

Novel Tools and Methods

Identification of Immediate Early Genes in the Nervous System of Snail *Helix lucorum*

Chuan Xu,^{1*} Qian Li,^{1*} Olga Efimova,² Xi Jiang,¹ Marina Petrova,¹ Alia K. Vinarskaya,³ Peter Kolosov,³ Nikolay Aseyev,³ Kira Koshkareva,² Victor N. Ierusalimsky,³  Pavel M. Balaban,³ and Philipp Khaitovich^{2,4,5,6}

<https://doi.org/10.1523/ENEURO.0416-18.2019>

¹CAS Key Laboratory of Computational Biology, Chinese Academy of Sciences-Max Planck Gesellschaft Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China, ²Skolkovo Institute of Science and Technology, Moscow 143026, Russia, ³Institute of Higher Nervous Activity and Neurophysiology, Moscow 117485, Russia, ⁴Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China, ⁵Comparative Biology Laboratory, Chinese Academy of Sciences-Max Planck Gesellschaft Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China, and ⁶School of Life Science and Technology, ShanghaiTech University, Shanghai 200031, China

Abstract

Immediate early genes (IEGs) are useful markers of neuronal activation and essential components of neuronal response. While studies of gastropods have provided many insights into the basic learning and memory mechanisms, the genome-wide assessment of IEGs has been mainly restricted to vertebrates. In this study, we identified IEGs in the terrestrial snail *Helix lucorum*. In the absence of the genome, we conducted *de novo* transcriptome assembly using reads with short and intermediate lengths cumulatively covering more than 98 billion nucleotides. Based on this assembly, we identified 37 proteins corresponding to contigs differentially expressed (DE) in either the parietal ganglia (PaG) or two giant interneurons located within the PaG of the snail in response to the neuronal stimulation. These proteins included homologues of well-known mammalian IEGs, such as *c-jun/jund*, *C/EBP*, *c-fos/fosl2*, and *Egr1*, as well as homologues of genes not yet implicated in the neuronal response.

Key words: snail; immediate early genes; *Helix lucorum*; nervous system

Significance Statement

Gastropods, which include snails and slugs, are widely used in studies of basic neuronal activity mechanisms. The first step of the transcriptional response to neuronal stimulation requires the activation of immediate early genes (IEGs). The identification of IEGs is important for the understanding of neuronal response mechanisms and for the visualization of activated neurons. However, genome-wide studies of IEGs have thus far been mainly restricted to vertebrates. Furthermore, a study of activity-regulated genes (ARGs), a gene group that includes IEGs, conducted in fruit flies did not reveal a clear overlap with vertebrate IEGs. In this study we present a transcriptome-wide study of snail IEGs, which reveals multiple homologues of well-known mammalian IEGs, as well as a number of novel IEG candidates.

Introduction

Immediate early genes (IEGs) are the composite group of genes rapidly and transiently upregulated in neuronal

cells in response to stimulation. An induction of IEGs does not require *de novo* protein synthesis and commonly occurs within an hour after the stimulation event (Fowler

Received October 27, 2018; accepted March 17, 2019; First published May 3, 2019.

The authors declare no competing financial interests.

Author contributions: C.X. and Q.L. analyzed data; C.X., Q.L., P.M.B., and P. Khaitovich wrote the paper; O.E., X.J., M.P., A.K.V., P. Kolosov, N.A., K.K., and V.N.I. performed research; P.M.B. and P. Khaitovich designed research.

et al., 2011). Stimulation experiments conducted in the presence of protein synthesis inhibitors revealed hundreds of mammalian IEGs, with most works conducted in mouse brain preparations (Thompson et al., 2010; Bojovic et al., 2015; Gerstner et al., 2016) and neuronal cultures (Kim et al., 2010; Saha et al., 2011; Spiegel et al., 2014). Although different types of stimulations induce different sets of IEGs, many genes were repeatedly found in most experiments (Dahmen et al., 1997; Bepari et al., 2012; Lacar et al., 2016). Many commonly induced IEGs, including the first identified one, *c-fos*, as well as *c-jun*, *C/EBP*, and *Egr1*, function as transcription factors triggering the expression of the secondary response genes (SRGs; West and Greenberg, 2011). As well as transcription factors, neuronal IEGs encode other functional proteins, such as cytoskeletal regulators (*Arc*), growth factors (β -actinin), metabolic enzymes (*Dusp1*), and signal transduction proteins (*Homer*; Lanahan and Worley, 1998).

While mammalian and vertebrate IEGs are relatively well characterized, less is known about IEGs in invertebrate species. Whereas several studies have been performed exploring genes involved in the long-term memory in the *Caenorhabditis elegans* (Lakhina et al., 2015; Freytag et al., 2017) and *Aplysia kurodai* (Lee et al., 2008), the genome-wide systematic investigation focusing on IEGs in invertebrates to date has been restricted to fruit flies, with two studies characterizing neuronal activity-regulated genes (ARGs). Similar to IEGs, ARGs are defined as genes induced in neurons within approximately 1 h after the stimulation, but without *de novo* protein synthesis inhibition. The first study using microarrays to assess the transcriptome alterations in heads of mutant flies after the seizure induction identified 122 genes showing rapid expression changes. These genes included fly homologs of known mammalian IEGs, such as *c-fos*, *c-jun*, *C/EBP*, and *Egr* (Guan et al., 2005). The second study used transcriptome sequencing (RNA-Seq) to look for ARGs in the fly brain and various neuron types after the activation by three stimulation protocols (Chen et al., 2016). Although the study identified known insect IEGs of *hr38* and *sr* induced in the fly brain by the three stimulation protocols, there was no detectable activation of many mammalian IEGs' homologues, including *c-fos* and *c-jun*. Instead, largely independent sets of genes, ~100 each, were induced by each stimulation procedure.

While no genome-wide screens for neuronal IEGs were conducted in other invertebrate species, previous studies

identified a number of individual genes. Specifically, the analysis of IEGs in a sea slug *Aplysia californica*, an organism widely used in the studies of memory mechanisms, identified homologues of mammalian IEGs *c-jun*, *C/EBP*, *CREB1*, and *Egr* (Alberini et al., 1994; Sung et al., 2006; Bonnicksen et al., 2012; Cyriac et al., 2013). In addition, studies conducted in the terrestrial slug *Limax valentianus* identified homologues of mammalian IEGs *C/EBP* and *KLF* (Fukunaga et al., 2006). Similarly, homologues of mammalian IEGs *C/EBP*, *CREB1*, and *CREB2* were identified in the extensively studied pond snail *Lymnaea stagnalis* (Sadamoto et al., 2004, 2010; Hatakeyama et al., 2006).

