACCGGGAGGCGCAGAACGTCACTTTCAACTCCTTCTACTTCTACCAGACGGAGCAGTTC
  95 T  V  L  W  I  L  F  A  L  I  V  L  G  N  L  A  V  L  A  A
 481 ACCGTGCTGTGGATCCTGTTCGCGCTGATCGTGCTGGGCAACCTGGCCGTGTTGGCGGCG
 115 L  A  M  N  K  R  R  K  S  R  M  N  F  F  I  M  Q  L  A  L
 541 CTCGCCATGAACAAGCGCCGCAAGTCGCGCATGAACTTCTTCATCATGCAGCTCGCCCTC
 135 A  D  L  S  V  G  L  I  S  V  L  T  D  I  V  W  R  H  T  I
 601 GCAGATTTGTCGGTGGGGCTGATCAGTGTGCTCACCGACATCGTTTGGCGTCATACCATC
 155 A  W  N  A  G  N  A  A  C  K  V  I  R  F  L  Q  A  V  V  T
 661 GCCTGGAATGCAGGAAATGCCGCCTGCAAGGTCATTCGCTTCCTGCAGGCTGTCGTCACC
 175 Y  S  S  T  Y  V  L  V  A  L  S  I  D  R  Y  D  A  I  T  H
 721 TACTCGTCCACCTACGTGCTGGTGGCGCTTAGCATCGATCGCTACGACGCCATCACCCAC
 195 P  M     F  S  G  S  W  R  R  A  R  L  L  V  G  V  A  W  V
 781 CCCATGAACTTCTCTGGCAGCTGGCGGCGCGCGCGGCTGCTGGTGGGAGTGGCGTGGGTG
 215 I  S  A  L  F  S  V  P  I  L  F  L  Y  Q  E  Q  P  V  E  G
 841 ATCAGCGCCCTCTTCTCGGTGCCCATCCTCTTCCTCTACCAGGAGCAGCCCGTTGAAGGT
 235 H  L  Q  C  W  I  D  F  E  K  Q  W  M  W  Q  L  W  V  T  L
 901 CACCTCCAGTGCTGGATCGACTTCGAGAAGCAGTGGATGTGGCAGTTGTGGGTGACGCTG
 255 V  A  L  T  L  F  V  L  P  A  F  I  I  S  A  C  Y  I  V  I
 961 GTGGCCCTGACGCTGTTCGTGCTGCCCGCCTTCATCATCTCGGCGTGCTACATCGTCATC
```

```
 275 V  S  T  I  W  T  K  S  K  Q  L  T  P  D  P  T  R  R  T  S
1021 GTCTCCACCATTTGGACCAAGAGCAAGCAACTCACGCCCGATCCCACGCGCCGCACGTCA
 295 R     S  S  V  Q  Q  L  Q  Q  T  H  Q  Q  L  Q  L  Q  N  S
1081 CGCAATAGTTCTGTCCAGCAACTGCAACAGACACATCAGCAACTACAGCTGCAAAACAGT
 315 G  K  P  R  G  V  A  S  R  L  Q  Q  L  Q  Q  E  D  Q  D  S
1141 GGCAAGCCGCGAGGTGTGGCTTCGCGATTGCAACAACTTCAGCAAGAGGACCAGGACAGT
 335 R  R  A  S  S  R  G  I  I  P  K  A  K  V  K  T  V  K  M  T
1201 CGAAGAGCTTCTTCGCGAGGAATCATTCCAAAAGCAAAAGTCAAGACTGTAAAGATGACC
 355 F  V  I  V  I  V  F  V  L  C  W  S  P  Y  F  V  F  D  L  L
1261 TTTGTAATTGTAATAGTATTTGTACTCTGCTGGAGTCCATATTTCGTATTTGACCTACTT
 375 Q  V  Y  G  Y  V  P  R  T  Q  T  N  I  A  V  A  T  F  I  Q
1321 CAAGTATATGGGTATGTTCCGAGGACACAGACCAACATTGCTGTAGCTACTTTCATTCAG
 395 S  L  A  P  L  N  S  A  A  N  P  V  I  Y  C  L  F  S  T  T
1381 AGTTTGGCTCCCCTAAATTCTGCAGCAAATCCTGTCATCTATTGTCTTTTCTCTACTACA
 415 I  G  R  T  L  R  K  I  P  P  I  S  W  A  V  A  A  L  Q  P
1441 ATTGGTCGAACACTCAGGAAAATCCCACCAATCAGTTGGGCAGTAGCTGCTCTCCAGCCT
 435 C  C  P  G  L  R  P  V  P  P  D  E  E  E  P  A  L  L  D  A
1501 TGCTGTCCCGGTTTGCGCCCTGTCCCTCCTGACGAGGAGGAACCAGCTCTTTTAGACGCA
 455 R  M  R  W  A  S  Q  H  H  K  A  H  P  H  L  D  A  A  V  R
1561 CGCATGAGATGGGCATCTCAACATCACAAAGCGCATCCACATTTGGATGCAGCAGTGAGA
 475 R  G  T  A  V  A  A  V  V  S  I  V  *
1621 AGAGGTACTGCTGTTGCAGCAGTGGTTTCTATTGTGTAAAGAGACAGTGATTTAATACTG
1681 CTCATTTTCCTATAAGAGATTCTCAGAAATGTGGACATGAGTGTATGTCAAAGAATGACT
1741 GACTGGAGGTACCTGCAAGTTATTATACTAGTGCACACACTTTCTCAAATTTGTATTAGT
1801 TCAAGAGTGCAAATTTTTCTAAAACGCTCAATGCAAGGAAGATGATTTACATGCAAAGTT
1861 TATTTCTCAGCTTAAAAGGTCAGTTTTCATAGTTTCTCAACAGATCTGTTCTATTCCAGA
1921 TGAAGCCTCATCTGTAAAAATTTGTTATATCCAGGTAGTGTAACATACAAAGAAATAATA
1981 CAGGACCTTCAAAAATTCTGAATTATCCAACATATGGCAATGTTACCTTTTCTCTATTCA
2041 AAAAAGGTCGAATAAATATGTGGTAAAATACATGCAGGGAAAGCCAGGAGACAT
```

C

```
   1  P  S  P  T  P  A  N  A     G  S  Y     L  S  Y  P  L  L  S
   1 CCCGTCCCCCACCCCCGCCAACGCCAACGGCTCCTACAACCTGAGCTACCCGCTGCTGTC
  21  T  V  A  S  A  S  A  V  P  D  L  N  S  F  Y  F  Y  E  T  E
  61 CACCGTTGCCTCCGCCTCCGCAGTCCCGGACCTCAACAGCTTCTACTTCTACGAGACTGA
  41  Q  F  A  V  L  W  I  L  F  T  L  I  V  L  G  N  S  A  V  L
 121 GCAGTTTGCCGTGCTGTGGATCCTCTTCACCCTGATCGTGTTGGGTAACTCTGCTGTGCT
  61  V  A  L  F  L  N  K  N  R  K  S  R  M  N  F  F  I  M  Q  L
 181 TGTCGCACTTTTCCTCAACAAAAACAGAAAATCGCGAATGAACTTCTTCATCATGCAGCT
  81  A  L  A  D  L  T  V  G  L  I  S  V  L  T  D  I  V  W  R  I
 241 CGCACTAGCCGATCTGACCGTGGGGCTGATCAGCGTGCTGACGGACATCGTGTGGCGCAT
 101  T  V  A  W  H  A  G  N  A  A  C  K  I  I  R  Y  L  Q  A  V
 301 CACCGTGGCGTGGCACGCGGGCAACGCAGCCTGCAAGATCATACGCTATCTGCAGGCCGT
 121  V  T  Y  S  S  T  Y  V  L  V  A  L  S  I  D  R  Y  D  A  I
 361 GGTCACCTACTCGTCCACGTACGTGCTGGTGGCGCTGAGCATCGACCGCTACGACGCCAT
 141  T  H  P  M     F  S  G  S  W  R  R  A  R  V  L  V  A  V  A
 421 CACGCACCCCATGAACTTCTCCGGCAGCTGGCGGCGGGCGCGCGTGCTGGTGGCTGTCGC
 161  W  F  L  S  A  I  F  S  I  P  I  I  V  F  Y  Q  E  Q  S  V
 481 CTGGTTCCTCAGCGCCATCTTTTCCATCCCGATCATCGTCTTCTACCAGGAGCAGTCGGT
 181  Q  G  H  Q  Q  C  W  I  D  F  P  E  Q  W  Q  W  Q  L  Y  M
 541 GCAAGGGCACCAGCAGTGCTGGATCGACTTCCCCGAGCAGTGGCAGTGGCAGCTATACAT
```

```
 201  T  L  V  A  V  L  L  F  V  F  P  A  I  I  I  S  F  C  Y  T
 601 GACGCTGGTGGCCGTGTTGCTGTTCGTGTTCCCGGCCATCATCATCTCCTTCTGCTACAC
 221  V  I  V  F  T  I  W  S  K  S  K  L  L  T  P  S  G  R  P  R
 661 CGTCATCGTCTTCACCATCTGGAGCAAGAGCAAGCTGCTCACCCCGTCCGGTCGCCCGCG
 241  G  A  T  R  N  S  E  R  R  A  L  R  Q  D  D  Q  D  S  R  R
 721 CGGGGCTACGCGCAACAGCGAGCGGCGCGCGCTGCGCCAGGACGACCAGGACAGCCGCCG
 261  A  S  S  R  G  L  I  P  R  A  K  I  K  T  V  K  M  T  F  V
 781 CGCGTCGTCGCGCGGACTTATCCCGCGCGCCAAGATCAAGACGGTCAAGATGACCTTCGT
 281  I  V  F  V  F  I  L  C  W  S  P  Y  F  V  F  D  L  L  Q  V
 841 CATCGTCTTCGTGTTCATCCTGTGCTGGAGCCCCTACTTCGTGTTCGACCTGCTGCAGGT
 301  Y  G  H  V  P  S  T  Q  T  N  I  A  V  A  T  F  I  Q  S  L
 901 GTACGGCCACGTGCCGTCCACGCAGACCAACATCGCCGTGGCCACTTTCATCCAGAGCCT
 321  A  P  L  N  S  A  A  N  P  L  I  Y  C  L  F  S  T  H  I  C
 961 GGCGCCGCTCAACTCGGCCGCCAACCCGCTCATCTACTGCCTCTTCTCCACGCACATCTG
 341  R  T  L  R  K  I  P  P  F  S  W  V  A  A  G  L  S  L  C  F
1021 CCGCACGCTCAGGAAGATCCCGCCCTTCAGCTGGGTGGCGGCCGGGCTGTCGTTGTGCTT
 361  P  A  L  R  S  G  D  G  R  C  L  R  Y  A  T  G  D  S  S  S
1081 CCCGGCGCTGCGTTCGGGTGACGGGCGCTGCTTGCGCTACGCCACCGGCGACTCCAGCTC
 381  T  V  T  E  T  L  T  Q  Q  S  S  R  R  S  T  S  L  R  H  T
1141 CACCGTCACCGAGACGCTCACGCAACAGTCCTCGCGCCGCTCCACCTCGCTGCGCCACAC
 401  M  Q  V  R  L  P  V  S  G  A  A  V  G  A  C  G  A  T  S  S
1201 CATGCAGGTGCGCCTGCCTGTGTCGGGCGCGGCGGTGGGCGCGTGCGGGGCGACCAGCTC
 421  H  R  R  K  V  A  V  S  V  V  *
1261 GCACCGCCGCAAGGTGGCCGTGTCTGTGGTGTAGGCGCCCGGGCCCAGCCTCCGCCCGCA
1321 CCGCGACGCCGCCTGGTGAAGGAAGTGGCGAGCGGCATGCACCGGCGCTGTGCACCGCTT
1381 TCTGCTCTACACTTCAACGCCGAGTTCATCCGGAATGGATTCACGCAGCTTCACCCTCCA
1441 GTGGTCATAGGTTGGATCACCCTCTGCGGTTTCGCAGCCGGGCAGATGGGAATGCGAGCG
1501 GTTTGAATCGGAGCTGTGCACCACTTGGTCCTCTATTCCTCAAGGCAGTGTTCAGCCAAA
1561 ATGGATTCATGGATCTGCTTCGCCCTCAAATGGTCTTAGTGTGGAATTGTATTGTTGTTT
1621 AATGTAAGAATCGGTTATCAAGTGGTTTCGGAGTACGAAGTGTCTATGACTTCTCACAAT
1681 GCAAGGGAAAGTTAGTTCTTTAAGATTTGGAAAATATAATATCCATAGTAAAAACCTTCA
1741 ATGTGGTTATTGGGTCTATTTAGTGATGTATGAAGTCTGACCTGTGACTCTGAACTGTGA
1801 CCTCAATTAAACCACCAATATCATTTGCGATCTGTGGATATTTTGAAAACTTGAATTAGT
1861 AAGTTCAATTTCTGCACGAAGCAGACCTATTGTATCATTGTATATGTGTGTGAAAACCCT
1921 AACGCAACACAGTGAACCGCAAGC
```

Figure S1 cDNAs and deduced amino acid sequences of GbCCAP (A), GbCCAPR-1 (B) and GbCCAPR-2 (C).

The mature sequence of GbCCAP is shallowly shaded in grey. The cleavage site is wavy underlined. The transmembrane areas of receptors are darkly shaded in grey. The putative glycosylation sites in receptors are shaded in black.

A