Here, we conducted a broad search for neuronal IEGs in another model invertebrate species, the terrestrial snail *Helix lucorum*. The nervous system of this organism containing five pairs of neuronal ganglia and one unpaired visceral ganglion has been used in electrophysiological studies for >35 years, yielding insights into basic learning and memory mechanisms (Balaban, 1980, 2002; Balaban et al., 2015). However, the absence of the genome sequence has limited molecular studies of the neuronal response mechanisms. To overcome this limitation, we conducted the *de novo* assembly of the snail neuronal transcriptome using >943 million reads cumulatively covering >98 billion nucleotides. The following analysis yielded 37 putative snail IEGs, including homologues of well-characterized mammalian ones: *c-jun/jund*, *C/EBP*, *c-fos/fos2*, *Egr1*, *ler5l*, *Socs2*, and *Dusp10*.

Materials and Methods

Sample preparations and RNA-Seq

We conducted experiments using adult *H. lucorum taurica* L. specimens weighing 30–35 g. The snails were kept in a wet environment and fed their usual diet of pieces of lettuce. The experimental procedures were in compliance with the Guide for the Care and Use of Laboratory Animals published by the National Institutes of Health and the protocol was approved by the Ethical Committee of the Institute of Higher Nervous Activity and Neurophysiology RAS. Before the experiment, the snails were kept in the active state for at least two weeks. Details of the preparation and identification of neurons are given elsewhere (Malyshev and Balaban, 2002). Briefly, animals were cooled to 4°C and injected with isotonic MgCl₂ before the CNS isolation to minimize pain. The central ganglia complex was surgically isolated from anesthetized snails, pinned to a silicone-elastomer (Sylgard)-coated dish, and kept in high-Ca²⁺, high-Mg²⁺ Ringer saline (80 mM NaCl, 4 mM KCl, 28 mM CaCl₂, 25 mM MgCl₂, and 10 mM Trisma; pH 7.6) to suppress electrical activity of the nervous system. Such treatment blocks electrical activity in the CNS and neuromuscular connections (Balaban and Chase, 1989), thereby minimizing the dissection effect on IEGs' expression in our experiments. The central ganglia complex was stripped of connective tissue sheath using a fine forceps and scissors, ensuring the integrity of the thin layer adherent to the neurons, kept at 4°C for 24 (\pm 1) h to further minimize any possible effect of dissection on IEGs' expression, and then kept at room temperature for 1 h.

This work was supported by the Russian Science Foundation Grant 16-14-00220 (to Q.L., O.E., V.N.I., P.M.B., and P.Kh.) and by the National Key R&D Program of China Grant 2017YFA0505700.

[^]C.X. and Q.L. contributed equally to this work.

Acknowledgements: We thank Zhisong He for his assistance, Grace Cudihy for her comments on this manuscript, and all members of the Comparative Biology Group in Shanghai for their helpful suggestions.

Correspondence should be addressed to Pavel M. Balaban at pmbalaban@gmail.com or Philipp Khaitovich at khaitovich@eva.mpg.de
<https://doi.org/10.1523/ENEURO.0416-18.2019>

Copyright © 2019 Xu et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Table 1. Sample information for *H. lucorum*

Sample ID	Experiment	Hemisphere	Condition	Individual
Sample1	E1	Right	Control	1
Sample2	E1	Left	Activated	1
Sample3	E1	Right	Control	2
Sample4	E1	Left	Activated	2
Sample5	E1	Right	Control	3
Sample6	E1	Left	Activated	3
Sample7	E1	Left	Control	4
Sample8	E1	Right	Activated	4
Sample9	E1	Left	Control	5
Sample10	E1	Right	Activated	5
Sample11	E1	Left	Control	6
Sample12	E1	Right	Activated	6
Sample13	E2	Right	Activated	7
Sample14	E2	Right	Control	8
Sample15	E2	Left	Activated	8
Sample16	E2	Left	Control	9
Sample17	E2	Right	Activated	9
Sample18	E2	Right	Control	10
Sample19	E2	Left	Activated	10
Sample20	E2	Right	Control	11
Sample21	E2	Left	Activated	11
Sample22	E2	Left	Control	12
Sample23	E2	Right	Control	13
Sample24	E2	Left	Activated	13

Part of CNS which was used as a self-control was cut out, transferred to the dry-ice cooled plastic tube, and frozen at -80°C . The remaining part was washed with 50 ml of normal Ringer saline (80 mM NaCl, 4 mM KCl, 8 mM CaCl_2 , 5 mM MgCl_2 , and 10 mM Trisma; pH 7.6) containing 20 μM anisomycin, and kept in this solution for 10 min, then was washed with 50 ml of normal Ringer saline containing 20 μM anisomycin, 100 μM caffeine and 5 μM 5-HT, and kept in this solution for 20–25 min. Part of activated CNS symmetrical to the control part was cut out, transferred to the dry-ice cooled plastic tube, and frozen at -80°C .

In the first experiment, we isolated the entire parieto-visceral complex. Self-control sets included three left parietal (+adhered visceral) ganglia and three right parietal ganglia (PaG). Activated CNS sets included three right PaG and three left parietal (+adhered visceral) ganglia, respectively. In the second experiment, a small medial part of the PaG mainly containing bodies of the two giant premotor interneurons Pa2 and Pa3 were surgically dissected. Self-control sets included similarly dissected parts of symmetrical ganglia with two left parietal Pa2+Pa3 neurons and four right Pa2+Pa3 neurons. Activated CNS sets included two right parietal Pa2+Pa3 neurons and four left parietal Pa2+Pa3 neurons, respectively (Table 1).

For RNA-Seq, we prepared cellular RNA from samples using the guanidine thiocyanate method (Chomczynski and Sacchi, 1987). A total of 24 RNA samples were analyzed using Agilent 2100 Bioanalyzer to confirm the RNA isolation purity and absence of RNA degradation. The peak of 28S rRNA is invisible in some species of snails because their 28S rRNA consists of two separate pieces held together by ribosome proteins, and after purification

Table 2. Read information and mapping summary of the 24 samples

Sample ID	Read length (nt)	Total reads	Mapped reads	Mapped rate (%)
Sample1	100	47132146	28173764	59.78
Sample2	100	10237649	6171779	60.29
Sample3	100	8472972	5174050	61.07
Sample4	100	32281924	19199711	59.48
Sample5	100	11572132	7005917	60.54
Sample6	100	40469603	24695241	61.02
Sample7	100	40876298	24490556	59.91
Sample8	100	19487663	11845560	60.78
Sample9	100	4677121	2798683	59.84
Sample10	100	16178231	9788917	60.51
Sample11	100	4742934	2913221	61.42
Sample12	100	6208915	3810534	61.37
Sample13	100	12194698	7185438	58.92
Sample14	100	15605531	9151020	58.64
Sample15	100	10636031	5895842	55.43
Sample16	100	11168095	6623440	59.31
Sample17	100	1181972	754320	63.82
Sample18	100	26521969	16231015	61.2
Sample19	100	25234534	15249538	60.43
Sample20	100	11450444	6902634	60.28
Sample21	100	38410752	22487496	58.54
Sample22	100	11772575	7275845	61.8
Sample23	100	36248990	21424916	59.1
Sample24	100	17470225	10952305	62.69

each half of 28S rRNA has the same length as 18S rRNA, so 28S peak merges with 18S peak. We therefore checked only the 18S peak integrity to estimate the total RNA quality.

A total of 500 ng RNA of each sample was depleted with rRNA Removal Mix (Ribo-Zero Human/Mouse/Rat) kit. cDNA preparations from RNA samples were performed using TruSeq Stranded Total RNA Sample Preparation kit (Illumina) following the supplier's instruction. Briefly, RNAs were fragmented to 120–200 bp with a median size of 150 bp and reverse transcribed using random hexamers and SuperScript II Reverse Transcriptase. Single stranded cDNAs were converted to double stranded cDNAs. End repair protocol and subsequent adenine nucleoside addition to 5'-end of DNA were made for ligation of barcoded adapters. The quality of each prepared cDNA library was evaluated using Qubit 2.0 Fluorometer (with Qubit dsDNA HS Assay kit) and Agilent 2100 Bioanalyzer (Agilent High Sensitivity DNA kit). The amount of short cDNA fragments in our samples with length of 25–160 bp did not exceed 10%. Sequencing was performed using the HiSeq Illumina platform (Table 2). One load contained 10–12 pooled libraries tagged with different barcodes.

Transcriptome assembly

The transcriptome was assembled from the dataset mentioned above, as well as two additional RNA-Seq datasets measured in *H. lucorum* (Table 3), by SOAPdenovo-Trans (version 1.04; Xie et al., 2014) using the recommended parameters, after the adapter sequence removal by cutadapt (version 1.11; Martin, 2011).

Table 3. Read pair information of additional samples

Sample ID	Total read pairs	Read length (nt)
SampleO1	31378888	101
SampleO2	31246438	101
SampleO3	26933069	101
SampleO4	28013479	101
SampleK1	13543270	101
SampleK2	24181331	101
SampleK3	28520249	101
SampleK4	45234073	101
SampleS1	4594007	251
SampleS2	2917579	251
SampleS3	5189873	251

Assessment of transcriptome assembly quality

The quality of the overall snail neuronal transcriptome assembly was assessed by Transrate (version 1.0.3; Smith-Unna et al., 2016), including N30/50/70 (the contig size at which 30/50/70% of bases are contained in contigs with greater sizes), GC content (percentage of nitrogenous bases), ORF percentage, and reciprocal best BLAST matches between the assembly and proteins of *A. californica* (California sea slug; Knudsen et al., 2006).

Quantification of contig expression levels

Raw reads from the 24 snail samples were mapped to the transcriptome assembly using bowtie (version 1.0.0) with the specifically chosen parameters by RSEM (version 1.2.18; Li and Dewey, 2011). We took advantage of the effective handling of ambiguously-mapping reads and the absence of reference genome implemented in RSEM to locate the expression abundance (raw count) at the contig level. In each experiment, only contigs with total counts across all 12 samples >10 were determined as expressed and used in the following analyses. Considering the intrinsically comparable character among samples, we normalized the abundance data using trimmed mean of M-values normalization method (TMM) provided by the Bioconductor package “edgeR” in R (Robinson et al., 2010). Moreover, for proper comparisons among genes, reads per kilobase per million mapped reads (RPKM) was subsequently obtained for each contig in each sample in consideration of the contig’s effective length.

Global pattern exploration

We performed multidimensional scaling (MDS) analyses to explore the global patterns of the snail samples in both experiments on the basis of sample dissimilarities defined as one minus Spearman’s rank correlation coefficient between pairwise samples based on expressed contigs using the “cmdscale” function in R.

Differential expression analysis

Differentially expressed (DE) contigs between activated and control samples in each experiment were identified using the Bioconductor package edgeR in R (Robinson et al., 2010). DE contigs were determined by the criteria of false discovery rate (FDR) <0.05 and fold change > 2. The significance of overlap of DE contigs between the two experiments was assessed by a hypergeometric test us-

ing the “phyper” function in R, with all expressed contigs in the first or second experiment as the statistical background.

Contig annotation

We resolved the homologous proteins of expressed contigs by the sequence search using BLASTX (version 2.2.24). Proteomes of five species were used as the search database: *A. californica* (Knudsen et al., 2006), *Biomphalaria glabrata* (VectorBase, version 1.2), *C. elegans* (Ensembl, release 77), *Drosophila melanogaster* (Ensembl, release 77), and *Takifugu rubripes* (Ensembl, release 77). Alignments with e-values below 1e-5 were selected as valid. If one contig was aligned to multiple proteins, we considered only the best alignment.

To define the consistent annotation across the five species, we performed BLASTP (version 2.2.24) between proteomes of pairwise species. The top five alignments with e-values below 1e-5 were considered as homologous. One contig was considered to be consistently annotated across species when the aligned protein in one species (e.g., the sea slug) and each of the aligned proteins in the other species (e.g., the remaining four species) were homologous based on the above criterion.

For each protein, we determined the direction of its change between activated and control samples by considering the expression changes of all the contigs mapped to the protein (both DE and non-DE contigs) in two experiments or one experiment. Specifically, for each protein, we calculated the percentage of contigs which were upregulated/downregulated in both experiments (E1 and E2) or in either experiment (E1 or E2) after activation out of all its mapped contigs. Under both experiments or one experiment, we next considered proteins with either upregulated or downregulated contigs’ percentage no <80% (i.e., no <80% of its mapped contigs showed consistent direction of expression change). The percentage of such proteins was then calculated and the significance of this percentage was estimated by randomly sampling 1000 times the same number of expressed contigs to be mapped to the protein, with other procedures being constant.

Design of *in situ* probe

The probe, which uniquely targeted the contig homologous to mammalian *fos* family IEGs *c-fos* and *fosl2* and fruit fly IEG *kayak* (*c-fos/fosl2* probe), was designed in an intron-spanned manner for the purpose of hybridizing only the transcript instead of the DNA. To find the junction site of the corresponding contig, we aligned the contig to the gene sequence of *c-Fos* in mice by BLASTN in Ensembl. With the aid of the gene structure and alignment visualization implemented in Ensembl, we chose the subset of sequence around the junction site from the contig as the probe (Table 4). Along with this criterion, a compromise was reached in consideration of the probe’s length (~50 nt), GC content (40–50%) for an appropriate hybridization.