```
    1       M  T  K  G  Q  E  L  A  V  V  T  E  E  L  L  H  G  K
    1 CGCCGCCATGACGAAGGGCCAGGAGCTGGCGGTCGTCACCGAGGAGCTGCTCCACGGGAA
  101  P  K  Q  S  F  G  S  Q  P  L  Q  P  Q  P  L  Q  G  L  G  R
  301 GCCGAAGCAGAGCTTCGGCTCGCAGCCGCTGCAGCCGCAGCCACTGCAGGGCCTCGGCCG
  121  Q  M  F  A  S  V  A  V  A  L  G  S  M  L  V  G  F  S  S  A
  361 ACAGATGTTCGCGTCGGTGGCGGTGGCGTTGGGCTCGATGCTCGTGGGCTTCTCTTCCGC
  141  Y  T  S  P  A  L  V  S  M  A  E  P  N  S  T  L  S  V  T  P
  421 CTACACGTCGCCCGCGCTCGTGTCCATGGCGGAGCCCAACAGCACGCTCTCAGTCACCCC
  161  S  Q  G  S  W  V  G  S  L  M  P  A  A  A  L  F  G  G  M  A
  481 GTCGCAGGGCTCGTGGGTGGGCAGCCTGATGCCGGCAGCGGCGCTGTTCGGAGGGATGGC
  181  G  G  F  L  I  E  W  L  G  R  K  T  T  I  L  A  T  A  P  P
  541 GGGCGGTTTCCTCATCGAGTGGCTGGGCCGCAAGACTACCATCCTCGCCACCGCGCCGCC
  201  S  A  I  A  W  L  L  I  F  G  A  S  N  V  G  M  L  Y  A  G
  601 GTCCGCCATTGCGTGGCTTCTCATCTTCGGCGCGAGCAACGTGGGCATGCTGTACGCGGG
  221  R  I  L  S  G  F  C  V  G  V  T  S  L  S  L  P  V  Y  M  G
  661 CCGCATCCTCTCGGGCTTCTGCGTGGGCGTCACGTCCCTCAGCCTGCCCGTCTACATGGG
  241  E  T  I  Q  P  E  V  R  G  M  L  G  L  I  S  T  T  L  G  N
  721 CGAGACCATTCAGCCCGAGGTGCGAGGCATGCTTGGCCTCATCTCCACAACACTAGGCAA
  261  I  G  I  L  L  C  F  V  L  G  K  F  L  D  W  R  Y  L  A  M
  781 CATCGGCATTCTGCTGTGCTTCGTCCTGGGCAAGTTCCTCGACTGGCGCTACCTGGCCAT
  281  M  G  A  L  C  S  L  P  Y  G  L  C  T  M  F  A  P  E  T  P
  841 GATGGGCGCGCTGTGCTCGCTGCCGTACGGGCTCTGCACCATGTTCGCCCCCGAAACGCC
  301  H  W  Y  L  S  K  G  R  R  G  D  A  E  T  A  L  R  W  L  R
  901 GCACTGGTACCTCTCCAAAGGGCGGCGGGGCGACGCGGAGACGGCGCTGCGCTGGCTGCG
  321  G  G  A  G  A  D  V  A  Q  E  L  R  A  T  E  Q  G  L  G  K
  961 CGGCGGCGCGGGCGCCGACGTGGCGCAGGAGCTGCGCGCCACCGAGCAGGGCCTGGGCAA
  341  G  A  E  A  A  I  G  L  P  R  S  V  T  S  I  N  L  Q  A  Q
 1021 GGGCGCCGAGGCCGCCATCGGCCTGCCGCGCAGCGTCACCAGCATCAACCTGCAGGCGCA
  361  A  A  P  S  D  K  S  G  L  A  L  L  F  T  R  D  N  M  R  P
 1081 GGCCGCGCCCTCCGACAAGAGCGGCCTGGCGCTGCTCTTCACGCGCGACAACATGCGGCC
  381  V  G  V  A  L  G  L  M  L  F  Q  Q  F  S  G  I  N  A  V  I
 1141 CGTCGGCGTCGCGCTGGGCCTCATGCTCTTCCAGCAGTTCAGCGGCATCAACGCCGTCAT
  401  F  Y  T  V  K  I  F  Q  D  A  G  S  T  I  D  E  F  V  S  T
 1201 CTTCTACACGGTCAAGATCTTCCAGGACGCGGGCAGCACCATTGACGAGTTCGTGAGCAC
  421  M  V  V  G  A  V  N  L  G  A  V  V  L  A  T  A  L  I  D  R
 1261 GATGGTGGTGGGCGCGGTGAACCTGGGCGCCGTGGTGCTGGCGACGGCGCTCATCGACCG
  441  L  G  R  R  V  L  L  A  A  S  A  A  A  M  A  L  A  L  A  T
 1321 CCTGGGGCGGCGGGTGCTGCTGGCGGCGTCGGCGGCCGCCATGGCGCTGGCGCTCGCCAC
  461  L  G  T  F  F  Y  L  S  T  C  P  D  V  D  V  S  G  F  G  W
 1381 GCTGGGCACCTTCTTCTACCTCAGCACCTGCCCCGACGTCGACGTCAGCGGCTTCGGCTG
  481  L  P  L  V  S  F  V  I  F  V  V  G  F  S  V  G  F  G  P  I
 1441 GCTGCCGCTGGTCAGCTTCGTCATCTTCGTGGTGGGCTTCTCCGTGGGCTTCGGGCCCAT
```

```
 501  P  W  L  M  M  G  E  I  L  P  P  D  I  R  G  P  A  A  S  L
1501 CCCCTGGCTCATGATGGGCGAAATTTTACCACCGGACATCCGCGGGCCGGCGGCGTCGCT
 521  A  T  A  V  N  W  S  C  T  F  M  V  T  K  M  F  A  D  V  V
1561 GGCGACGGCCGTCAACTGGTCGTGCACGTTCATGGTGACGAAGATGTTCGCGGACGTGGT
 541  E  G  V  G  S  H  A  A  F  W  G  F  C  A  V  C  V  G  A  F
1621 GGAGGGCGTGGGCAGCCACGCGGCCTTCTGGGGCTTCTGCGCCGTGTGCGTGGGCGCCTT
 561  V  F  V  W  A  C  V  P  E  T  R  G  R  T  L  A  D  I  Q  R
1681 CGTGTTCGTGTGGGCGTGCGTGCCCGAGACGCGCGGCCGCACGCTCGCAGACATCCAGCG
 581  R  M  A  G  R  P  A  P  Q  R  K  R  R  L  T  V  S  E  S  V
1741 GCGCATGGCCGGGCGGCCGGCCCCGCAGCGCAAGCGCCGCCTCACCGTCAGCGAGTCGGT
 601  A  G  W  T  A  S  M  A  S  L  K  P  M  P  T  G  A  *
```

B

```
   1  M  A  D  V  A  S  S  T  G  S  M  I  V  V  S  N  R  L  P  F
   1 ATGGCTGATGTAGCGTCGTCCACTGGCAGTATGATCGTAGTATCGAACAGATTACCTTTC
  21  V  L  K  R  N  A  L  T  Q  L  L  E  R  K  A  S  A  G  G  L
  61 GTTTTAAAGCGAAATGCACTGACACAATTATTGGAACGAAAGGCCAGCGCCGGCGGTCTT
  41  V  T  A  V  A  P  V  V  I  Q  S  G  G  L  W  V  G  W  P  G
 121 GTAACTGCTGTAGCACCTGTAGTAATACAGAGTGGCGGATTGTGGGTGGGCTGGCCAGGT
  61  I  H  L  N  N  P  N  E  P  I  P  E  S  D  P  N  D  K  T  P
 181 ATTCACTTGAATAATCCAAATGAGCCCATACCTGAATCAGACCCAAATGATAAAACACCC
  81  T  A  G  L  L  S  K  K  V  V  S  V  H  I  N  A  D  V  F  D
 241 ACTGCTGGACTTCTGTCAAAGAAGGTTGTTTCAGTTCATATAAACGCTGATGTGTTTGAT
 101  S  Y  Y  N  G  C  C  N  G  T  F  W  P  L  F  H  S  M  P  D
 301 TCATATTACAATGGTTGTTGTAATGGTACATTTTGGCCCTTGTTTCACTCAATGCCAGAT
 121  R  A  V  F  S  A  D  N  W  K  C  Y  Y  D  V  N  Q  L  F  A
 361 AGAGCTGTCTTTTCTGCAGATAATTGGAAGTGCTATTATGATGTGAATCAATTGTTTGCG
 141  D  K  T  I  E  A  F  D  K  L  I  A  E  Q  G  N  S  A  G  T
 421 GATAAAACCATTGAAGCATTTGACAAGCTGATAGCTGAACAGGGAAACAGTGCCGGAACT
 161  P  L  I  W  I  H  D  Y  H  L  M  L  A  A  N  W  V  R  Q  V
 481 CCGTTAATATGGATTCATGATTACCATTTGATGTTGGCTGCAAATTGGGTTAGACAGGTG
 181  A  E  E  R  D  L  R  C  K  L  G  F  F  L  H  I  P  F  P  P
 541 GCTGAAGAAAGGGATCTTCGTTGCAAATTGGGTTTCTTTTTACATATTCCTTTTCCCCCA
 201  W  D  I  F  R  L  F  P  W  A  D  E  I  L  Q  G  M  L  G  N
 601 TGGGATATTTTTCGCCTTTTCCCATGGGCTGATGAAATTCTGCAAGGAATGCTGGGTAAT
 221  D  M  V  G  F  H  I  E  D  Y  C  L  N  F  V  D  C  C  Q  R
 661 GACATGGTTGGTTTTCATATTGAAGACTATTGCCTGAACTTTGTGGATTGTTGTCAGCGA
 241  R  L  G  C  R  V  D  R  K  A  L  L  V  E  H  G  G  R  T  V
 721 AGGCTAGGCTGCCGAGTGGATCGTAAAGCTCTTTTAGTAGAACATGGTGGCCGCACAGTC
 261  R  V  R  P  L  P  I  G  I  P  Y  D  R  F  V  Q  L  A  S  N
 781 AGGGTTCGACCTCTCCCCATAGGCATCCCATATGATCGCTTTGTTCAACTTGCAAGTAAT
 281  A  P  K  V  M  L  T  N  Q  K  I  I  L  G  V  D  R  L  D  Y
 841 GCTCCCAAGGTAATGCTGACCAATCAGAAGATTATCCTTGGTGTTGATCGCTTAGATTAC
```

```
 301 T  K  G  L  V  H  R  L  K  A  I  E  T  L  F  E  K  Y  P  E
 901 ACAAAAGGCCTTGTTCACAGATTAAAGGCAATAGAGACTCTATTTGAAAAATATCCTGAA
 321 H  I  E  K  A  T  M  L  Q  I  S  V  P  S  R  T  D  V  K  E
 961 CACATTGAAAAGGCAACAATGTTGCAAATATCCGTCCCTTCACGCACAGATGTGAAAGAG
 341 Y  Q  D  L  K  E  E  M  D  Q  L  V  G  R  I  N  G  R  F  T
1021 TATCAAGATCTGAAGGAAGAAATGGATCAACTTGTGGGACGCATCAATGGGCGTTTCACA
 361 T  Y  N  W  S  P  I  R  Y  I  Y  G  C  V  S  Q  D  E  L  A
1081 ACTTACAACTGGTCTCCTATTCGTTATATATATGGTTGTGTCAGCCAAGATGAATTGGCA
 381 A  F  Y  R  D  A  A  V  A  L  V  T  P  L  R  D  G  M  N  L
1141 GCTTTTTACCGGGATGCAGCTGTAGCTCTTGTCACACCCTTGAGAGATGGCATGAACCTG
 401 V  A  K  E  F  V  A  C  Q  I  N  E  P  P  G  V  L  I  V  S
1201 GTGGCCAAGGAGTTCGTGGCGTGCCAGATCAACGAGCCGCCCGGCGTGCTCATCGTGTCG
 421 P  F  A  G  A  G  E  M  M  H  E  A  L  I  C  N  P  Y  E  I
1261 CCCTTCGCCGGCGCGGGCGAGATGATGCACGAGGCGCTCATCTGCAACCCCTACGAGATC
 441 N  A  A  A  E  V  I  H  R  A  L  T  M  P  E  D  E  R  T  L
1321 AACGCCGCCGCCGAGGTCATCCACAGGGCTCTTACAATGCCAGAAGATGAAAGGACCTTG
 461 R  M  N  Y  L  R  R  R  E  K  L  H  D  V  N  Y  W  M  R  S
1381 CGAATGAACTACCTGAGAAGGCGAGAGAAGCTTCATGATGTGAACTACTGGATGAGGTCT
 481 F  L  K  A  M  G  S  L  I  E  E  D  G  D  E  V  R  P  T  T
1441 TTCCTCAAAGCGATGGGTTCTCTCATTGAGGAAGATGGAGACGAAGTCAGGCCCACTACG
 501 M  Q  P  V  T  M  D  D  F  D  E  Y  L  A  K  Y  I  G  H  T
1501 ATGCAACCTGTCACAATGGATGACTTCGATGAGTACCTTGCTAAATACATTGGGCACACA
 521 H  K  L  A  L  L  L  D  Y  D  G  T  L  A  P  I  A  P  H  P
1561 CATAAATTAGCTCTACTGTTGGATTATGATGGCACCCTGGCACCTATTGCCCCACATCCT
 541 D  L  A  I  M  P  E  E  T  K  N  V  L  E  R  L  S  N  M  P
1621 GATCTGGCTATCATGCCTGAAGAAACCAAGAATGTTCTTGAGAGATTGTCCAATATGCCA
 561 D  V  Y  I  A  I  I  S  G  R  N  V  N  N  V  K  S  M  V  G
1681 GATGTGTATATTGCTATCATATCTGGACGAAATGTCAACAATGTGAAATCCATGGTTGGG
 581 I  E  G  I  T  Y  A  G  S  H  G  L  E  I  L  H  P  D  G  S
1741 ATTGAAGGTATTACGTATGCTGGTAGTCATGGTTTGGAAATCCTTCATCCTGATGGAAGC
 601 K  F  V  H  P  M  P  V  E  F  E  D  K  V  T  E  L  L  K  A
1801 AAGTTTGTACATCCCATGCCTGTGGAATTTGAAGACAAAGTTACTGAATTACTAAAAGCT
 621 L  Q  D  Q  V  C  K  D  G  A  W  V  E  N  K  G  A  L  L  T
1861 CTGCAAGACCAAGTGTGCAAGGACGGCGCGTGGGTGGAGAACAAGGGGGCGCTGCTGACG
 641 F  H  Y  R  E  T  P  V  Q  K  R  E  A  L  V  S  K  A  Q  Q
1921 TTCCACTACCGCGAGACGCCCGTGCAGAAGCGCGAGGCGCTGGTGAGCAAGGCGCAGCAG
 661 M  I  V  A  A  G  F  K  A  G  A  A  H  C  A  L  E  A  K  P
1981 ATGATCGTGGCGGCGGGCTTCAAGGCCGGCGCGGCGCACTGCGCCCTCGAGGCGAAGCCG
 681 P  V  Q  W  N  K  G  R  A  S  I  Y  I  L  R  T  A  F  G  V
2041 CCCGTGCAGTGGAACAAGGGCCGCGCCTCCATCTACATCCTGCGCACGGCGTTCGGCGTC
 701 D  W  S  E  R  I  R  I  I  Y  A  G  D  D  V  T  D  E  D  A
2101 GACTGGAGCGAGCGCATCCGCATCATCTACGCCGGCGACGACGTCACCGACGAGGACGCC
 721 M  E  A  L  K  G  M  A  A  T  F  R  V  T  S  S  H  I  V  K
```

```
2161 ATGGAGGCCTTAAAAGGTATGGCAGCCACTTTTCGTGTCACTTCATCGCATATCGTCAAG
 741 T  A  A  E  R  R  L  P  S  T  D  S  V  L  T  M  L  K  W  V
2221 ACTGCAGCAGAGAGAAGACTGCCTAGTACTGATTCAGTGCTAACCATGTTGAAATGGGTG
 761 E  R  H  F  S  R  R  E  V  R  R  D  P  E  S  N  L  L  Y  R
2281 GAACGACACTTTTCTAGGCGTGAAGTACGGCGGGACCCTGAGTCAAATCTCCTGTACAGA
 781 R  Q  S  S  S  K  T  N  G  G  L  K  T  E  M  S  Y  T  P  K
2341 AGGCAATCCAGTAGTAAAACAAATGGTGGCTTGAAAACAGAAATGTCTTACACACCAAAG
 801 A  V  T  P  E  P  L  *
2401 GCTGTGACACCAGAACCATTGTAG
```

C