Table 4. Information of the *c-fos/fos2* probe used for *in situ* hybridization

Contig ID	Sequence of the probe	Length	GC content
1398029	TCGACCATGCTTTGCTTTTGGCTCCCATTCTGCATCAGAATATTCGG	47	46.81%

Electrophysiological experiment

Sample preparations were the same as the preparations before RNA-Seq. The CNS was treated with proteinase Type XIV for 2 min, followed by washing using Ringer saline (80 mM NaCl, 4 mM KCl, 7 mM CaCl₂, 5 mM MgCl₂, and 10 mM Tris; pH 7.8). We recorded intracellularly the activity of the readily identifiable giant premotor neurons located in both PaG involved in triggering withdrawal (RPa2, RPa3, LPa2, LPa3). Microelectrodes were filled with 2 M potassium acetate and had the conductance 10–30 MOhm. Nerve stimulations were performed via the polyethylene suction electrodes. The experimental protocol included the nerve stimulation with 10-stimulus trains (10 Hz) once per 30 s, the intensity of stimulation chosen to cause the appearance of several action potentials in the giant premotor neurons at the beginning of the training. The total time of stimulation was 2 h. Intracellular signals were recorded with preamplifiers (Axoclamp 2B, Molecular Devices), digitized and stored on a computer (Digidata 1400 A A/D converter and Axoscope 10.0 software, both from Molecular Devices).

Nerve backfills

To detect the population of neurons projecting to the stimulated nerve and to compare it with the population of neurons expressing genes of interest (here *c-fos/fos2*), we performed the retrograde labeling of the neurons in the CNS. The cut end of the nerve was sucked into a pipette filled with 10% neurobiotin in 0.1 M KCl. The end of the nerve was left in place for 12–24 h at 18–22°C. The time of backfill was chosen experimentally. The ganglia were fixed with 4% paraformaldehyde in 0.1 M phosphate buffer for 2 h. The brains were processed with the Vectastain ABC kit.

In situ hybridization

For *in situ* hybridization, the CNS was processed as whole-mounts. The experimental procedure was described earlier (Balaban et al., 2001). Differently, as the short probe to the RNA of interest was used, we slightly modified the protocol. That is, the pre-hybridization was conducted at 50°C, and the hybridization itself was conducted at 37°C, with other procedure details being constant.

RT-qPCR

Five snails were used for RT-qPCR, with the sample preparation conducted as described above in Sample preparations and RNA-Seq. RNA was extracted from snail's CNS with PureLink RNA Mini kit (Ambion) according to manufacturer's recommendations. One RNA extraction failed for the stimulated part of the snail CNS. The quality of purified RNA was examined with Nanodrop spectrophotometer and 1 µg of RNA was taken for DNase I (Thermo Scientific) treatment conducted according to

the manufacturer's recommendations. DNA-free RNA was subsequently transcribed into cDNA for RT-qPCR procedure using Maxima First Strand cDNA Synthesis kit (Thermo Scientific). Quantitative PCR was performed using QuantStudio 3 Real-Time PCR System (Applied Biosystems) in triplicates for each sample. The protocol used SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) according to manufacturer's recommendations. Primer sequences used for *c-fos* and *β-actin* are provided in Table 5. The relative gene expression was calculated using 2^{-ΔΔCt} method (Livak and Schmittgen, 2001). The relative expression levels of *c-fos* mRNA were normalized by the geometric mean of *β-actin* mRNA expression.

Training and unilateral stimulation

Training experiments were performed on adult *H. lucorum* weighing 25–30 g. Animals were housed in large plastic boxes with increased humidity and fed with cabbage ad libitum. The animals were food deprived for 3 d before experiments. A total of 36 snails were involved in a food-aversion experiment. 12 of them were trained by the association of the novel food (carrot, conditioned stimulus, CS) with a bitter taste (10% quinine hydrochloride solution, unconditioned stimulus, US). Three CS-US paired stimulations were applied to the 12 snails with 10-min intertrial interval. Another 12 out of the 36 snails were used in the unpaired training, with three carrot and three quinine presentations applied in a random order. The remaining 12 snails were used as controls. For each experimental group, five snails were randomly selected and subjected to immunohistochemistry (IHC) analyses 2 h later from the start of the training (Zangenehpour and Chaudhuri, 2002; Barry and Commins, 2017). The remaining seven snails were subjected to behavioral test 24 h later to assess the long-term memory formation.

To reveal the immediate gene activation, 2 h after the stimulus presentation trained, unpaired trained, and control snails ($n = 5$ in each experimental group) were anesthetized; control snails were anesthetized immediately after the removal from their home boxes. CNS was quickly removed and frozen in liquid nitrogen vapors for IHC analyses. The remaining snails ($n = 7$ in each experimental group) were used to assess the reaction time latency 24 h after the training. The cabbage was used as a

Table 5. Primer sequences for RT-qPCR

Gene	Primer direction	Sequence	Product length
<i>c-fos</i>	Forward	5'-TTACACCCCA TTCATCCGC-3'	138
	Reverse	5'-AATACGTCCCC AGCGAAGTG-3'	
<i>β-Actin</i>	Forward	5'-AGTGCTTGCCCT TGTATGCCT-3'	185
	Reverse	5'-AACTTAAGCCC CTTCTGCC-3'	

Table 6. Antibodies/dyes used in the IHC

Antibodies	Made in	Dilution	Conjugate	Source	RRID
c-Fos antibody	Mouse	1:250		Santa Cruz sc-8047	AB_627253
				Santa Cruz sc-166940	AB_10609634
Anti-serotonin	Rabbit	1:1000		Sigma-Aldrich S5545	AB_477522
Anti-mouse	Horse	1:500	Biotinylated	Vector Labs BA-2000	AB_2313581
Anti-rabbit	Donkey	1:500	Alexa Fluor 488	Invitrogen A-21206	AB_2535792
Streptavidin		1:500	Alexa Fluor 568	Invitrogen S-11226	AB_2315774
DAPI		1:500		Invitrogen D1306	AB_2629482

differentiation stimulus (DS). Three CS and three DS were presented to each snail with 10-min intertrial intervals.

To prepare semi-intact preparations for unilateral stimulation experiments, five snails were anesthetized in ice-cold water for 30–40 min followed by the injection of 100- to 150-mg MgCl₂ in 1-ml physiologic saline for mollusks. Quinine solution (10%, 600 μl) was applied unilaterally to the chemoreceptive snail lip area and triggered ipsilateral contractions of the skin and mantle bolster. After 2 h from the stimulation, CNS was extracted and frozen in liquid nitrogen vapors for the following IHC.

IHC

For the multiple immunofluorescence reaction, 20-μm serial sections were prepared from snail CNS on freezing microtome Leica CM1950 by the freeze–thaw method. Sections were fixed in fresh ice-cold 4% PFA solution for 7 min and washed in 0.01 M PBS (1× PBS; pH 7.4) three times for 5 min. Then sections were incubated in permeabilization buffer [5% Triton X-100, 5% DMSO and 5% normal horse serum (NHS) in 1× PBS] for 1 h at room temperature and washed. The reaction with primary antibodies (Table 6) was performed in the blocking buffer (1% Triton X-100, 5% DMSO, 5% NHS, 0.01% NaN₃ in 1× PBS) overnight at 4°C followed by the washing. For the first experimental series, mouse anti-c-Fos antibodies (sc-8047, Santa-Cruz) were used. In the second series, to confirm the results, we used mouse anti-c-Fos antibodies (sc-166940, Santa-Cruz; Table 6). Then sections were incubated in biotinylated horse anti-mouse immunoglobulin IgG (BA-2000, VectorLabs; Table 6) in the blocking buffer for 2 h at room temperature and washed. After that sections were stained with streptavidin conjugated to Alexa Fluor 568 (S-11226, Invitrogen; Table 6) in the blocking buffer for 2 h at room temperature and washed. For double IHC, mouse anti-c-Fos (sc-166940, Santa-Cruz), rabbit anti-serotonin (S5545, Sigma), and corresponding biotinylated horse anti-mouse followed by streptavidin conjugated to Alexa Fluor 568 and donkey anti-rabbit Alexa Fluor 488 (A-21206, Invitrogen) were used (Table 6). After autofluorescence reduction in 1% Sudan black in 70% ethanol for 20 min followed by washing in 1× PBS, sections were mounted in FluoroMount aqueous mounting medium (Sigma) with fluorescent nuclear counterstain DAPI, coverslipped and sealed with nail polish. No signal was seen in negative control sections processed with primary antibody omission.

Images were obtained by Olympus FluoView 10i confocal laser scanning microscope with UPLSAPO 60×/1.20 W objective and Zeiss LSM800 AiryScan system with

LD Plan-Neofluar 40×/0.6 objective. The image analysis was performed in Imaris (Bitplane) and ImageJ (NIH) software. Fluorescence intensities were measured in the outlined nuclei of identified neurons on three sections from each neuron and were further averaged for each snail. Statistical difference was determined using Mann–Whitney test.

Data and code accessibility

The RNA-Seq data and the assembled snail neuronal transcriptome were deposited in the Gene Expression Omnibus (GEO) under the accession number GSE123558. The code is available as [Extended Data Code 1](#).

Results

Snail transcriptome assembly

We searched for IEGs in the terrestrial snail *H. lucorum* by comparing the transcriptome composition between the activated and control sections of the nervous system. In the first experiment (E1), we stimulated one half of the PaG network of six snails using serotonin (5-HT) in the presence of the protein synthesis inhibitor anisomycin, with the other half of the PaG network serving as a control (Fig. 1). In the second experiment (E2), we stimulated in the presence of anisomycin a medial part of the PaG containing mostly two giant premotor interneurons (Pa2/3) surgically dissected from the PaG network of the other six snails, and used the dissected Pa2/3 interneurons from the non-stimulated side of the network as a control. This resulted in a total of 12 activated and 12 control samples, dissected from 13 snails (Table 1).

We measured the poly A+ transcriptome of activated and control samples using RNA-Seq yielding a total of 460,233,404 reads with the read length of 100 nucleotides (nt; Table 2). Further, we generated 229,050,797 transcriptome read pairs with the read length of 101 nt and 12,701,459 transcriptome read pairs with the read length of 251 nt from *H. lucorum* neurons in additional experiments (Table 3). As the genome sequence of the snail *H. lucorum* or any other closely related invertebrate species was not yet deciphered, we conducted the *de novo* assembly of the neuronal transcriptome of the snail using all obtained reads.

The *de novo* transcriptome assembly yielded 693,041 contigs, far more than the anticipated transcript number, indicating the presence of alignment gaps. Nonetheless, the assessment criteria indicated reasonable assembly quality: N50 representing the minimal size of the contigs covering half of the assembled transcriptome sequence equaled 865 bp, the median GC content equaled 41%,

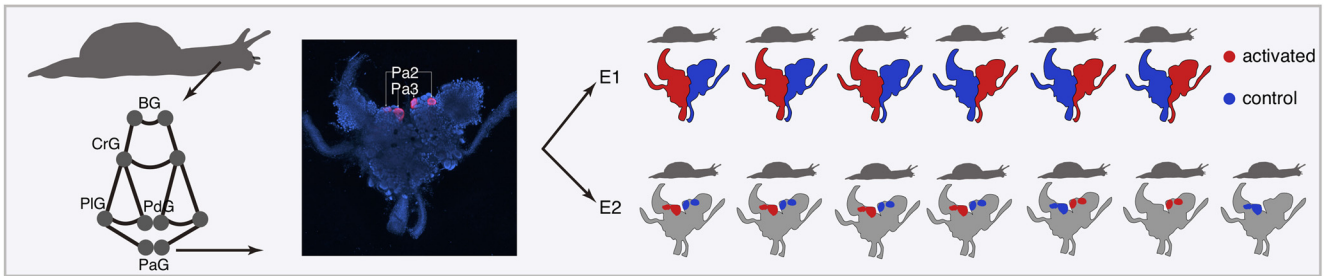


Figure 1. Experimental design. Left, Silhouette of the snail *H. lucorum* and the schematic representation of its nervous system showing the five pairs of ganglia: buccal ganglia (BG), cerebral ganglia (CrG), pedal ganglia (PdG), pleural ganglia (PlG), and PaG. The visceral ganglion was not used in experiments and was not shown. Middle, Fluorescence microscopy image of the PaG (blue) with two giant interneurons used in E2 (Pa2/3) shown in pink. Right, Schematic representation of two stimulation experiments conducted using the left or right PaG (E1) or a surgically dissected medial part of the PaG containing the Pa2/3 interneurons (E2). Extended data [Figures 1-1, 1-2](#) showed the quality of the assembled snail neuronal transcriptome and the snail sample distribution based on the RNA-Seq measurements, respectively.

and the median open reading frame (ORF) percentage equaled 71% (Extended Data [Fig. 1-1](#)).

To test the reliability of our assembly, we conducted conditional reciprocal best BLAST (CRBB) between the snail assembly and the annotated proteome of another gastropod, the California sea slug *A. californica* ([Knudsen et al., 2006](#)). Almost half (49%) of the slug proteins could be reciprocally matched with the assembled snail transcripts, despite >450 million years of evolution separating these two species.

Detection of DE contigs

To quantify gene expression at the contig level, we mapped RNA-Seq reads from each sample to the transcriptome assembly. On average, 60% of reads were mapped to the assembled transcripts ([Table 2](#)), resulting in 404,678 (58%) contigs classified as expressed in E1 and 332,423 (48%) contigs in E2. The MDS analysis based on the expression of these contigs revealed the absence of outliers and clustering of samples according to individuals (Extended Data [Fig. 1-2](#)).