```
   1 TTGACATTTTAAAGTGCCTCCACTCAGCATTCTCGGAACTTGTCGAAAGCGTTCGGGAAA
   1          M  D  P  W  M  I  G  F  Y  I  V  S  V  L  V  V  V
  61 ATCTTGGACATGGATCCGTGGATGATAGGATTTTACATTGTGTCCGTGCTGGTGGTCGTA
  18 N  T  E  K  L  P  P  P  C  E  S  E  I  Y  C  H  G  D  L  L
 121 AATACAGAAAAATTACCTCCTCCGTGCGAAAGTGAAATATATTGTCATGGAGACCTGCTA
  38 H  Y  V  Q  L  G  S  G  I  R  D  S  K  E  F  V  D  M  K  M
 181 CATTACGTGCAGTTGGGAAGTGGAATACGAGATTCTAAAGAATTTGTAGACATGAAAATG
  58 K  Y  G  P  S  R  T  L  E  I  F  N  E  M  M  N  E  T  Q  G
 241 AAATATGGACCTTCCAGGACCTTAGAAATATTCAATGAAATGATGAATGAAACACAAGGA
  78 K  P  N  K  Q  D  I  L  N  F  L  A  K  A  F  E  E  P  G  S
 301 AAACCTAATAAGCAAGACATTTTGAATTTCCTTGCAAAAGCATTTGAAGAACCAGGGAGC
  98 E  F  E  F  W  D  P  V  D  W  I  P  N  P  K  F  L  N  K  I
 361 GAATTTGAATTCTGGGATCCAGTCGATTGGATACCAAACCCCAAATTTTTAAATAAAATT
 118 Q  D  P  D  F  R  R  F  A  E  H  L  N  S  L  W  K  S  L  G
 421 CAAGATCCAGACTTCCGTCGATTCGCAGAACATTTAAATAGTCTCTGGAAATCATTGGGC
 138 R  K  M  K  D  A  V  K  Q  S  P  E  F  Y  S  I  I  Y  V  P
 481 CGTAAAATGAAAGATGCAGTTAAACAAAGTCCTGAATTCTATTCCATTATCTATGTGCCT
 158 H  P  V  I  V  P  G  G  R  F  R  E  F  Y  Y  W  D  S  F  W
 541 CATCCTGTAATTGTACCTGGAGGTCGGTTTAGAGAATTTTACTATTGGGATTCATTTTGG
 178 I  I  R  G  L  L  L  S  E  M  Q  S  T  V  K  G  M  L  E  N
 601 ATTATCAGAGGTTTACTTTTGTCAGAAATGCAGTCAACTGTCAAGGGAATGCTAGAGAAC
 198 F  L  S  I  V  E  R  F  G  F  I  P  N  G  G  R  I  Y  Y  A
 661 TTTTTATCCATTGTTGAAAGGTTTGGATTTATACCCAATGGAGGTCGTATATATTACGCA
 218 M  R  S  Q  P  P  F  L  I  P  M  V  Q  T  Y  V  D  T  T  Q
 721 ATGCGTTCACAACCACCATTCCTCATACCGATGGTTCAAACTTATGTAGATACCACTCAG
 238 D  Y  E  F  I  E  K  N  I  D  T  L  D  K  E  F  Q  H  W  M
 781 GATTATGAGTTTATTGAAAAGAATATAGATACTCTGGATAAAGAATTTCAACATTGGATG
 258 V  N  H  T  V  N  I  K  K  D  G  K  V  Y  T  L  A  R  Y  K
 841 GTGAATCACACAGTTAACATTAAAAAAGATGGAAAAGTGTACACTCTAGCAAGATACAAG
 278 D  N  S  N  G  P  R  P  E  S  Y  K  E  D  Y  E  S  A  H  F
 901 GACAATTCTAATGGCCCAAGACCCGAGTCATACAAAGAAGATTATGAGAGTGCGCACTTC
```

```
 298 F  H  T  E  E  D  K  E  N  Y  Y  S  E  L  K  A  A  A  E  S
 961 TTTCACACTGAAGAAGACAAAGAAAATTACTACTCTGAACTTAAAGCAGCTGCTGAATCT
 318 G  W  D  F  S  S  R  W  F  V  L  N  G  T  N  K  G  N  L  T
1021 GGCTGGGACTTTTCGAGTAGATGGTTTGTTCTGAATGGCACTAACAAAGGCAACCTGACA
 338 N  L  K  V  R  S  I  I  P  V  E  L  N  A  V  L  Y  G  N  A
1081 AATTTGAAAGTCAGGTCCATAATTCCTGTGGAACTTAATGCAGTTCTATATGGAAATGCC
 358 K  T  L  A  N  Y  Y  S  R  F  K  D  F  N  K  S  E  M  Y  N
1141 AAAACATTGGCCAATTATTATTCAAGATTTAAAGACTTTAACAAATCTGAAATGTATAAT
 378 D  I  A  E  K  L  K  E  A  V  T  A  V  L  W  H  K  E  V  G
1201 GACATTGCTGAGAAGTTAAAAGAAGCTGTGACAGCTGTTCTGTGGCACAAAGAAGTTGGT
 398 A  W  L  D  Y  D  M  I  N  D  K  R  R  D  Y  F  Y  P  T  N
1261 GCTTGGTTAGATTATGATATGATAAATGATAAACGACGTGATTATTTCTATCCAACCAAT
 418 I  S  P  L  W  T  G  C  Y  D  T  T  H  K  E  Q  Y  V  G  R
1321 ATTTCTCCTTTATGGACTGGCTGCTATGATACAACACACAAAGAGCAATATGTAGGACGT
 438 V  L  K  Y  L  E  H  S  R  A  T  V  L  L  G  G  I  P  T  T
1381 GTACTCAAATACTTGGAGCATAGTCGTGCCACAGTACTCCTAGGAGGAATCCCAACAACA
 458 M  E  L  S  N  E  Q  W  D  Y  P  N  A  W  A  P  L  Q  H  L
1441 ATGGAACTGTCAAATGAACAGTGGGACTACCCTAATGCATGGGCTCCATTGCAGCATCTG
 478 F  V  E  S  L  D  A  T  G  D  P  W  A  K  E  L  A  Y  Q  I
1501 TTTGTGGAAAGCCTGGATGCTACAGGTGATCCATGGGCAAAAGAATTGGCATATCAAATT
 498 A  Q  K  W  V  R  S  N  F  K  A  W  N  E  T  G  N  M  Y  E
1561 GCACAGAAGTGGGTACGCTCCAACTTTAAAGCTTGGAATGAAACAGGAAACATGTATGAA
 518 K  Y  D  A  T  K  L  G  G  H  G  E  G  G  E  Y  R  V  Q  L
1621 AAGTATGACGCAACCAAACTAGGTGGTCATGGAGAAGGTGGTGAATATCGAGTTCAGTTG
 538 G  F  G  W  S  N  G  V  V  M  D  F  L  D  K  Y  G  H  N  I
1681 GGATTTGGATGGTCCAATGGGGTTGTCATGGATTTCTTGGATAAATATGGTCACAACATA
 558 T  L  R  D  K  F  E  V  V  Q  T  S  S  V  A  T  I  L  S  A
1741 ACTCTCAGAGACAAATTTGAAGTGGTTCAGACCTCCAGTGTGGCAACAATCTTGTCAGCA
 578 T  K  E  S  Q  I  L  T  A  I  Y  A  L  L  A  T  L  A  A  G
1801 ACCAAAGAAAGTCAGATCCTGACAGCTATTTATGCTCTACTTGCCACACTTGCGGCAGGA
 598 S  I  G  *
```

D

```
   1 Q  G  E  L  L  S  V  V  Q  N  A  H  L  Y  P  D  S  K  K  F
   1 CAAGGCGAGCTCCTCAGTGTCGTCCAGAACGCTCACTTGTATCCCGACTCGAAGAAATTC
  21 V  D  K  R  L  I  Y  S  P  D  E  T  L  E  Q  F  D  A  L  M
  61 GTCGACAAGCGGCTCATCTACTCGCCGGACGAGACGCTGGAGCAGTTCGACGCGTTGATG
  41 K  Q  T  G  G  N  P  S  K  E  Q  I  Q  K  F  V  D  D  Y  F
 121 AAGCAAACCGGAGGCAACCCGTCCAAGGAGCAAATCCAAAAGTTCGTAGACGATTACTTC
  61 E  D  G  N  E  L  E  G  A  D  L  P  D  W  V  A  E  P  A  L
 181 GAGGACGGCAACGAGCTGGAGGGGGCGGACCTGCCCGACTGGGTGGCCGAGCCGGCGCTG
  81 L  R  R  I  A  D  P  K  L  R  A  W  A  S  E  L  N  G  I  W
```

```
 241 CTCCGCCGTATCGCCGACCCCAAGCTGCGCGCCTGGGCGTCCGAGCTGAACGGCATATGG
 101 K  N  L  S  R  R  V  S  D  E  V  R  D  H  T  D  R  Y  S  L
 301 AAGAATCTCAGCAGAAGGGTTTCCGACGAAGTCAGGGACCACACGGACAGATACTCGCTC
 121 I  Y  V  P  H  H  F  I  V  P  G  G  R  F  R  E  L  Y  Y  W
 361 ATCTACGTTCCTCACCACTTCATCGTGCCAGGCGGGCGCTTTAGGGAGCTGTACTATTGG
 141 D  T  Y  W  I  V  R  G  L  L  L  S  D  M  K  D  T  V  K  G
 421 GATACCTACTGGATTGTTCGAGGGCTTCTGCTTAGCGATATGAAGGACACTGTAAAGGGT
 161 I  I  E  N  F  L  F  L  L  D  T  H  G  I  I  P  N  G  A  R
 481 ATAATAGAAAATTTCTTGTTCCTGCTGGACACCCATGGAATAATTCCTAATGGAGCAAGA
 181 I  Y  Y  L  E  R  S  Q  P  P  L  L  T  P  M  V  Y  N  Y  F
 541 ATTTACTACTTGGAACGCTCGCAACCCCCGCTTCTGACTCCCATGGTCTACAATTACTTT
 201 L  A  T  K  D  E  E  F  V  K  K  N  I  H  L  L  E  R  E  L
 601 CTGGCTACCAAGGATGAAGAATTTGTGAAGAAAAACATCCACTTGCTTGAAAGAGAATTA
 221 D  F  W  M  K  N  R  T  V  N  V  H  K  D  G  K  T  Y  T  L
 661 GATTTCTGGATGAAAAATCGTACAGTAAACGTGCACAAGGATGGAAAAACCTACACTCTT
 241 L  R  Y  Y  S  P  S  Q  G  P  R  P  E  S  Y  R  E  D  Y  M
 721 CTCAGATACTATTCTCCATCGCAGGGCCCTAGACCAGAATCATACAGGGAGGATTACATG
 261 N  G  Q  V  F  S  S  D  Q  R  K  E  E  F  Y  I  D  L  K  S
 781 AACGGCCAAGTATTCTCTTCTGATCAAAGAAAAGAAGAATTTTACATAGACTTGAAATCT
 281 A  A  E  S  G  W  D  F  S  S  R  W  F  I  K  D  G  S  N  K
 841 GCTGCAGAATCTGGTTGGGATTTCTCCAGTCGATGGTTCATCAAAGATGGTAGCAACAAA
 301 G  N  L  T  D  I  H  T  K  Y  I  I  P  V  D  V  N  A  F  V
 901 GGAAACTTGACAGACATTCATACAAAATATATAATTCCAGTAGATGTCAATGCTTTTGTG
 321 Y  G  N  A  R  L  L  S  T  L  H  T  E  V  V  G  N  L  D  K
 961 TATGGTAATGCTCGTTTGCTTTCTACATTGCACACGGAAGTTGTAGGAAACTTGGATAAA
 341 A  K  K  Y  S  Q  I  A  H  E  L  K  E  A  V  T  A  V  L  W
1021 GCCAAAAAATATTCCCAAATTGCCCATGAACTGAAGGAGGCAGTCACAGCAGTATTATGG
 361 N  G  T  T  G  S  W  F  D  Y  D  L  L  N  A  K  Q  R  T  Y
1081 AATGGTACAACTGGTTCTTGGTTTGATTATGATCTTCTGAATGCAAAACAGCGCACCTAC
 381 F  Y  P  S  N  L  A  P  L  W  T  K  C  Y  N  E  K  E  T  P
1141 TTTTACCCATCAAACCTTGCTCCTTTATGGACAAAGTGTTACAATGAGAAAGAGACTCCC
 401 I  I  A  Q  K  V  V  K  Y  L  Q  S  I  E  I  L  Q  H  Y  L
1201 ATCATTGCACAAAAAGTAGTTAAATACTTACAAAGTATTGAAATTTTACAACACTATCTT
 421 G  G  I  P  T  S  L  N  N  T  G  E  Q  W  D  L  P  N  A  W
1261 GGTGGAATTCCCACATCACTGAATAACACAGGAGAACAATGGGATCTTCCCAATGCCTGG
 441 P  P  L  Q  N  I  I  I  Q  G  L  Q  S  T  K  E  P  S  A  Q
1321 CCACCTCTCCAAAACATTATTATTCAAGGCTTGCAGTCCACTAAGGAACCTAGTGCCCAG
 461 K  L  A  Y  E  F  A  E  Q  W  I  R  S  N  Y  K  G  F  V  E
1381 AAATTAGCTTATGAATTCGCTGAACAATGGATTCGATCAAACTACAAAGGATTTGTGGAA
 481 H  H  D  M  Y  E  K  Y  D  A  E  V  P  G  N  S  G  G  G  G
1441 CACCATGATATGTATGAAAAGTATGATGCGGAGGTTCCAGGAAATTCTGGAGGAGGGGGA
 501 E  Y  I  P  Q  S  G  F  G  W  T  N  G  V  V  L  E  L  L  D
1501 GAATACATTCCACAATCAGGATTTGGATGGACAAATGGAGTGGTTCTTGAACTATTGGAC
```

```
521 T  Y  G  T  R  L  T  A  D  N  T  M  R  R  E  V  *
```

E

```
  1                                                M  S  V  Q  S  D
  1                                                ATGTCGGTTCAGAGCGAT
  7 I  E  K  R  K  Q  I  S  V  R  G  I  A  Q  V  E  N  V  S  T
 19 ATCGAGAAGCGCAAGCAGATCAGCGTCCGGGGCATAGCCCAGGTGGAAAACGTCTCCACC
 27 I  K  K  T  F  N  R  H  L  H  Y  T  L  V  K  D  R  N  V  A
 79 ATAAAGAAGACGTTCAACCGACACCTGCATTACACGCTCGTCAAGGACCGGAACGTTGCC
 47 T  P  R  D  Y  Y  F  A  L  A  H  S  V  K  D  N  L  V  S  R
139 ACCCCGCGAGACTACTACTTCGCGTTGGCCCACTCCGTCAAAGACAACCTGGTGTCGAGA
 67 W  I  R  T  Q  Q  Y  Y  Y  E  K  D  P  K  R  V  Y  Y  L  S
169 TGGATCCGCACACAGCAGTACTACTACGAGAAGGACCCCAAGAGAGTTTACTACCTATCC
 87 L  E  Y  Y  M  G  R  T  L  Q  N  T  M  I  N  L  G  I  Q  G
229 TTGGAGTACTACATGGGCCGTACCCTACAGAACACCATGATCAACCTGGGCATACAAGGA
107 A  C  D  E  A  M  Y  Q  L  G  L  D  I  E  E  L  E  D  L  E
289 GCCTGTGATGAAGCTATGTACCAGCTGGGCTTGGACATTGAGGAGTTGGAGGACCTGGAG
127 E  D  A  G  L  G  N  G  G  L  G  R  L  A  A  C  F  L  D  S
349 GAGGATGCTGGTTTGGGAAATGGAGGCCTTGGTCGTTTAGCAGCCTGCTTCCTGGACTCC
147 M  A  T  L  G  L  A  A  Y  G  Y  G  I  R  Y  E  Y  G  I  F
409 ATGGCCACACTCGGTCTGGCAGCGTATGGATACGGTATCAGATATGAGTATGGCATCTTC
167 T  Q  K  I  R  N  G  E  Q  V  E  E  P  D  D  W  L  R  F  G
469 ACGCAGAAGATTCGTAATGGAGAACAGGTAGAGGAGCCGGATGACTGGCTGAGATTTGGA
187 N  P  W  E  K  A  R  P  E  Y  M  L  P  V  N  F  Y  G  R  V
529 AACCCCTGGGAGAAGGCAAGACCAGAATATATGTTACCAGTTAATTTCTACGGTCGAGTG
207 E  Q  T  P  Q  G  V  K  W  V  D  T  Q  V  I  F  A  M  P  Y
589 GAACAGACTCCTCAAGGCGTCAAATGGGTGGATACACAGGTGATATTTGCGATGCCCTAT
227 D  S  P  V  P  G  Y  G  N  N  V  V  N  T  L  R  L  W  S  A
649 GACAGCCCTGTTCCAGGTTATGGTAATAATGTTGTGAACACATTGAGGCTGTGGTCTGCT
247 K  S  P  V  E  F  N  L  K  F  F  N  D  G  D  Y  I  Q  A  V
709 AAATCCCCTGTGGAATTTAACCTCAAATTTTTTAATGATGGTGACTACATCCAAGCAGTA
267 I  D  R  N  L  A  E  N  I  S  R  V  L  Y  P  N  D  N  F  F
769 ATAGATCGAAACTTGGCAGAGAACATCTCAAGAGTGCTGTATCCAAATGATAATTTCTTT
287 E  G  K  E  L  R  L  K  Q  E  Y  F  M  C  A  A  T  L  Q  D
829 GAAGGAAAAGAACTTCGTCTTAAACAAGAGTATTTCATGTGTGCTGCTACTTTGCAGGAT
307 L  I  R  R  F  K  A  S  K  F  G  V  R  T  A  V  R  T  N  F
869 CTCATACGCCGTTTTAAGGCTTCAAAATTTGGTGTCCGGACTGCAGTGAGAACAAATTTT
327 E  H  F  P  D  K  V  A  I  Q  L  N  D  T  H  P  S  L  A  I
929 GAACACTTCCCTGACAAGGTTGCCATTCAGCTGAATGACACACATCCATCACTGGCAATC
```