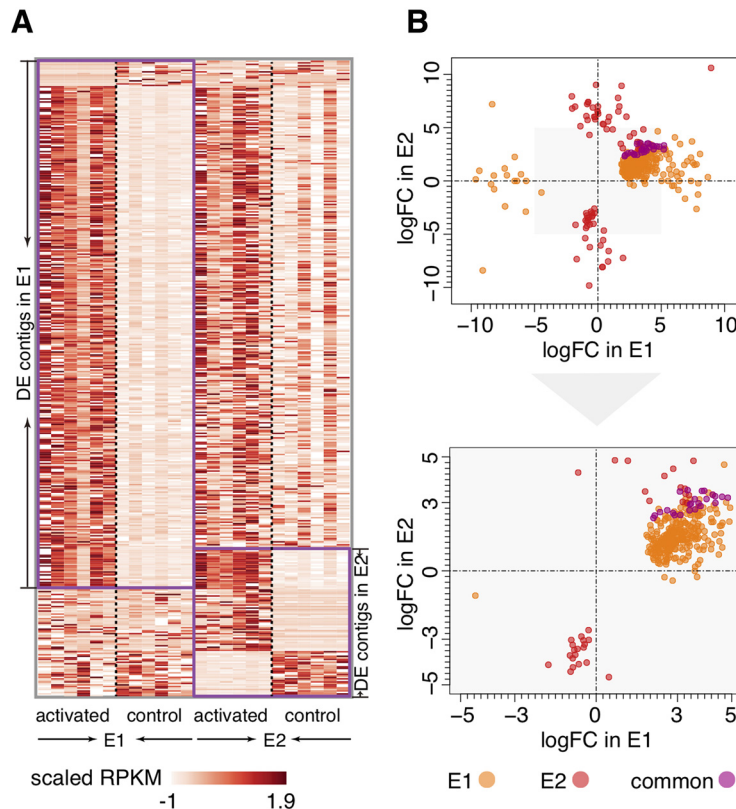


Figure 2 Differential expression after the neuronal stimulation. **A**, Heat map showing expression levels as standard-normalized RPKM values of 422 DE contigs classified as DE in at least one of the two experiments. Purple boundaries indicate contigs showing significant expression differences in each experiment. **B**, up, Scatter plot showing the amplitude and the direction of expression differences of the 422 DE contigs in E1 and E2 as log₂-transformed fold changes (logFC). The colors indicate significant expression differences in one or both experiments. Down, Zoomed in area of the upper plot shaded in gray.