```
 347 P  E  L  M  R  I  L  I  D  V  E  G  L  T  W  E  Q  A  W  P
 989 CCAGAGCTTATGCGGATCCTCATTGACGTGGAAGGCCTGACATGGGAACAGGCATGGCCC
 367 I  V  V  R  T  C  A  Y  T  N  H  T  V  L  P  E  A  L  E  R
1049 ATTGTGGTTCGCACATGTGCCTATACAAATCACACTGTGCTACCTGAAGCTTTGGAGCGA
 387 W  P  V  H  L  L  E  T  I  L  P  R  H  L  Q  I  I  Y  H  I
1109 TGGCCTGTTCATTTACTGGAGACCATCTTGCCACGCCATCTGCAGATCATCTATCACATT
 407 N  F  L  H  L  Q  E  V  G  K  K  Y  P  G  D  L  D  R  L  R
1169 AATTTCCTTCACCTGCAGGAAGTGGGAAAAAAATACCCTGGAGATCTTGATCGTCTTCGG
 427 R  M  S  L  V  E  E  G  G  E  K  R  I  N  M  A  H  L  C  I
1229 AGAATGTCCCTTGTGGAGGAAGGCGGAGAGAAAAGAATTAATATGGCTCATCTCTGCATT
 447 V  G  S  H  A  V  N  G  V  A  R  I  H  S  D  I  I  K  S  D
1289 GTGGGTTCTCATGCAGTCAATGGAGTTGCAAGAATTCATTCTGACATCATTAAGAGTGAC
 467 I  F  K  D  F  Y  E  M  T  P  E  K  F  Q  N  K  T  N  G  I
1349 ATCTTCAAAGATTTCTATGAAATGACTCCAGAGAAATTTCAAAATAAGACAAATGGTATC
 487 T  P  R  R  W  L  L  L  C  N  P  G  L  S  D  L  I  A  E  K
1409 ACTCCTCGTCGCTGGTTGCTGCTATGTAATCCTGGATTATCTGATTTGATTGCAGAGAAA
 507 I  G  E  E  W  I  T  H  L  E  Q  L  Q  Q  L  K  E  F  A  K
1469 ATTGGCGAGGAATGGATTACTCACCTGGAGCAACTGCAGCAATTGAAAGAATTTGCAAAG
 527 D  P  G  F  Q  R  A  V  Q  K  V  K  Q  E  N  K  L  R  L  A
1529 GATCCAGGCTTTCAGAGAGCTGTCCAAAAAGTGAAGCAAGAGAACAAATTGAGGCTGGCT
 547 Q  I  L  E  K  D  Y  G  V  K  V  N  P  A  S  M  F  D  I  Q
1589 CAAATTCTGGAGAAGGACTATGGTGTCAAGGTGAATCCAGCATCCATGTTTGACATTCAG
 567 V  K  R  I  H  E  Y  K  R  Q  L  L  N  C  L  H  I  I  T  L
1649 GTTAAACGTATTCATGAATACAAACGTCAGTTGCTGAATTGTCTGCATATAATAACACTG
 587 Y  N  R  I  K  R  D  P  S  A  N  I  T  A  R  T  V  M  I  G
1709 TACAATCGCATAAAGAGAGATCCATCTGCAAATATCACTGCAAGAACTGTGATGATTGGA
 607 G  K  A  A  P  G  Y  Y  M  A  K  K  I  I  K  L  I  N  A  V
1769 GGCAAGGCTGCTCCTGGCTATTATATGGCAAAGAAGATTATTAAGCTGATCAATGCTGTT
 627 G  N  I  V  N  N  D  P  I  V  G  D  K  L  K  V  I  F  L  E
1829 GGCAACATTGTGAACAATGACCCCATTGTTGGAGATAAATTGAAAGTAATTTTCCTAGAA
 647 N  Y  R  V  T  L  A  E  K  I  M  P  A  A  D  L  S  E  Q  I
1889 AACTACAGAGTGACACTTGCTGAAAAGATAATGCCAGCTGCAGATTTGAGTGAACAGATC
 667 S  T  A  G  T  E  A  S  G  T  G  N  M  K  F  M  L  N  G  A
1949 TCCACTGCTGGTACAGAAGCATCTGGAACTGGAAATATGAAGTTTATGTTGAATGGAGCC
 687 L  T  I  G  T  L  D  G  A  N  V  E  M  A  E  E  M  G  R  E
2009 CTCACAATTGGAACCTTGGATGGTGCCAATGTAGAAATGGCTGAGGAAATGGGTCGGGAA
 707 N  I  F  I  F  G  M  T  V  D  E  V  E  A  L  K  K  R  G  Y
2069 AACATCTTCATCTTTGGAATGACTGTGGACGAAGTAGAGGCACTGAAGAAACGAGGATAC
 727 N  A  H  D  Y  Y  A  A  N  P  E  I  Q  Q  C  V  D  Q  I  R
2129 AATGCCCATGACTACTATGCTGCAAATCCTGAGATCCAACAGTGTGTGGACCAGATCCGT
 747 N  G  F  F  S  P  E  N  P  S  E  F  A  D  V  A  D  V  L  M
2169 AATGGATTTTTCAGTCCTGAAAACCCATCAGAGTTTGCTGATGTGGCAGATGTGCTGATG
 767 K  Y  D  R  F  L  T  L  A  D  F  D  A  Y  I  K  C  Q  D  T
```

```
2229 AAGTATGATCGCTTCCTCACTCTTGCAGACTTCGATGCATACATCAAGTGCCAAGATACA
 787 V  S  A  V  Y  Q  D  Q  A  K  W  S  E  M  A  I  N  N  I  A
2289 GTATCAGCTGTATACCAGGATCAAGCAAAATGGTCAGAGATGGCTATCAACAACATTGCT
 807 S  S  G  K  F  S  S  D  R  T  I  A  E  Y  A  R  E  I  W  S
2349 TCATCGGGCAAGTTCTCAAGTGACCGCACAATTGCTGAATATGCTCGTGAAATTTGGAGT
 827 V  E  P  S  W  E  K  L  P  A  P  H  E  P  R  D  D  D  T  A
2409 GTGGAACCATCATGGGAAAAGTTACCTGCTCCGCATGAACCACGTGACGATGACACAGCT
 847 A  N  G  L  A  A  G  G  K  *
2469 GCAAATGGGCTAGCAGCAGGTGGAAAGTAG
```

F

```
   1  G  L  T  F  F  L  S  F  A  I  L  A  A  V  L  G  M  L  Q  F
   1 GGGCTTGACGTTTTTCCTATCTTTTGCAATACTGGCAGCAGTGCTGGGGATGCTACAGTT
  21  G  Y  N  T  G  V  I  N  A  P  E  A  N  I  E  N  F  M  K  D
  61 TGGTTATAATACTGGGGTCATAAACGCTCCCGAGGCGAACATTGAAAATTTTATGAAAGA
  41  V  Y  K  H  R  Y  G  E  D  M  T  D  D  T  V  K  L  L  Y  S
 121 TGTGTACAAGCATCGCTATGGTGAGGATATGACTGACGATACCGTGAAGTTGCTATATTC
  61  L  A  V  S  I  F  A  I  G  G  M  F  G  G  F  C  G  G  M  I
 181 CTTGGCAGTATCGATATTTGCCATTGGAGGCATGTTCGGTGGCTTTTGTGGAGGAATGAT
  81  A  N  K  F  G  R  K  G  G  L  L  L  N  N  I  L  G  V  L  G
 241 TGCCAACAAATTTGGACGAAAAGGTGGTTTGTTACTAAATAATATTCTCGGAGTTCTTGG
 101  A  V  L  M  G  C  T  K  E  T  Q  W  Y  E  V  L  F  L  G  R
 301 TGCTGTATTAATGGGTTGTACGAAAGAAACACAATGGTATGAAGTACTGTTTCTCGGGCG
 121  L  I  I  G  V  N  C  G  L  N  T  S  L  V  P  M  Y  I  S  E
 361 TCTTATCATTGGTGTCAATTGTGGTCTAAATACATCATTGGTTCCAATGTACATCTCAGA
 141  I  A  P  L  N  L  R  G  G  L  G  T  V  N  Q  L  A  V  T  V
 421 AATTGCTCCTCTCAATTTAAGAGGGGGTCTGGGTACCGTTAACCAATTGGCTGTGACTGT
 161  G  L  L  L  S  Q  V  L  G  I  E  Q  I  L  G  T  D  E  G  W
 481 TGGCTTGCTGCTTTCACAAGTACTTGGCATCGAGCAAATTCTTGGCACAGACGAAGGATG
 181  P  I  L  L  G  L  A  V  C  P  A  L  L  Q  L  T  L  L  P  F
 541 GCCAATCCTTCTTGGTCTAGCAGTTTGTCCAGCACTTCTACAGCTGACTTTGTTGCCTTT
 201  C  P  E  S  P  R  Y  L  L  I  T  K  Q  W  E  E  Q  A  R  Q
 601 TTGCCCTGAGTCACCAAGATACTTGCTCATTACGAAGCAGTGGGAAGAACAAGCGAGGCA
 221  A  L  R  R  L  R  A  S  N  Q  V  E  E  D  I  E  E  M  R  A
 661 AGCGCTACGGAGGTTGCGTGCTTCAAATCAAGTGGAGGAAGACATCGAAGAAATGCGAGC
 241  E  E  R  A  Q  Q  A  E  A  T  I  S  M  M  E  L  I  C  S  P
 721 AGAAGAAAGAGCCCAGCAAGCGGAAGCCACGATTTCCATGATGGAATTAATTTGCTCGCC
 261  T  L  R  S  P  L  I  I  G  V  V  M  Q  L  S  Q  Q  L  S  G
 781 GACACTTAGATCGCCTCTTATCATTGGGGTGGTAATGCAGCTTTCGCAACAACTATCAGG
 281  I  N  A  V  F  Y  Y  S  T  G  L  F  V  S  S  G  L  S  E  E
 841 GATTAATGCTGTTTTCTATTATTCGACAGGTCTGTTTGTGAGCTCTGGACTGTCGGAAGA
```

```
 301  Q  A  K  F  A  T  I  G  I  G  S  I  M  V  V  M  T  L  V  S
 901 GCAAGCAAAATTTGCCACAATTGGTATAGGTTCGATAATGGTGGTGATGACCCTTGTGTC
 321  I  F  L  M  D  R  M  G  R  R  S  L  H  L  Y  G  L  G  G  M
 961 CATATTCTTGATGGATCGCATGGGTCGGAGATCACTGCACCTGTACGGCCTGGGTGGCAT
 341  F  I  F  S  I  F  I  T  I  S  F  L  I  K  E  F  F  G  Y  V
1021 GTTTATCTTCTCAATATTTATCACCATCTCTTTCCTTATAAAGGAGTTTTTTGGTTACGT
 361  Q  D  M  I  D  W  M  S  Y  W  S  V  V  S  T  L  S  F  V  V
1081 CCAGGACATGATTGACTGGATGTCATACTGGTCTGTGGTATCTACGTTGAGCTTTGTGGT
 381  F  F  A  V  G  P  G  S  I  P  W  M  I  T  A  E  L  F  S  Q
1141 GTTCTTCGCTGTGGGTCCTGGATCCATCCCGTGGATGATAACTGCGGAGCTCTTTTCCCA
 401  G  P  R  P  A  A  M  S  I  A  V  L  V  N  W  M  A  N  F  L
1201 AGGACCACGACCAGCAGCTATGTCCATTGCTGTATTGGTCAACTGGATGGCGAACTTCCT
 421  V  G  I  G  F  P  S  M  N  T  A  L  E  N  Y  T  F  L  P  F
1261 TGTTGGAATTGGATTTCCCAGTATGAATACTGCCCTGGAAAACTACACATTTCTTCCCTT
 441  S  L  F  L  A  I  F  W  I  F  T  Y  K  K  V  P  E  T  K  N
1321 CAGCTTGTTCCTTGCAATCTTTTGGATTTTTACGTACAAGAAAGTGCCAGAGACCAAGAA
 461  K  T  F  E  E  I  L  A  L  F  R  Q  G  N  G  R  G  S  L  R
1381 TAAAACATTCGAAGAAATTCTAGCACTCTTTCGACAGGGAAATGGCAGGGGCAGTCTTCG
 481  D  S  R  L  Y
1441 GGACAGTAGGTTGTATGG
```

Figure S1 cDNAs and deduced amino acid sequences of Trehalose transporter (Tret 1) (A), Trehalose phosphate synthase (B), Trehalase-1 (C), Trehalase-2 (D), Glycogen polyphosphate (E) and Glucose transporter (Glut 1) (F).

Materials and methods.

Insects

The two-spotted crickets *G. bimaculatus* were reared in plastic containers (55×39×31cm) in the temperature conditioned room at 27±1℃, with the long-day photoperiod by 16h:8h as described (Tsukamoto et al., 2014).

Bibliometric network

The literature data for Chapter 1 were collected from the publications dating from 1990 to 2018 via PubMed (https://www.ncbi.nlm.nih.gov/pubmed/). All titles and abstracts were extracted to select the items by the binary counting method over the limitation by the number of 10 as the limitation. Finally, screened keywords for nodes were selected manually to generate the bibliometric mapping.

RNA-sequencing

Total RNA was extracted using TRI reagent® as described in the commercially available protocol (Molecular Research Center, Inc., Cincinnati, OH, USA). The extracted total RNA was further purified by RNeasy® Mini Kit (QIAGEN, Venlo, Netherlands). RNA sequencing was subjected to HiSeq® 2000 system (Illumina, Inc., San Diego, CA, USA). Contigs were de novo assembled using the software CLC genomic workbench (https://www.qiagenbioinformatics.com/products/clc-main-workbench/) and comprehensively annotated using blastx.

Chemically synthesis of neuropeptides

The mature neuropeptides were manually synthesized according to the Fmoc method C-terminally primed by a Rink-amide resin (Merck Millipore, Darmstadt, Germany). Fmoc amino acids were purchased from Kokusan Chemical (Tokyo, Japan). Fmoc moieties from the chemically extended peptide on the resin were de-protected by incubation with 20% piperidine in dimethylformamide at room temperature for 30 min. Fmoc amino acids were then linked to the first amine group on the synthesizing peptide on the resin in the presence of 1-hydroxybenzotriazole and N, N-dicyclohexylcarbodiimide in

N-methylpyrolidone. The resulting peptidyl resin was incubated to deprotect all protecting groups and release the peptide from the resin with a cleavage cocktail composed of trifluoroacetic acid (TFA)/ dichloromethane (DCM)/ anisole/ trimethylsilyl bromide (TMSBr)/ 3,4-ethoxylene dioxy thiophene (EDT) of the ration 10/5/2/2/1, v/v/v/v/v on ice for 30 min. The resulting crude peptide was applied to reversed-phase HPLC (Jasco SC-802, PU-880, UV-875; Jasco INt., Tokyo, Japan) on a Senshu Pak PEGASIL-300 ODS column (4.6 mm i.d. ×250 nm; Senshu Kagaku, Tokyo, Japan). The adsorbed peptide was eluted over the program of a linear gradient of 10-60% acetonitrile containing 0.1% TFA over 30 min, at a flow rate of 1 ml/min monitored by absorbance at 225 nm. The highly purified synthetic peptides were confirmed by measurement the molecular mass using MALDI-TOF MS as described below.

MALDI-TOF MS analysis

MALDI-TOF MS and MSMS analyses were performed by TOF/TOF 5800 System (AB SCIEX, Framingham, MA, USA). Peptide solutions were subjected to the target of sample stage after mixing with a semi-saturated solution of α-cyano-4-hydroxycinnamic acid (CHCA; 1:1 v/v) and followingly dried. Amino acid sequencing MSMS analyses were analyzed using the resulting spectra acquired by accumulation of 250–1000 laser shots in total.

RNA interference (RNAi)

Double-stranded RNA (dsRNA) was prepared using template plasmid DNA carrying the fragment DNAs obtained from tissue distribution analysis. DsRed-dsRNA was used as an experimental control of dsRNA, corresponding to 14 to 420 nt of DsRed2 nucleotide sequence. The used primers were listed below. RNA was synthesized using T7 RNA polymerase according to manufacturer's protocol using 500 ng of PCR products as template DNA primed by T7 sequence overhanged on the 5' portion of primers. The resulting RNA was purified by phenol/chloroform extraction and ethanol precipitation. The purified RNA was dissolved in diethylpyrocarbonate-treated water to a final concentration of 0.5 μg/μl. dsRNA was generated by gradually cooling the denatured synthesized RNA in 100℃ for 5 min and then cooled down to room temperature. Resulting dsRNA was then stored in -20℃ until experimental use.

Expression analysis

cDNAs which were reverse transcripted from total RNA derived from various tissues of adult crickets were prepared as described above. Reverse transcription was primed by an oligo dT18 primer using ReverTra Ace® (Toyobo Co. Ltd., Osaka, Japan). The cDNAs were used as template DNAs for RT-PCR using Gotaq Green Master Mix (Promega Corporation, Madison, WI, USA). The used primers were listed below. PCR products were electrophoresed on a 1.0% agarose gel and visualized by staining by ethidium bromide. DNA bands were extracted and purified using Wizard SV Gel and PCR Clean-up system (Promega Corporation, Madison, WI, USA) and subsequently perform TA-cloning into pGEM-T Vector (Promega Corporation, Madison, WI, USA) according to the manufacturer's protocol. The extraction of the resulting plasmid DNA was primed by the universal M13 primer as listed below, and the inserted cDNAs were sequenced for confirmation using 3500 Genetic Analyzer (Applied Biosystems, Waltham, MA, USA). Real-time quantitative RT-PCR was performed by Thermal Cycler Dice with SYBR-premix Ex Taq-II (TaKaRa, Bio, Japan) with the primers as listed below. The transcriptional levels were normalized to Elongation factor by ΔΔCt method to calculate the relative transcriptional level. Primers were designed according to http://www.primer3plus.com/cgi-bin/dev/primer3plus.cgi.

Lipid and carbohydrate extraction from hemolymph

Lipid and carbohydrate were extracted according to Lorenz's method (Lorenz, 2003). Collected 5 µl of the crickets' hemolymph from the abdomen was transferred in a tube containing 20 mg of solid sodium sulfate and 200 µl of 75% methanol. The mixture was sonicated for 5 min. After adding 600 µl of chloroform/methanol (1:1), the sonicated solution was then vortexed and centrifuged (15,000 rpm, 4°C, 10 min). The resulting supernatant was transferred into a new centrifuging tube. Next, 300 µl of chloroform/methanol (1:1) was then added. After vortexing and centrifugation (15,000 rpm, 4°C, 10 min), the supernatant was mixed to the former tube containing aqueous layer. The collected aqueous layer was vortexed and centrifuged again (15,000 rpm, 4°C, 10 min) after addition of 500 µl chloroform and 300 µl 1 M NaCl. The organic lower layer and the aqueous upper layer, which were then dried under vacuum with centrifugation, were used as lipid fraction and carbohydrate fraction for quantification of lipid and trehalose (insect blood sugar), respectively.

Measurement of lipid

Lipid extracted from cricket hemolymph as described above was quantified according to Sulfo-phospho-vanillin method (Zöllner and Kirsch, 1962). Lipid fractions were dissolved in 40 µl of chloroform/ methanol (1:1). 5 µl of the lipid solution was mixed with 100 µl of sulfuric acid, followed by heated at 100℃ for 10 min. After cooling to room temperature, the heated lipid solution was mixed with 500 µl of vanillin reagent (0.2% vanillin in 67% o-phosphoric acid). Absorbance at 595 nm was measured using a spectrophotometer. Quantification of lipids were performed compared with the standard curve generated by Cholesterol (Sigma-Aldrich Japan, Tokyo, Japan); a standard curve of 100, 20, 4, 0.8, 0.16 µg/µl and blank.

Measurement of carbohydrate

Carbohydrate fraction was dissolved in 100 µl distilled water. 20 µl of the carbohydrate fraction was incubated with 1 µl trehalase (Sigma-Aldrich Japan, Tokyo, Japan) and 5 µl of 0.27 M Citric acid (pH 5.7) at 37℃ for 4 hours. After cooling to room temperature, solution is measured by Glucose assay kit (Wako, Tokyo, Japan) according to the manufacture's protocol. The glucose concentration was compared with standard samples 100, 20, 4, 0.8, 0.16 µg/µl of glucose, and then calculated as carbohydrate levels as trehalose level.

Food intake assay

On third day after adult emergence, 5 µg of dsRNA or 100 pmol of chemically synthesized neuropeptides were injected. These treated crickets were kept isolated in a plastic container. Diet consumption was quantified the difference of diet before and after experiments (24 hours). Before weighing the tablets, the diet tablets were dried by baking at 80 °C for 2 hours. For preparation of dietary tablets, indigestible cellulose was mixed with three different nutrients, which were isocaloric (4 kj/g). All examined diets contained 33% casein, 33% dextrin and 33% soybean oil for total energy.

Excretion assay

Excretion activity was measured by counting the number of fecal pellets, which has been used

previously as an index of food intake (Konuma et al., 2012). After treatment of dsRNA and peptides, I counted on the next day of injection. In the experiment of RNAi, I counted the third day after treatment.

First-bite time assay

Crickets were kept isolated in a plastic container with water only for 24 hours. At 1 hour after treated with 100 pmol of chemically synthesized neuropeptides, normal diets were replaced. Time were measured until crickets firstly accessed to the food.

Crop storage assay

Crickets were dissected for intestine 1 hour after application of diet and injection with 100 pmol chemically synthesized GbCCAP. Width of crop was measured by a scale.

RNA probe synthesis

DNA templates were amplified by Ex Taq with 5 ng of vector and gene specific primer with T7 sequence. After purifying PCR products by phenol/chloroform/isopropanol and 70% ethanol, products are incubated with transcription buffer, 0.1 M DTT, BSA, DIG-labeling mixture and T7 polymerase at 37℃ for 2 hours. After DNase treatment and purification. The quality of riboprobe was confirmed by electrophoresis with 1.5% TAE agarose gel.

In situ hybridization

Tissues were dissected in PBS (Wako, Tokyo, Japan) and fixed in 4% paraformaldehyde (PFA)/PBS for 12 hours at 4℃. After washing with PBST (PBS + 0.5% Triton X-100, Invitrogen) (3 times, 10 min each), tissues were digested by Proteinase K (2 μg/ml, Wako) for proper time. After re-fixing with 4% PFA solution and 0.2% glutaraldehyde, tissues are washed with PBST 3 times. RNA probe (0.1 ng/μl) is added to hybridization buffer (1.5% blocking regent, 50% Formamide, 25% 20xSSC (pH 7.0), 0.1% TritonX-100, 0.1% CHAPS, 1% DNA from salmon sperm) and incubated in a moisture chamber for 48 h at 70℃ after a 120 min pre-hybridization. Tissues were then blocked in 1.5% blocking buffer/KTBT (1.5 M NaCl, 0.1 M KCl, 1 M Tris-HCl (pH 7.5)) for 1h at room temperature. After washed by wash

solution (10% 20xSCC (pH=7.0), 0.1% CHAPS) (3 times, 60 min each), the tissues were incubated with secondary antibody (Anti-DIP AP, Fab fragments from sheep, 1:200, Sigma-Alrich, Tokyo, Japan) for 24 h at 4℃ and washed again with KTBT (5 times, 10 min each). CDP-star (Sigma-Aldrich Japan, Tokyo, Japan) is used for staining according to the manufacture's instruction.

Immunohistochemistry