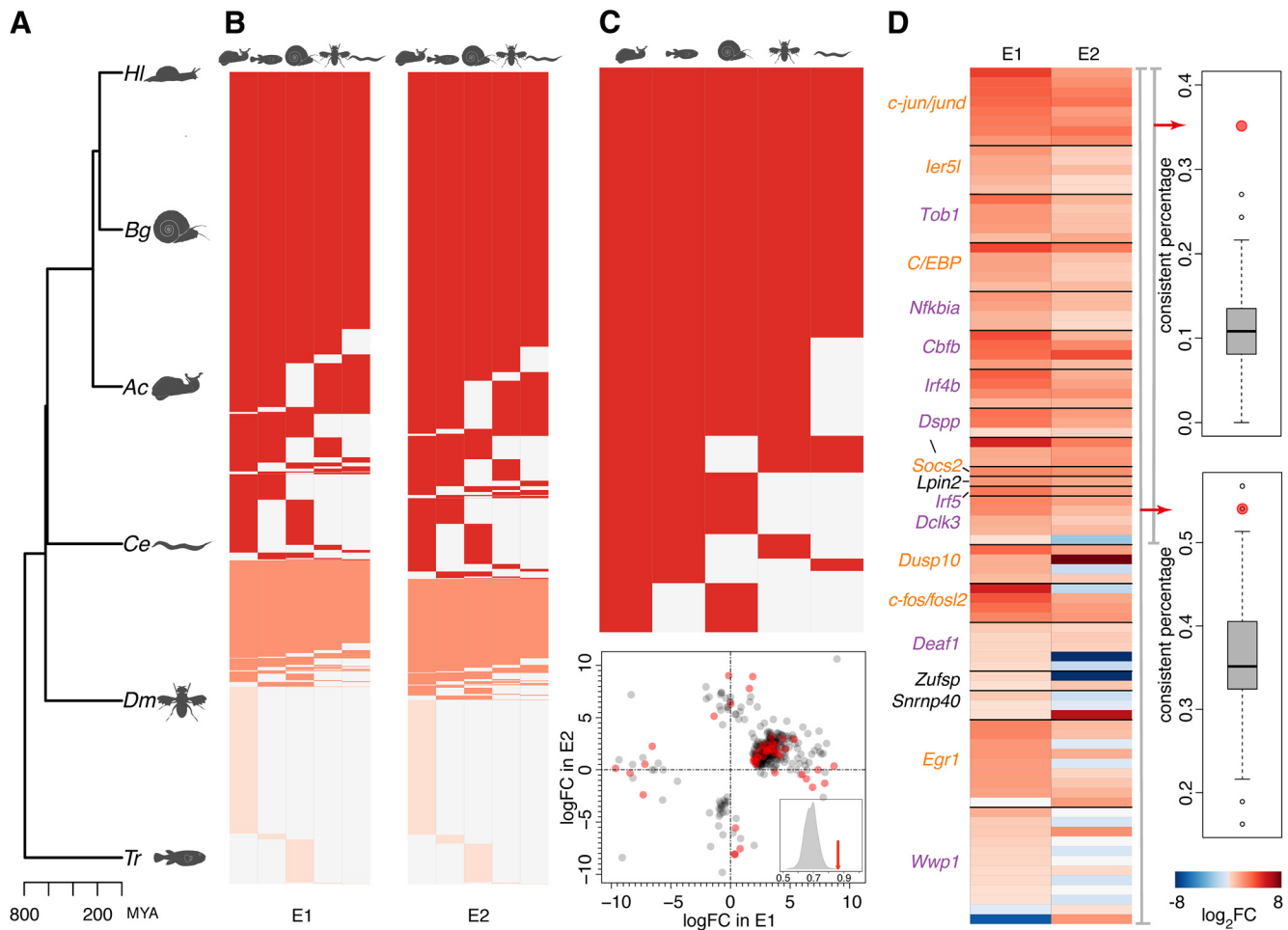


Figure 3. Annotation of expressed contigs. **A**, Phylogenetic tree showing the relationship between the snail *H. lucorum* (HI) and the five species used in the contig annotation: *B. glabrata* (Bg), *A. californica* (Ac), *C. elegans* (Ce), *D. melanogaster* (Dm), and *T. rubripes* (Tr). MYA, million years ago. **B**, Snail contigs annotated and expressed in E1 (left) and E2 (right). The colors indicate contigs mapped to one species (light red); contigs mapped to multiple species, but inconsistently annotated among them (coral); and contigs mapped to multiple species and consistently annotated among them (red). **C**, up, 46 DE contigs consistently mapped to 37 proteins in multiple species. Down, Scatter plot showing the amplitude and the direction of expression differences of 422 DE contigs in E1 and E2 (the same as in Fig. 2B) with the 46 DE contigs shown in red. The inset shows the percentage of upregulated non-DE contigs mapped to the 37 proteins (red arrow) and the percentage expected by chance (gray distribution). **D**, left, Heat map showing the log₂-transformed fold changes between activated and control samples in E1 and E2. The fold changes are shown for all contigs mapped to the 20 proteins, independent of the significance of the difference. Colors of protein names indicate known IEGs (orange) and stimulus response and immune response genes (purple). Right, Percentage of proteins showing a consistent difference direction based on no <80% of all mapped contigs in both experiments (upper) or within one experiment (lower). The red dots show the observed percentage and gray boxplots show the distributions of percentages expected by chance. Extended data Figure 3-1 showed the expression regulation of all the 37 consistently annotated proteins.

The statistical analysis identified 350 contigs DE between control and activated samples in E1 and 98 contigs in E2 [exact test, Benjamini and Hochberg (BH) FDR-corrected $p < 0.05$; fold change > 2]. Of them, 26 contigs were DE in both experiments. All 26 were upregulated after the neuronal activity stimulation, an observation not expected by chance (hypergeometric test, $p < 0.0001$). Overall, 95% of DE contigs in E1 and 70% of DE contigs in E2 were upregulated after the neuronal activity stimulation, consistent with the known IEGs' response mechanism (Fig. 2A). Furthermore, the direction of expression difference was in good agreement between the experiments: 83% of the 422 DE contigs representing the union

of the two experiments were upregulated and 7% downregulated in both experiments after the neuronal activation (χ^2 test, $p < 0.0001$; Fig. 2A,B).

Annotation of expressed contigs

We then annotated all expressed contigs by translating their nucleotide sequences into the amino acid sequences in all six possible frames and aligning them to the protein sequences from five species: bloodfluck planorb (*B. glabrata*, Bg), California sea slug (*A. californica*, Ac), roundworm (*C. elegans*, Ce), fruit fly (*D. melanogaster*, Dm), and tiger puffer (*T. rubripes*, Tr; Fig. 3A).