Tissues were dissected in PBS and were fixed in 4% paraformaldehyde (PFA)/PBS for 12 hours at 4℃. After washing in PBST (PBS + 0.5% Triton X-100) (3 times, for 10 min each), tissues were blocked in 10% normal goat serum (Wako, Tokyo, Japan) in PBST for 1 h at room temperature, and then incubated with a primary antibody (1:1500) for 48 h at 4℃. After washing with PBST (3 times, for 10 min each), the tissues were incubated with a secondary antibody (Alexa Fluor 488 goat anti-rabbit IgG, 1:1000, Invitrogen, A27034) for 24 hours at 4℃ and washed again with PBST (5 times, 10 min each). Control tissues are treated with secondary antibody only. For CCAP immunohistochemistry and MSR in situ hybridization, first antibody of immunohistochemistry is incubated together with secondary antibody in MSR in situ hybridization.

Gut contraction assay

After removal of head, crickets on the fifth days after adult emergence are spread out along the abdomen by showing foregut clearly with no harm on peripheral neurons in the 0.9% NaCl. Their foreguts were exposed to 100 pmol synthetic MS for the foregut was conducted observation of gut contraction for 40 s. Experimental control was treated with PBS in the same methods. The contractile movement was observed and recorded using a digital microscope USB2.0 DigiScope II v2TM (CHRONOS, Taiwan). The contractile movements were analyzed using the ImageJ software (https://imagej.net/Fiji) with the Optic flow plug-in (Gaussian window MSE). The output results were visualized as a kymograph.

Semi-quantification of neuropeptides by MALDI-TOF MS

Extracts of crickets' brains are supersonic treated with 300μl of 60% acetonitrile containing 0.1%

TFA after immersing for 30 minutes, following Zip-Tip treatment for isolation of peptides. Products are semi-quantification measured by MALDI-TOF MS and MSMS with internal standard of *B. mori* peptide GSRYamide-1.

Statistical analyses

Statistical analyses were performed using a software, Prism 7 (GraphPad). Each statistical method was described in the figure legends.

Primer used in this study

Elongation factor RT-PCR-f: 5'- ATGCCTGTATCTTGACTGCTCA -3'
Elongation factor RT-PCR-r: 5'- ATGGTTTGCTTCCAGTTCAGT -3'
Elongation factor real time PCR-f: 5'- CCCTGCTGCTGTTGCTTT -3'
Elongation factor real time PCR-r: 5'- CCCATTTTGTCGGAGTGC -3'
DsRed-f: 5'- GAACGTCATCACCGAGTTCA -3'
DsRed-r: 5'- TGGTCTTCTTCTGCATCACG -3'
M13 RT-PCR-f: 5'- GTTTTCCCAGTCACGACGTT -3'
M13 RT-PCR-r: 5'- GGAAACAGCTATGACCATGA -3'
T7-Special-f: 5'-TAATACGACTCACTATAGGGAGACCGCGGGAATTCGAT-3'
T7-Special-r: 5'-TAATAGACTCACTATAGGGAGAGAATTCACTAGTGAT -3'
GbAKH RT-PCR-f: 5'- TCCATGGAGGAATTGCACCC -3'
GbAKH RT-PCR-r: 5'- CTTGGTTCTGGTTTTGGCGG -3'
GbAKH real time PCR-f: 5'- CGCCAAAACCAGAACCAAGG -3'
GbAKH real time PCR-r: 5'- TGGAGTTCAGACGTGTGCTC -3'
GbACP RT-PCR-f: 5'- TCCGCAACTCATGGCTTCCT -3'
GbACP RT-PCR-r: 5'- TGGAGAAGGTGATCTGCGCG -3'
GbACP real time PCR-f: 5'- CCACCCATCTGCGGTAGAAG -3'
GbACP real time PCR-r: 5'- CCATGAGTTGCGGAGTAGCT -3'
GbCRZ RT-PCR-f: 5'- CGAAGGATCTCCCACAGACG -3'
GbCRZ RT-PCR-r: 5'- TGAGGAATACTCGTCGCTGC -3'
GbCRZ real time PCR-f: 5'- GGTGCGTGCTCCTGCTGAGT -3'
GbCRZ real time PCR-r: 5'- CCGCGGCTGTACTGGAAGGT -3'
GbCCAP RT-PCR-f: 5'- CCACAAGAGGCAAGCGGAC -3'
GbCCAP RT-PCR-r: 5'- CCGGTGAAGGCGTTGCAGAA -3'
GbCCAP real time PCR-f: 5'- CCACAAGAGGCAAGCGGAC -3'
GbCCAP real time PCR-r: 5'- CCGGTGAAGGCGTTGCAGAA -3'
GbMS RT-PCR-f: 5'- CCCGTCATCAAGAGGCAAGA-3'
GbMS RT-PCR-r: 5'- ACATGCTTTGTACAGGGCCG -3'
GbMS real time PCR-f: 5'- AGAGGCAAGACGTCGATCAC-3'
GbMS real time PCR-r: 5'- ACATGCTTTGTACAGGGCCG -3'
GbAst-A RT-PCR-f: 5'- AAGGCCGCATGTACTCCTTC -3'
GbAst-A RT-PCR-r: 5'- GGTCCCCATGTACGACTTCG -3'
GbAst-A real time PCR-f: 5'- TCCGAGTACAAGCGCCTGCC -3'
GbAst-A real time PCR-r: 5'- GAAGCTGTACATGCCGGCGC -3'
GbMIP RT-PCR-f: 5'-GGGGTGGGCAGTTACTCTTC -3'
GbMIP RT-PCR-r: 5'- TCGTACTGGCCCTCGATCTC -3'
GbMIP real time PCR-f: 5'- ACGTGGCCGGTTCCTCCTTC -3'
GbMIP real time PCR-r: 5'- TGAGCGCTGTGTGGGGGAAG -3'
GbAst-C RT-PCR-f: 5'- CAGGAGACGGCGTTGAAGG -3'
GbAst-C RT-PCR-r: 5'- CTACCATCGTCGTCCACCAG -3'
GbAst-C real time PCR-f: 5'- CAGTAGCTGCGCTTGCGCTG -3'

GbAst-C real time PCR-r: 5’- CGTCGAGACCGCGCTCATCA-3’
GbAT RT-PCR-f: 5’- AGCAAGCCCCGCACGATACG -3’
GbAT RT-PCR-r: 5’- AGGGGCTCAGTACAGTGGCCT -3’
GbAT real time PCR-f: 5’- TTCGCCCAGCACGACAGCTT -3’
GbAT real time PCR-r: 5’- ACAGCTCGCCGTCCTGGTTC -3’
GbsNPF RT-PCR-f: 5’- CGCTGTCCCTGATGCTGCTACTC-3’
GbsNPF RT-PCR-r: 5’- CCTCACAATAAATGGCGGTGCTT -3’
GbsNPF real time PCR-f: 5’- TTTCCGGAATTGCTTTGCCG-3’
GbsNPF real time PCR-r: 5’- CGCTTGCGAACATACTCAGC -3’
GbFMRFa RT-PCR-f: 5’- TCGGAAGGGCCAACCA -3’
GbFMRFa RT-PCR-r: 5’- TCCAGCGACACCTGTGC -3’
GbFMRFa real time PCR-f: 5’- CGCGCGGGAGAGGAAATGGA -3’
GbFMRFa real time PCR-r: 5’- AAGAAGTTGGATGCGGCGCG -3’
GbDH31 RT-PCR-f: 5’- CGTCGTTGGGCGTGATGAGTGA -3’
GbDH31 RT-PCR-r: 5’- CTAGCGGCGAGGGAGGAAGAA -3’
GbDH31 real time PCR-f: 5’- GCTGGAGATGCTGGCACGGA -3’
GbDH31 real time PCR-r: 5’- GGGCGCGCTTGTTCTCAACG -3’
GbDH44 RT-PCR-f: 5’- GCAATCGCATCCAGCAGAACC -3’
GbDH44 RT-PCR-r: 5’- AAGGGAAACAGGGTGACAGAACAGA -3’
GbDH44 real time PCR-f: 5’- AGTCGCTGTCCATCGTGGCG -3’
GbDH44 real time PCR-r: 5’- GCTGGATGCGATTGCCTTGC -3’
GbAst-A receptor RT-PCR-f: 5’- AGATCACGCAGAAGGTGGTG -3’
GbAst-A receptor RT-PCR-r: 5’- CGGATGGACATCGACGTGAT -3’
GbAst-A receptor real time PCR-f: 5’- TTTTGCGTGCCGTTCACCGC -3’
GbAst-A receptor real time PCR-r: 5’- TTGCACCAGACGTCCCCGAA -3’
GbBur-α receptor RT-PCR-f: 5’- CAAGCTGTTCCTGTCCGACT -3’
GbBur-α receptor RT-PCR-r: 5’- CGTTGTTGCCCAGGTTCAAG -3’
GbBur-α receptor real time PCR-f: 5’- ACCGCATTGGCAGCCTGGAA -3’
GbBur-α receptor real time PCR-r: 5’- TGCTCTGGCAGGTGCGTGAT -3’
GbBur-β receptor RT-PCR-f: 5’- CTCCGACCAGTACGCTTACC -3’
GbBur-β receptor RT-PCR-r: 5’- GAGCTCGACAGGTTGGAGTC -3’
GbBur-β receptor real time PCR-f: 5’- AGCAGAAGCAGCACAAGCGC -3’
GbBur-β receptor real time PCR-r: 5’- AGCGCGACGAGCTGAAGTTCT -3’
GbCCHa-2 receptor RT-PCR-f: 5’- GCTGTGCATCATCGCGTGCT -3’
GbCCHa-2 receptor RT-PCR-r: 5’- CACCTTTTTGCGGGCGCGAA -3’
GbCCHa-2 receptor real time PCR-f: 5’- GCTGTGCATCATCGCGTGCT -3’
GbCCHa-2 receptor real time PCR-r: 5’- CACCTTTTTGCGGGCGCGAA -3’
GbMIP receptor RT-PCR-f: 5’- ACTTCTTCGAGCAGTGCTACT -3’
GbMIP receptor RT-PCR-r: 5’- TGACCTGCGTCACGTGGTT -3’
GbMIP receptor real time PCR-f: 5’- AATGTTTCGGCCGCGCTTCC -3’
GbMIP receptor real time PCR-r: 5’- ACCTGCGTCACGTGGTTGGT -3’
GbNPF receptor RT-PCR-f: 5’- CAGTACGTGGTGCCCTTCAT -3’
GbNPF receptor RT-PCR-r: 5’- CTGGAACTCCTTGCGGAAGT -3’

GbNPF receptor real time PCR-f: 5’- TACTCGCGCATCTGCCGCAA -3’
GbNPF receptor real time PCR-r: 5’- TTCGTCTTGCGCCGAGTGGA -3’
GbRYa receptor RT-PCR-f: 5’- CTGTGCTACGGCTCCATCTC -3’
GbRYa receptor RT-PCR-r: 5’- GAAGATGCTCACGTTCACGC -3’
GbRYa receptor real time PCR-f: 5’- GCGCATGCAGAACGTCACCA -3’
GbRYa receptor real time PCR-r: 5’- AAGGTTCCAGCGCTGCAGCA -3’
GbFMRFa receptor RT-PCR-f: 5’- CGCCTCATGCTCTCCATCTT -3’
GbFMRFa receptor RT-PCR-r: 5’- AGATGAAGGGGTTGACGACG -3’
GbFMRFa receptor real time PCR-f: 5’- TGCTTCCTGCCGCTCATGCT -3’
GbFMRFa receptor real time PCR-r: 5’- AGATGAAGGGGTTGACGACGGC -3’
GbGlyco-A2 receptor RT-PCR-f: 5’- TGCTGGTAGCGAGTGCATTGTTG -3’
GbGlyco-A2 receptor RT-PCR-r: 5’- GGGTTGATACGCAGCGTGTTGG -3’
GbGlyco-A2 receptor real time PCR-f: 5’- ACTCGCACAGTCAGCATACC -3’
GbGlyco-A2 receptor real time PCR-r: 5’- CCAGGATTCGCAGTAGCCTC -3’
GbGlyco-B5 receptor RT-PCR-f: 5’- CTACAAGGTGGTGAAGGCGGACTC -3’
GbGlyco-B5 receptor RT-PCR-r: 5’- TGTTTGGGAAGCGCCAGTCC -3’
GbGlyco-B5 receptor real time PCR-f: 5’- CTACAAGGTGGTGAAGGCGGACTC -3’
GbGlyco-B5 receptor real time PCR-r: 5’- TGTTTGGGAAGCGCCAGTCC -3’
GbInotocin receptor RT-PCR-f: 5’- CCAATGAGGAATTGCGCGCG -3’
GbInotocin receptor RT-PCR-r: 5’- TCTGTGAGGTGGCAAG-3’
GbInotocin receptor real time PCR-f: 5’- TGTCTTGCCACACCGACCCC -3’
GbInotocin receptor real time PCR-r: 5’- AGGTCCGAGGGTCCATCCGA -3’
GbOrcokinin receptor RT-PCR-f: 5’- CGTGCCTCTACACCGATTCCTCA -3’
GbOrcokinin receptor RT-PCR-r: 5’- TATCTCGACGTTGGCTCAACTCATG -3’
GbOrcokinin receptor real time PCR-f: 5’- ACCGGCTTCGACAACTTCAT -3’
GbOrcokinin receptor real time PCR-r: 5’- CGTTGACGCCGACGAGGT -3’
GbPDF receptor RT-PCR-f: 5’- CTGCTCGACAAGGAGGTAGC -3’
GbPDF receptor RT-PCR-r: 5’- GCCCTGCATCATTCAGCAC-3’
GbPDF receptor real time PCR-f: 5’- GAAGCCAATGCTGCTCCTTCTCC -3’
GbPDF receptor real time PCR-r: 5’- GTGAGTTGATGATTTCCGAGTTCCTTT -3’
GbCCAPR-1 RT-PCR-f: 5’- GGGAGGCCCGACTTGTATTC -3’
GbCCAPR-1 RT-PCR-r: 5’- TTGAAAGTGACGTTCTGCGC -3’
GbCCAPR-1 real time PCR-f: 5’- GCAAGTGGTGTTGGATATTC -3’
GbCCAPR-1 real time PCR-r: 5’- ACCTCCAGTCAGTCATTCT -3’
GbCCAPR-2 RT-PCR-f: 5’- CGCGCGCCAAGATCAAGACGGT -3’
GbCCAPR-2 RT-PCR-r: 5’- AACGCAGCGCCGGGAAGCACAA -3’
GbCCAPR-2 real time PCR-f: 5’- TCGGAAATCATCAACTCACTCTTAGGTC -3’
GbCCAPR-2 real time PCR-r: 5’- CCACGGCGATGTTGGTCTGC -3’
GbMSR RT-PCR-f: 5’- CCACCAACTCCATCCTCACC -3’
GbMSR RT-PCR-r: 5’- GTTGTGGTCGCTCTTGTTG -3’
GbMSR real time PCR-f: 5’- TGCTCAGCCCGCTGCTCT -3’
GbMSR real time PCR-r: 5’- TGTGGTCGCTCTTGTTGTAGTCG -3’
GbAKHR RT-PCR-f: 5’- ATGGACAGGTACTTCGCCATTTT -3’

GbAKHR RT-PCR-r: 5’- ATGAAGAAGAACAGCACGATGAC -3’
GbAKHR RNAi-f: 5’- CCATTCCTGTCCCAGCCATT -3’
GbAKHR RNAi-r: 5’- AACTCGTGCATGAACCCGAT -3’
GbAKHR real time PCR-f: 5’- TCTGGGTACTCTCGCAGCGC -3’
GbAKHR real time PCR-r: 5’- CGTGTACGGCGCGTTCAACG -3’
GbIRS RT-PCR-f: 5’- CTACATGGAAGTGGGCAGGT -3’
GbIRS RT-PCR-r: 5’- GGGGTAGAAGGTGACATGGA -3’
GbIRS real time PCR-f: 5’- ACTTGCTCTGCCTTACGGAC -3’
GbIRS real time PCR-r: 5’- CTGTTGGCCTTCTCCTCCTG -3’
T7-*GbIRS*-f: 5’- GCTTCTAATACGACTCACTATAGCTACATGGAAGTGGGCAGGT -3’
T7-*GbIRS*-r: 5’- GCTTCTAATACGACTCACTATAGGGGGTAGAAGGTGACATGGA -3’
GbMSR-probe-f: 5’- GCCTGCTTCGTGCTGTTCC -3’
GbMSR-probe-r: 5’- TCCGTCTGGCGGCTCTTG -3’
T7-*GbMSR*-probe-f: 5’- TAATACGACTCACTATAGGGAGAGCCTGCTTCGTGCTGTTCC -3’
T7-*GbMSR*-probe-r: 5’- TAATACGACTCACTATAGGGAGATCCGTCTGGCGGCTCTTG -3’
Gb*Tret1* RT-PCR-f: 5’- GCGGCGAAAGACTAAAGACG -3’
GbTret1 RT-PCR-r: 5’- TGCAGGTTGATGCTGGTGA -3’
GbTret1 real time PCR-f: 5’- AGACTACCATCCTCGCCACC -3’
GbTret1 real time PCR-r: 5’- TGACGCCCACGCAGAAG -3’
GbGlut1 RT-PCR-f: 5’- GTCGGAGATCACTGCACCTGTAC -3’
GbGlut1 RT-PCR-r: 5’- CCCCTGCCATTTCCCTGT -3’
GbGlut1 real time PCR-f: 5’- CGCTACGGAGGTTGCGTGCTT-3’
GbGlut1 real time PCR-r: 5’- TTGCTGGGCTCTTTCTTCTGCTC-3’
GbGp RT-PCR-f: 5’- GATGGCTATCAACAACATTGCTTC -3’
GbGp RT-PCR-r: 5’- TTGTCTACGGCACTTCTGTTTCA -3’
GbGp real time PCR-f: 5’- TCTTCGGAGAATGTCCCTTGTGG -3’
GbGp real time PCR-r: 5’- CTCCATTGACTGCATGAGAACCC-3’
GbTPS RT-PCR-f: 5’- AGGCTAGGCTGCCGAGTG -3’
GbTPS RT-PCR-r: 5’- GCCCATTGATGCGTCCC -3’
GbTPS real time PCR-f: 5’- TGGCGGATTGTGGGTGGG -3’
GbTPS real time PCR-r: 5’- TGATTCAGGTATGGGCTCATTTGG -3’
GbTPS RT-PCR-f2: 5’- AGGCTAGGCTGCCGAGTG -3’
GbTPS RT-PCR-r2: 5’- GCCCATTGATGCGTCCC -3’
GbTPS real time PCR-f2: 5’- CGGCGTGCTCATCGTGTCG -3’
GbTPS real time PCR-r2: 5’- GGCGTTGATCTCGTAGGGGTTG -3’
GbTre1 RT-PCR-f: 5’- TCTGCTTAGCGATATGAAGGACAC -3’
GbTre1 RT-PCR-r: 5’- CATCGACTGGAGAAATCCCAAC -3’
GbTre1 real time PCR-f: 5’- CGTATCGCCGACCCCAAGCT -3’
GbTre1 real time PCR-r: 5’- CCTGACTTCGTCGGAAACCCTTC -3’
GbTre2 RT-PCR-f: 5’- CTCCTCCGTGCGAAAGTGAA-3’
GbTre2 RT-PCR-r: 5’- CAGGATGAGGCACATAGATAATGG -3’
GbTre2 real time PCR-f: 5’- ATGGACTGGCTGCTATGATACAAC -3’
GbTre2 real time PCR-r: 5’- TCCAGGCTTTCCACAAACAGAT -3’

References

AHMAD, M., HE, L. & PERRIMON, N. 2020. Regulation of insulin and adipokinetic hormone/glucagon production in flies. *Wiley Interdiscip Rev Dev Biol,* 9**,** e360.

ANTIN, J., GIBBS, J., HOLT, J., YOUNG, R. C. & SMITH, G. P. 1975. Cholecystokinin elicits the complete behavioral sequence of satiety in rats. *J Comp Physiol Psychol,* 89**,** 784-90.

ARAKANE, Y., LI, B., MUTHUKRISHNAN, S., BEEMAN, R. W., KRAMER, K. J. & PARK, Y. 2008. Functional analysis of four neuropeptides, EH, ETH, CCAP and bursicon, and their receptors in adult ecdysis behavior of the red flour beetle, Tribolium castaneum. *Mech Dev,* 125**,** 984-95.

AUDSLEY, N. & WEAVER, R. J. 2009. Neuropeptides associated with the regulation of feeding in insects. *Gen Comp Endocrinol,* 162**,** 93-104.

BANERJEE, K. K., DESHPANDE, R. S., KOPPULA, P., AYYUB, C. & KOLTHUR-SEETHARAM, U. 2017. Central metabolic sensing remotely controls nutrient-sensitive endocrine response in Drosophila via Sir2/Sirt1-upd2-IIS axis. *J Exp Biol,* 220**,** 1187-1191.

BECKER, A., SCHLODER, P., STEELE, J. E. & WEGENER, G. 1996. The regulation of trehalose metabolism in insects. *Experientia,* 52**,** 433-439.

BEDNAROVA, A., KODRIK, D. & KRISHNAN, N. 2013. Unique roles of glucagon and glucagon-like peptides: Parallels in understanding the functions of adipokinetic hormones in stress responses in insects. *Comp Biochem Physiol A Mol Integr Physiol,* 164**,** 91-100.

BEDNAROVA, A., TOMCALA, A., MOCHANOVA, M., KODRIK, D. & KRISHNAN, N. 2018. Disruption of adipokinetic hormone mediated energy homeostasis has subtle effects on physiology, Behavior and Lipid Status During Aging in Drosophila. *Front Physiol,* 9**,** 949.

BENDENA, W. G., DONLY, B. C., FUSE, M., LEE, E., LANGE, A. B., ORCHARD, I. & TOBE, S. S. 1997. Molecular characterization of the inhibitory myotropic peptide leucomyosuppressin. *Peptides,* 18**,** 157-63.

BILLINGSLEY, P. F. 1988. Morphometric analysis of Rhodnius prolixus Stal (Hemiptera:Reduviidae) midgut cells during blood digestion. *Tissue Cell,* 20**,** 291-301.

BLATT, J. & ROCES, F. 2001. Haemolymph sugar levels in foraging honeybees (Apis mellifera carnica): dependence on metabolic rate and in vivo measurement of maximal rates of trehalose synthesis. *J Exp Biol,* 204**,** 2709-16.

BLATT, J. & ROCES, F. 2002. The control of the proventriculus in the honeybee (Apis mellifera carnica L.) I. A dynamic process influenced by food quality and quantity? *J Insect Physiol,* 48**,** 643-654.

BOITEUX, A., HESS, B. & SEL'KOV, E. E. 1980. Creative functions of instability and oscillations in metabolic systems. *Curr Top Cell Regul,* 17**,** 171-203.

BRANCH, A. & SHEN, P. 2017. Central and peripheral regulation of appetite and food intake in Drosophila. *In:* ND & HARRIS, R. B. S. (eds.) *Appetite and Food Intake: Central Control.* Boca Raton (FL).

BRAY, G. A. 1991. Reciprocal relation between the sympathetic nervous system and food intake. *Brain Res Bull,* 27**,** 517-20.

CANNON, W. B. 1929. Organization for physiological homeostasis. *Physiol Rev,* 9**,** 32.

CARREIRO, A. L., DHILLON, J., GORDON, S., HIGGINS, K. A., JACOBS, A. G., MCARTHUR, B. M., REDAN, B. W., RIVERA, R. L., SCHMIDT, L. R. & MATTES, R. D. 2016. The Macronutrients, appetite, and energy Intake. *Annu Rev Nutr,* 36**,** 73-103.

CASTILLO, K., NASSIF, M., VALENZUELA, V., ROJAS, F., MATUS, S., MERCADO, G., COURT, F. A., VAN ZUNDERT, B. & HETZ, C. 2013. Trehalose delays the progression of amyotrophic lateral sclerosis by enhancing autophagy in motoneurons. *Autophagy,* 9**,** 1308-20.

CHAPELOT, D. & CHARLOT, K. 2019. Physiology of energy homeostasis: Models, actors, challenges and the glucoadipostatic loop. *Metabolism,* 92**,** 11-25.

CHEESEMAN, C. I. & PARSONS, D. S. 1976. The role of some small peptides in the transfer of amino nitrogen across the wall of vascularly perfused intestine. *J Physiol,* 262**,** 459-76.

CHEESEMAN, P. & GOLDSWORTHY, G. J. 1979. The release of adipokinetic hormone during flight and starvation in Locusta. *Gen Comp Endocrinol,* 37**,** 35-43.

CHEN, Y. D. & DAHANUKAR, A. 2018. DH44 neurons: gut-brain amino acid sensors. *Cell Res,* 28**,** 1048-1049.

CLEGG, J. S. & EVANS, D. R. 1961. Blood trehalose and flight metabolism in the blowfly. *Science,* 134**,** 54-5.

CLYNEN, E., BAGGERMAN, G., VEELAERT, D., CERSTIAENS, A., VAN DER HORST, D., HARTHOORN, L., DERUA, R., WAELKENS, E., DE LOOF, A. & SCHOOFS, L. 2001. Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, Locusta migratoria. *Eur J Biochem,* 268**,** 1929-39.

COGNIGNI, P., BAILEY, A. P. & MIGUEL-ALIAGA, I. 2011. Enteric neurons and systemic signals couple nutritional and reproductive status with intestinal homeostasis. *Cell Metab,* 13**,** 92-104.

CROWE, J. H., CARPENTER, J. F. & CROWE, L. M. 1998. The role of vitrification in anhydrobiosis. *Annu Rev Physiol,* 60**,** 73-103.

DONINI, A., NGO, C. & LANGE, A. B. 2002. Evidence for crustacean cardioactive peptide-like innervation of the gut in Locusta migratoria. *Peptides,* 23**,** 1915-23.

DUBOS, M. P., BERNAY, B. & FAVREL, P. 2017. Molecular characterization of an adipokinetic hormone-related neuropeptide (AKH) from a mollusk. *Gen Comp Endocrinol,* 243**,** 15-21.

DUS, M., LAI, J. S., GUNAPALA, K. M., MIN, S., TAYLER, T. D., HERGARDEN, A. C., GERAUD, E., JOSEPH, C. M. & SUH, G. S. 2015. Nutrient sensor in the brain directs the action of the brain-gut axis in Drosophila. *Neuron,* 87**,** 139-51.

FUKUMURA, K., KONUMA, T., TSUKAMOTO, Y. & NAGATA, S. 2018. Adipokinetic hormone signaling determines dietary fatty acid preference through maintenance of hemolymph fatty acid composition in the cricket Gryllus bimaculatus. *Sci Rep,* 8**,** 4737.

GALIKOVA, M., DIESNER, M., KLEPSATEL, P., HEHLERT, P., XU, Y. J., BICKMEYER, I., PREDEL, R. & KUHNLEIN, R. P. 2015. Energy homeostasis control in Drosophila Adipokinetic hormone mutants. *Genetics,* 201**,** 665-+.

GAVRIELI, A. & MANTZOROS, C. S. 2016. Novel molecules regulating energy homeostasis: physiology and regulation by macronutrient intake and weight loss. *Endocrinol Metab (Seoul),* 31**,** 361-372.

GLANTZ, R. M., MILLER, C. S. & NASSEL, D. R. 2000. Tachykinin-related peptide and GABA-mediated presynaptic inhibition of crayfish photoreceptors. *J Neurosci,* 20**,** 1780-90.

GOLDSWORTHY, G. J., OPOKU-WARE, K. & MULLEN, L. M. 2005. Adipokinetic hormone and the immune responses of locusts to infection. *Ann N Y Acad Sci,* 1040**,** 106-13.

GRILL, H. J. & HAYES, M. R. 2009. The nucleus tractus solitarius: a portal for visceral afferent signal processing, energy status assessment and integration of their combined effects on food intake. *Int J Obes (Lond),* 33 Suppl 1**,** S11-5.

HADLEY WH, TARO M, et al. *The crickets as a model organism: Development, Regeneration, and Behavior*. Springer. 2017.

HAND, S. C. & SOMERO, G. N. 1984. Influence of osmolytes, thin filaments, and solubility state on elasmobranch phosphofructokinase in vitro. *J Exp Zool,* 231**,** 297-302.

HAVULA, E., TEESALU, M., HYOTYLAINEN, T., SEPPALA, H., HASYGAR, K., AUVINEN, P., ORESIC, M., SANDMANN, T. & HIETAKANGAS, V. 2013. Mondo/ChREBP-Mlx-regulated transcriptional network is essential for dietary sugar tolerance in Drosophila. *PLoS Genet,* 9**,** e1003438.

HOLMES, F. L. 1986. Claude Bernard, the milieu interieur, and regulatory physiology. *Hist Philos Life Sci,* 8**,** 3-25.

HONDA, Y., TANAKA, M. & HONDA, S. 2010. Trehalose extends longevity in the nematode Caenorhabditis elegans. *Aging Cell,* 9**,** 558-69.

HU, C., PETERSEN, M., HOYER, N., SPITZWECK, B., TENEDINI, F., WANG, D., GRUSCHKA, A., BURCHARDT, L. S., SZPOTOWICZ, E., SCHWEIZER, M., GUNTUR, A. R., YANG, C. H. & SOBA, P. 2017. Sensory integration and neuromodulatory feedback facilitate Drosophila mechanonociceptive behavior. *Nat Neurosci,* 20**,** 1085-1095.

IBRAHIM, E., DOBES, P., KUNC, M., HYRSL, P. & KODRIK, D. 2018. Adipokinetic hormone and adenosine interfere with nematobacterial infection and locomotion in Drosophila melanogaster. *J Insect Physiol,* 107**,** 167-174.

IGNELL, R., ROOT, C. M., BIRSE, R. T., WANG, J. W., NASSEL, D. R. & WINTHER, A. M. 2009. Presynaptic peptidergic modulation of olfactory receptor neurons in Drosophila. *Proc Natl Acad Sci U S A,* 106**,** 13070-5.

INAGAKI, H. K., PANSE, K. M. & ANDERSON, D. J. 2014. Independent, reciprocal neuromodulatory control of sweet and bitter taste sensitivity during starvation in Drosophila. *Neuron,* 84**,** 806-20.

INGOLIA, N. T., BRAR, G. A., ROUSKIN, S., MCGEACHY, A. M. & WEISSMAN, J. S. 2012. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc,* 7**,** 1534-50.

JOHARD, H. A., COAST, G. M., MORDUE, W. & NASSEL, D. R. 2003. Diuretic action of the peptide locustatachykinin I: cellular localisation and effects on fluid secretion in Malpighian tubules of locusts. *Peptides,* 24**,** 1571-9.

KAHSAI, L., KAPAN, N., DIRCKSEN, H., WINTHER, A. M. & NASSEL, D. R. 2010. Metabolic stress responses in Drosophila are modulated by brain neurosecretory cells that produce multiple neuropeptides. *PLoS One,* 5**,** e11480.

KANAMORI, Y., SAITO, A., HAGIWARA-KOMODA, Y., TANAKA, D., MITSUMASU, K., KIKUTA, S., WATANABE, M., CORNETTE, R., KIKAWADA, T. & OKUDA, T. 2010. The trehalose transporter 1 gene sequence is conserved in insects and encodes proteins with different kinetic properties involved in trehalose import into peripheral tissues. *Insect Biochem Mol Biol,* 40**,** 30-7.

KAPAN, N., LUSHCHAK, O. V., LUO, J. & NASSEL, D. R. 2012. Identified peptidergic neurons in the Drosophila brain regulate insulin-producing cells, stress responses and metabolism by coexpressed short neuropeptide F and corazonin. *Cell Mol Life Sci,* 69**,** 4051-66.

KIM, S. K. & RULIFSON, E. J. 2004. Conserved mechanisms of glucose sensing and regulation by Drosophila corpora cardiaca cells. *Nature,* 431**,** 316-20.

KIM, S. M., SU, C. Y. & WANG, J. W. 2017. Neuromodulation of innate behaviors in Drosophila. *Annu Rev Neurosci,* 40**,** 327-348.

KIM, Y. J., ZITNAN, D., CHO, K. H., SCHOOLEY, D. A., MIZOGUCHI, A. & ADAMS, M. E. 2006a. Central peptidergic ensembles associated with organization of an innate behavior. *Proc Natl Acad Sci U S A,* 103**,** 14211-6.

KIM, Y. J., ZITNAN, D., GALIZIA, C. G., CHO, K. H. & ADAMS, M. E. 2006b. A command chemical triggers an innate behavior by sequential activation of multiple peptidergic ensembles. *Curr Biol,* 16**,** 1395-407.

KLEIN, S. & GORAN, M. 1993. Energy metabolism in response to overfeeding in young adult men. *Metabolism,* 42**,** 1201-5.

KO, K. I., ROOT, C. M., LINDSAY, S. A., ZANINOVICH, O. A., SHEPHERD, A. K., WASSERMAN, S. A., KIM, S. M. & WANG, J. W. 2015. Starvation promotes concerted modulation of appetitive olfactory behavior via parallel neuromodulatory circuits. *Elife,* 4.

KONUMA, T., MOROOKA, N., NAGASAWA, H. & NAGATA, S. 2012. Knockdown of the adipokinetic hormone receptor increases feeding frequency in the two-spotted cricket Gryllus bimaculatus. *Endocrinology,* 153**,** 3111-3122.

KOSTRON, B., KALTENHAUSER, U., SEIBEL, B., BR, AUML, UNIG, P. & HONEGGER, H. 1996. Localization of bursicon in CCAP-immunoreactive cells in the thoracic ganglia of the cricket Gryllus bimaculatus. *J Exp Biol,* 199**,** 367-77.

KUKURBA, K. R. & MONTGOMERY, S. B. 2015. RNA sequencing and analysis. *Cold Spring Harb Protoc,* 2015**,** 951-69.

LAHR, E. C., DEAN, D. & EWER, J. 2012. Genetic analysis of ecdysis behavior in Drosophila reveals partially overlapping functions of two unrelated neuropeptides. *J Neurosci,* 32**,** 6819-29.

LANCHA, A., FRUHBECK, G. & GOMEZ-AMBROSI, J. 2012. Peripheral signalling involved in energy homeostasis control. *Nutr Res Rev,* 25**,** 223-48.

LEE, G. & PARK, J. H. 2004. Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in Drosophila melanogaster. *Genetics,* 167**,** 311-23.

LIMMER, S., WEILER, A., VOLKENHOFF, A., BABATZ, F. & KLAMBT, C. 2014. The Drosophila blood-brain barrier: development and function of a glial endothelium. *Front Neurosci,* 8**,** 365.

LIU, S. & BORGLAND, S. L. 2015. Regulation of the mesolimbic dopamine circuit by feeding peptides. *Neuroscience,* 289**,** 19-42.

LO, C. M., OBICI, S., DONG, H. H., HAAS, M., LOU, D., KIM, D. H., LIU, M., D'ALESSIO, D., WOODS, S. C. & TSO, P. 2011. Impaired insulin secretion and enhanced insulin sensitivity in cholecystokinin-deficient mice. *Diabetes,* 60**,** 2000-7.

LOI, P. K., EMMAL, S. A., PARK, Y. & TUBLITZ, N. J. 2001. Identification, sequence and expression of a crustacean cardioactive peptide (CCAP) gene in the moth Manduca sexta. *J Exp Biol,* 204**,** 2803-16.

LORENZ, M. W. 2003. Adipokinetic hormone inhibits the formation of energy stores and egg production in the cricket Gryllus bimaculatus. *Comp Biochem Physiol B Biochem Mol Biol,* 136**,** 197-206.

LORENZ, M. W., KELLNER, R. & HOFFMANN, K. H. 1995. Identification of two allatostatins from the cricket, Gryllus bimaculatus de Geer (Ensifera, Gryllidae): additional members of a family of neuropeptides inhibiting juvenile hormone biosynthesis. *Regul Pept,* 57**,** 227-36.

MAIER, A. & ORION, B. 1969. [Open system ergometry with study of the various respiratory and circulatory parameters]. *Poumon Coeur,* 25**,** 609-26.

MARIC, G., GAZIBARA, T., ZALETEL, I., LABUDOVIC BOROVIC, M., TOMANOVIC, N., CIRIC, M. & PUSKAS, N. 2014. The role of gut hormones in appetite regulation (review). *Acta Physiol Hung,* 101**,** 395-407.

MATSUMOTO, S., OHARA, A., NAGAI-OKATANI, C., ZHOU, Y. J., FUJINAGA, D., SEIKE, H. & NAGATA, S. 2019. Antagonistic Effect of Short Neuropeptide F on Allatotropin-Inhibited Feeding Motivation of the Silkworm Larva, Bombyx mori. *Zoolog Sci,* 36**,** 58-67.

MATSUSHITA, R. & NISHIMURA, T. 2020. Trehalose metabolism confers developmental robustness and stability in Drosophila by regulating glucose homeostasis. *Commun Biol,* 3**,** 170.

MEISELMAN, M., LEE, S. S., TRAN, R. T., DAI, H., DING, Y., RIVERA-PEREZ, C., WIJESEKERA, T. P., DAUWALDER, B., NORIEGA, F. G. & ADAMS, M. E. 2017. Endocrine network essential for reproductive success in Drosophila melanogaster. *Proc Natl Acad Sci U S A,* 114**,** E3849-E3858.

MIGUEL-ALIAGA, I., JASPER, H. & LEMAITRE, B. 2018. Anatomy and Physiology of the Digestive Tract of Drosophila melanogaster. *Genetics,* 210**,** 357-396.

MIKANI, A., WANG, Q.-S. & TAKEDA, M. 2012. Brain-midgut short neuropeptide F mechanism that inhibits digestive activity of the American cockroach, Periplaneta americana upon starvation. *Peptides,* 34**,** 135-144.

MIKANI, A., WATARI, Y. & TAKEDA, M. 2015. Brain-midgut cross-talk and autocrine metabolastat via the sNPF/CCAP negative feed-back loop in the American cockroach, Periplaneta americana. *Cell Tissue Res,* 362**,** 481-96.

MILLER, T. A. 1997. Control of circulation in insects. *Gen Pharmacol,* 29**,** 23-38.

MIRABEAU, O. & JOLY, J. S. 2013. Molecular evolution of peptidergic signaling systems in bilaterians. *Proc Natl Acad Sci U S A,* 110**,** E2028-37.

MOYA-ANEGÓN, S. G. F. D., VARGAS-QUESADA, B., CHINCHILLA-RODRÍGUEZ, Z., CORERA-ÁLVAREZ, E., MUNOZ-FERNÁNDEZ, F. J. & HERRERO-SOLANA, V. 2007. Visualizing the marrow of science. *Journal of the American Society for Information Science and Technology,* 58**,** 2167-2179.

MURPHY, K. G., DHILLO, W. S. & BLOOM, S. R. 2006. Gut peptides in the regulation of food intake and energy homeostasis. *Endocr Rev,* 27**,** 719-27.

MURPHY, T. A. & WYATT, G. R. 1965. The enzymes of glycogen and trehalose synthesis in silk moth fat body. *J Biol Chem,* 240**,** 1500-8.

NAGATA, S., MATSUMOTO, S., NAKANE, T., OHARA, A., MOROOKA, N., KONUMA, T., NAGAI, C. & NAGASAWA, H. 2012. Effects of starvation on brain short neuropeptide F-1,-2, and-3 levels and short neuropeptide F receptor expression levels of the silkworm, Bombyx mori. *Frontiers in endocrinology,* 3**,** 3.

NAGATA, S., MOROOKA, N., MATSUMOTO, S., KAWAI, T. & NAGASAWA, H. 2011. Effects of neuropeptides on feeding initiation in larvae of the silkworm, Bombyx mori. *Gen Comp Endocrinol,* 172**,** 90-5.

NAKAZATO, M., MURAKAMI, N., DATE, Y., KOJIMA, M., MATSUO, H., KANGAWA, K. & MATSUKURA, S. 2001. A role for ghrelin in the central regulation of feeding. *Nature,* 409**,** 194-8.

NASSEL, D. R. 2018. Substrates for neuronal cotransmission with neuropeptides and small molecule neurotransmitters in Drosophila. *Front Cell Neurosci,* 12**,** 83.

NASSEL, D. R. & ZANDAWALA, M. 2019. Recent advances in neuropeptide signaling in Drosophila, from genes to physiology and behavior. *Prog Neurobiol,* 179**,** 101607.

NEUHUSER, T., SORGE, D., STAY, B. & HOFFMANN, K. H. 1994. Responsiveness of the adult cricket (Gryllus bimaculatus and Acheta domesticus) retrocerebral complex to allatostatin-1 from a cockroach, Diploptera punctata. *J Comp Physiol B,* 164**,** 23-31.

OH, Y., LAI, J. S., MILLS, H. J., ERDJUMENT-BROMAGE, H., GIAMMARINARO, B., SAADIPOUR, K., WANG, J. G., ABU, F., NEUBERT, T. A. & SUH, G. S. B. 2019. A glucose-sensing neuron pair regulates insulin and glucagon in Drosophila. *Nature,* 574**,** 559-564.

OOMURA, Y., OOYAMA, H., SUGIMORI, M., NAKAMURA, T. & YAMADA, Y. 1974. Glucose inhibition of the glucose-sensitive neurone in the rat lateral hypothalamus. *Nature,* 247**,** 284-6.

OOMURA, Y. & YOSHIMATSU, H. 1984. Neural network of glucose monitoring system. *J Auton Nerv Syst,* 10**,** 359-72.

PAGLIASSOTTI, M. J., ESTRADA, A. L., HUDSON, W. M., WEI, Y., WANG, D., SEALS, D. R., ZIGLER, M. L. & LAROCCA, T. J. 2017. Trehalose supplementation reduces hepatic endoplasmic reticulum stress and inflammatory signaling in old mice. *J Nutr Biochem,* 45**,** 15-23.

PARK, D. & TAGHERT, P. H. 2009. Peptidergic neurosecretory cells in insects: organization and control by the bHLH protein DIMMED. *Gen Comp Endocrinol,* 162**,** 2-7.

PARK, J. H., CHEN, J., JANG, S., AHN, T. J., KANG, K., CHOI, M. S. & KWON, J. Y. 2016. A subset of enteroendocrine cells is activated by amino acids in the Drosophila midgut. *FEBS Lett,* 590**,** 493-500.

PARK, J. H., SCHROEDER, A. J., HELFRICH-FORSTER, C., JACKSON, F. R. & EWER, J. 2003. Targeted ablation of CCAP neuropeptide-containing neurons of Drosophila causes specific defects in execution and circadian timing of ecdysis behavior. *Development,* 130**,** 2645-56.

PARK, S., BUSTAMANTE, E. L., ANTONOVA, J., MCLEAN, G. W. & KIM, S. K. 2011. Specification of Drosophila corpora cardiaca neuroendocrine cells from mesoderm is regulated by Notch signaling. *PLoS Genet,* 7**,** e1002241.

PASCO, M. Y. & LEOPOLD, P. 2012. High sugar-induced insulin resistance in Drosophila relies on the lipocalin Neural Lazarillo. *PLoS One,* 7**,** e36583.

PRICE, D. A. & GREENBERG, M. J. 1977. Structure of a molluscan cardioexcitatory neuropeptide. *Science,* 197**,** 670-1.

REHFELD, J. F., LARSSON, L. I., GOLTERMANN, N. R., SCHWARTZ, T. W., HOLST, J. J., JENSEN, S. L. & MORLEY, J. S. 1980. Neural regulation of pancreatic hormone secretion by the C-terminal tetrapeptide of CCK. *Nature,* 284**,** 33-8.

REN, G. R., HAUSER, F., REWITZ, K. F., KONDO, S., ENGELBRECHT, A. F., DIDRIKSEN, A. K., SCHJOTT, S. R., SEMBACH, F. E., LI, S., SOGAARD, K. C., SONDERGAARD, L. & GRIMMELIKHUIJZEN, C. J. 2015. CCHamide-2 is an orexigenic brain-gut peptide in Drosophila. *PLoS One,* 10**,** e0133017.

RICHARD, D. 2015. Cognitive and autonomic determinants of energy homeostasis in obesity. *Nat Rev Endocrinol,* 11**,** 489-501.

ROLLS, B. J. 2009. The relationship between dietary energy density and energy intake. *Physiol Behav,* 97**,** 609-15.

ROOT, C. M., KO, K. I., JAFARI, A. & WANG, J. W. 2011. Presynaptic Facilitation by Neuropeptide Signaling Mediates Odor-Driven Food Search. *Cell,* 145**,** 133-144.

SAKAI, T., SATAKE, H., MINAKATA, H. & TAKEDA, M. 2004. Characterization of crustacean cardioactive peptide as a novel insect midgut factor: isolation, localization, and stimulation of alpha-amylase activity and gut contraction. *Endocrinology,* 145**,** 5671-8.

SAKAI, T., SATAKE, H. & TAKEDA, M. 2006. Nutrient-induced alpha-amylase and protease activity is regulated by crustacean cardioactive peptide (CCAP) in the cockroach midgut. *Peptides,* 27**,** 2157-64.

SANO, H., NAKAMURA, A., TEXADA, M. J., TRUMAN, J. W., ISHIMOTO, H., KAMIKOUCHI, A., NIBU, Y., KUME, K., IDA, T. & KOJIMA, M. 2015. The Nutrient-responsive hormone CCHamide-2 controls growth by regulating insulin-like peptides in the brain of Drosophila melanogaster. *PLoS Genet,* 11**,** e1005209.

SAYIN, S., BOEHM, A. C., KOBLER, J. M., DE BACKER, J. F. & GRUNWALD KADOW, I. C. 2018. Internal state dependent odor processing and perception-the role of neuromodulation in the fly olfactory system. *Front Cell Neurosci,* 12**,** 11.

SCOPELLITI, A., BAUER, C., YU, Y., ZHANG, T., KRUSPIG, B., MURPHY, D. J., VIDAL, M., MADDOCKS, O. D. K. & CORDERO, J. B. 2019. A neuronal relay mediates a nutrient responsive gut/fat body axis regulating rnergy homeostasis in adult Drosophila. *Cell Metab,* 29**,** 269-284 e10.

SIVITER, R. J., COAST, G. M., WINTHER, A. M., NACHMAN, R. J., TAYLOR, C. A., SHIRRAS, A. D., COATES, D., ISAAC, R. E. & NASSEL, D. R. 2000. Expression and functional characterization of a Drosophila neuropeptide precursor with homology to mammalian preprotachykinin A. *J Biol Chem,* 275**,** 23273-80.

SONG, W., VEENSTRA, J. A. & PERRIMON, N. 2014. Control of lipid metabolism by tachykinin in Drosophila. *Cell Rep,* 9**,** 40-47.

STANGIER, J., HILBICH, C., BEYREUTHER, K. & KELLER, R. 1987. Unusual cardioactive peptide (CCAP) from pericardial organs of the shore crab Carcinus maenas. *Proc Natl Acad Sci U S A,* 84**,** 575-9.

THOMPSON, S. N. 2003. Trehalose - The insect 'blood' sugar. *Advances in Insect Physiology, Vol 31,* 31**,** 205-285.

THOMPSON, S. N., REDAK, R. A. & WANG, L. W. 2001. Altered dietary nutrient intake maintains metabolic homeostasis in parasitized larvae of the insect Manduca sexta L. *J Exp Biol,* 204**,** 4065-80.

TSCHOP, M., SMILEY, D. L. & HEIMAN, M. L. 2000. Ghrelin induces adiposity in rodents. *Nature,* 407**,** 908-13.

TSUKAMOTO, Y. & NAGATA, S. 2016. Newly identified allatostatin Bs and their receptor in the two-spotted cricket, Gryllus bimaculatus. *Peptides,* 80**,** 25-31.

TUBLITZ, N. J. & EVANS, P. D. 1986. Insect cardioactive peptides: cardioacceleratory peptide (CAP) activity is blocked in vivo and in vitro with a monoclonal antibody. *J Neurosci,* 6**,** 2451-6.

UGRANKAR, R., BERGLUND, E., AKDEMIR, F., TRAN, C., KIM, M. S., NOH, J., SCHNEIDER, R., EBERT, B. & GRAFF, J. M. 2015. Drosophila glucome screening identifies Ck1alpha as a regulator of mammalian glucose metabolism. *Nat Commun,* 6**,** 7102.

VAN ECK, N. J. & WALTMAN, L. 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics,* 84**,** 523-538.

VAUGHAN, M. 1998. Signaling by heterotrimeric G proteins minireview series. *J Biol Chem,* 273**,** 667-8.

VEELAERT, D., PASSIER, P., DEVREESE, B., VANDEN BROECK, J., VAN BEEUMEN, J., VULLINGS, H. G., DIEDEREN, J. H., SCHOOFS, L. & DE LOOF, A. 1997. Isolation and characterization of an adipokinetic hormone release-inducing factor in locusts: the crustacean cardioactive peptide. *Endocrinology,* 138**,** 138-42.

VEENSTRA, J. A. 1989. Do insects really have a homeostatic hypotrehalosaemic hormone? *Biol Rev Camb Philos Soc,* 64**,** 305-16.

VOLKENHOFF, A., HIRRLINGER, J., KAPPEL, J. M., KLAMBT, C. & SCHIRMEIER, S. 2018. Live imaging using a FRET glucose sensor reveals glucose delivery to all cell types in the Drosophila brain. *J Insect Physiol,* 106**,** 55-64.

VOLKENHOFF, A., WEILER, A., LETZEL, M., STEHLING, M., KLAMBT, C. & SCHIRMEIER, S. 2015. Glial glycolysis is essential for neuronal survival in Drosophila. *Cell Metab,* 22**,** 437-47.

VULLINGS, H. G., DIEDEREN, J. H., VEELAERT, D. & VAN DER HORST, D. J. 1999. Multifactorial control of the release of hormones from the locust retrocerebral complex. *Microsc Res Tech,* 45**,** 142-53.

WANG, S., TULINA, N., CARLIN, D. L. & RULIFSON, E. J. 2007. The origin of islet-like cells in Drosophila identifies parallels to the vertebrate endocrine axis. *Proc Natl Acad Sci U S A,* 104**,** 19873-8.

WASIELEWSKI, O. & SKONIECZNA, M. 2008. Pleiotropic effects of the neuropeptides CCAP and myosuppressin in the beetle, Tenebrio molitor L. *J Comp Physiol B,* 178**,** 877-85.

WEN, X., WANG, S., DUMAN, J. G., ARIFIN, J. F., JUWITA, V., GODDARD, W. A., 3RD, RIOS, A., LIU, F., KIM, S. K., ABROL, R., DEVRIES, A. L. & HENLING, L. M. 2016. Antifreeze proteins govern the precipitation of trehalose in a freezing-avoiding insect at low temperature. *Proc Natl Acad Sci U S A,* 113**,** 6683-8.

WICHER, D., AGRICOLA, H. J., SOHLER, S., GUNDEL, M., HEINEMANN, S. H., WOLLWEBER, L., STENGL, M. & DERST, C. 2006. Differential receptor activation by cockroach adipokinetic hormones produces differential effects on ion currents, neuronal activity, and locomotion. *J Neurophysiol,* 95**,** 2314-25.

WILLIAMS, M. J., AKRAM, M., BARKAUSKAITE, D., PATIL, S., KOTSIDOU, E., KHEDER, S., VITALE, G., FILAFERRO, M., BLEMINGS, S. W., MAESTRI, G., HAZIM, N., VERGONI, A. V. & SCHIOTH, H. B. 2020. CCAP regulates feeding behavior via the NPF pathway in Drosophila adults. *Proc Natl Acad Sci U S A,* 117**,** 7401-7408.

WINTHER, A. M., SIVITER, R. J., ISAAC, R. E., PREDEL, R. & NASSEL, D. R. 2003. Neuronal expression of tachykinin-related peptides and gene transcript during postembryonic development of Drosophila. *J Comp Neurol,* 464**,** 180-96.

WOODS, S. C. & RAMSAY, D. S. 2007. Homeostasis: beyond Curt Richter. *Appetite,* 49**,** 388-98.

WYATT, G. R. & KALE, G. F. 1957. The chemistry of insect hemolymph. II. Trehalose and other carbohydrates. *J Gen Physiol,* 40**,** 833-47.

YAMANAKA, N., YAMAMOTO, S., ZITNAN, D., WATANABE, K., KAWADA, T., SATAKE, H., KANEKO, Y., HIRUMA, K., TANAKA, Y., SHINODA, T. & KATAOKA, H. 2008. Neuropeptide receptor transcriptome reveals unidentified neuroendocrine pathways. *PLoS One,* 3**,** e3048.

YAO, Z. & SCOTT, K. 2020. Glucose-sensing neurons reciprocally regulate insulin and glucagon. *Trends Neurosci,* 43**,** 2-5.

YU, Y., HUANG, R., YE, J., ZHANG, V., WU, C., CHENG, G., JIA, J. & WANG, L. 2016. Regulation of starvation-induced hyperactivity by insulin and glucagon signaling in adult Drosophila. *Elife,* 5.

ZANDAWALA, M., HADDAD, A. S., HAMOUDI, Z. & ORCHARD, I. 2015. Identification and characterization of the adipokinetic hormone/corazonin-related peptide signaling system in Rhodnius prolixus. *FEBS J,* 282**,** 3603-17.

ZÖLLNER, N. & KIRSCH, K. 1962. On the quantitative determination of lipids (micromethod) by means of the general sulfophosphovanillin reaction of the may natural lipids (all know plasmalipids). *Z ges exp Med,* 135**,** 545.