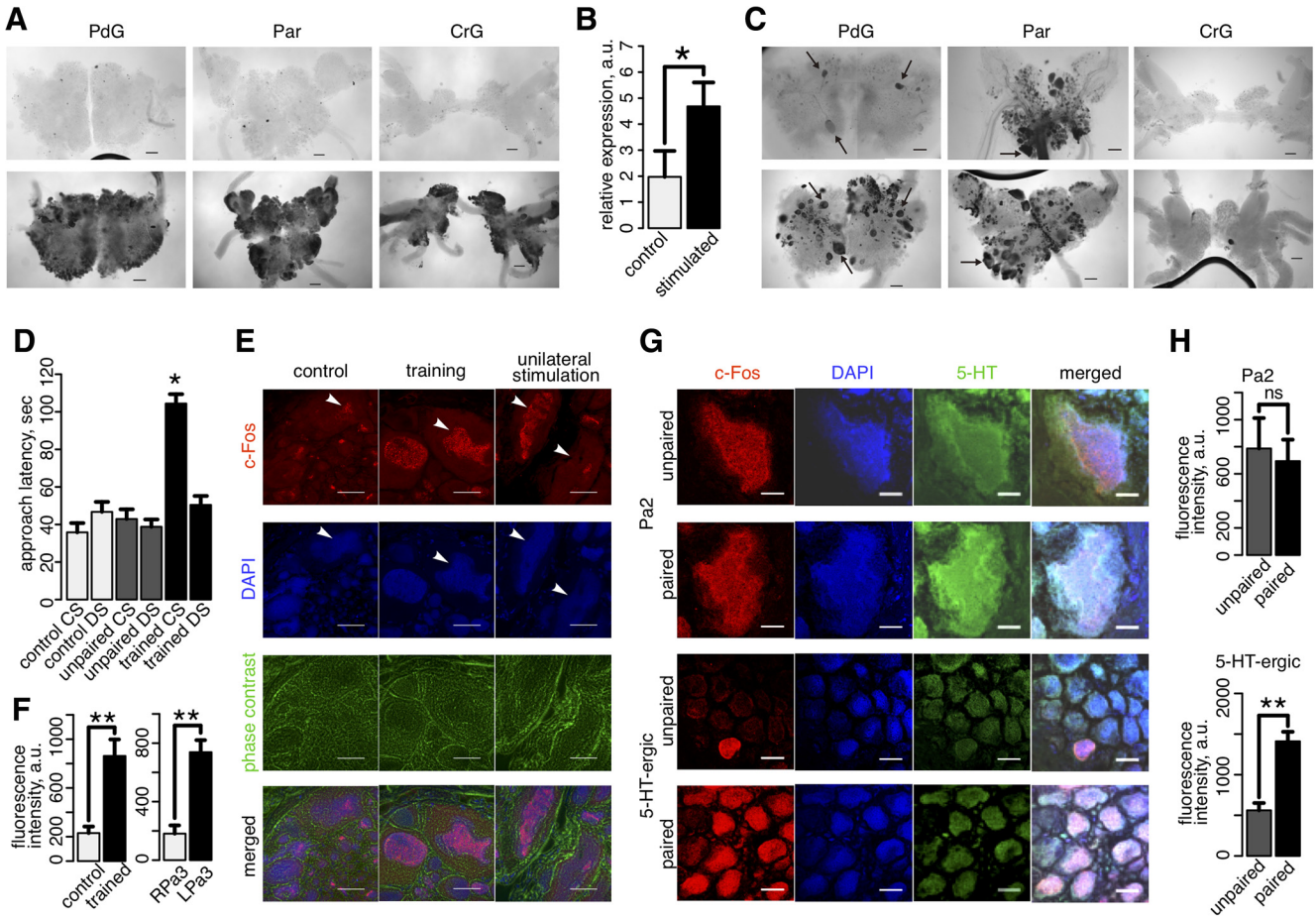


Figure 4. Expression of the snail *c-Fos* homolog. **A**, The *in situ* hybridization pattern of the *c-fos/fos2* probe in three parts of the snail nervous system, PdG, Par, and CrG after the stimulation by the bath application of 5-HT for 2 h (lower panel) and in the non-stimulated controls (upper panel). **B**, Expression of *c-fos* (mean \pm SEM) in non-stimulated control parts of PaG (control, $n = 5$) and 5-HT activated parts of PaG (stimulated, $n = 4$) measured by RT-qPCR. The asterisk (*) indicates the significance of one-sided Mann-Whitney test, $p < 0.05$. **C**, Localization of neurons projecting to the anal nerve revealed by the backfill with neurobiotin (upper panel), and the *in situ* hybridization pattern of the *c-fos/fos2* probe after the stimulation of the anal nerve (lower panel). Arrows mark activated neurons projecting to the anal nerve revealed by the backfill. Scale bar = 200 μ m. **D**, The approach latency of snails (in seconds) from control, CS-US unpaired, and CS-US paired training groups ($n = 7$) for different stimuli: CS (carrot), DS (cabbage). The asterisk indicates the significance of the difference in CS stimulus effect between the CS-US paired group and all the other conditions (Mann-Whitney test, Bonferroni corrected $p < 0.05$). **E**, IHC images showing the fluorescent staining by the mouse monoclonal antibodies against conserved amino acid regions within human *c-Fos* protein (red); the general nuclear marker DAPI (blue); the phase contrast microscopy image (green); and a merged image. The images show sections of the PaG network taken from control snails (control), from snails subjected to the behavioral training using taste aversion (CS-US paired, training), and from the unilaterally stimulated semi-intact snail CNS preparations (unilateral stimulation). The arrowheads indicate the location of the giant Pa2 interneuron in control and training experiments. In the unilateral stimulation experiment, the arrowheads show Pa3 interneurons from the stimulated (left) and non-stimulated (right) sides of the PaG network. Scale bar = 50 μ m. **F**, Fluorescence intensity (mean \pm SEM) measured in the outlined nuclei of Pa2 neurons from control snails ($n = 5$) and trained snails (CS-US paired training, $n = 5$; left), as well as from right non-stimulated (RPa3) and left stimulated (LPa3) Pa3 neurons in the unilateral stimulation experiment ($n = 5$; right). The asterisk (**) indicates the significance of one-sided Mann-Whitney test, $p < 0.005$. **G**, IHC images showing the fluorescent staining by the mouse monoclonal antibodies against conserved amino acid regions within human *c-Fos* protein (red); the general nuclear marker DAPI (blue); anti-serotonin antibody (5-HT, green); and a merged image. The images show sections containing Pa2 neurons of PaG (upper two panels) and serotonergic (5-HT-ergic) neurons of PdG (lower two panels) taken from snails subjected to CS-US paired behavioral training (paired) and CS-US unpaired presentations (unpaired). Scale bar = 20 μ m. **H**, Fluorescence intensity (mean \pm SEM) measured in the outlined nuclei of Pa2 neurons from snails trained using paired stimuli (CS-US paired training, $n = 5$) and snails trained using unpaired stimuli ($n = 5$; upper panel), as well as in serotonergic neurons from PdG of snails trained using paired and unpaired stimuli ($n = 5$, lower panel). The symbols indicate the significance of one-sided Mann-Whitney test, ** $p < 0.005$; ns, non-significant.

Out of all expressed contigs in E1 and E2, 49,695 (12.3%) and 44,958 (13.5%) were, respectively, mapped to a total of 12,711 and 12,104 protein sequences in at least one species (Fig. 3B). Of them, 76% and 77% were

mapped in multiple species, yielding the consistent annotation in 79% and 81% of the cases (Fig. 3B). On average, each consistently annotated protein was represented by three contigs in each of the two experiments.

Table 7. Expression of c-Fos measured using RT-qPCR and IHC

RT-qPCR for 5-HT activation experiment (Fig. 4B)		Behavioral test for control, unpaired trained and paired trained snails using CS or DS (approach time latency in seconds; Fig. 4D)					
Control	Stimulated	Control CS	Control DS	Unpaired CS	Unpaired DS	Trained CS	Trained DS
5.53	2.99	23.3	60.7	36.7	42	112.7	45.3
2.73	6.5	58.7	57.7	57.7	33.7	112.7	54
1.11	NA	45.3	41.3	41.3	28.7	83	31.3
0.24	3.21	26	64.3	49.3	51.7	109	38.7
0.25	6.03	36.7	34	16.7	39	90.7	60.7
		37.3	40.3	44.7	25.3	120	56.3
		23.3	28.7	54	51	102.7	66.7
Fluorescence intensity of c-Fos in Pa2 2 h after training (Fig. 4F)		Fluorescence intensity of c-Fos in Pa3 after unilateral stimulation (Fig. 4F)		Fluorescence intensity of c-Fos in Pa2 2 h after training (Fig. 4H)		Fluorescence intensity of c-Fos in serotonergic neurons 2 h after training (Fig. 4H)	
Control	Trained	Stimulated	Control	Unpaired	Paired	Unpaired	Paired
283.3	819.7	735.7	284.7	603.3	1274.7	355.7	1789
332.7	1223	976	38.7	1449.3	536.3	766.3	1344
71.3	470.3	637	310	1181.7	736	459	1511
131.7	698.3	498.3	214.3	347.7	582.7	788.3	1329.7
328.7	1100.3	844.7	58.7	352.7	339.3	436.3	1082

Among DE contigs, 46 were consistently mapped in multiple species to 37 proteins (Fig. 3C). Further, 138 contigs not classified as DE were mapped to these 37 proteins. Despite failing to pass the stringent significance cutoff, 86% of these 138 contigs were upregulated in activated samples in at least one experiment, an observation consistent with the expression behavior of DE contigs, and not expected by chance (permutations, $p < 0.001$; Fig. 3C). Accordingly, for 20 of the 37 proteins, >80% of its all mapped contigs, including both DE and non-DE contigs, showed consistent upregulation in activated samples (permutations, $p < 0.004$; Fig. 3D; Extended Data Fig. 3-1).

Notably, these 20 proteins contained snail homologs of seven previously characterized mammalian IEGs: *c-jun/jund*, *C/EBP*, *c-fos/fosl2*, *Egr1*, *ler5l*, *Socs2*, and *Dusp10* (Fig. 3D). Among the remaining 13 proteins, nine fell within well-defined Gene Ontology (GO) terms: “response to stimulus” and “immune system process” (Fig. 3D).

In vivo assessment of identified IEGs

To characterize the spatial expression of putative snail IEGs detected in our study, we conducted *in situ* hybridization experiments in snails’ pedal ganglia (PdG), pleuro-viscero-PaG complex (Par), and cerebral ganglia (CrG) using the customized probe targeting the snail transcript homologues to mammalian *fos* family IEGs *c-fos* and *fosl2* and fruit fly IEG *kayak* (*c-fos/fosl2* probe; Table 4).

The global neuronal activation of the circumphageal ganglia complex (cerebral, pleural, parietal, pedal and visceral ganglia) by 5 μ M 5-HT bath application produced widespread hybridization signals of the *c-fos/fosl2* probe within neurons compared to non-stimulated controls in all ganglia (Fig. 4A), which was further confirmed by the increased *c-fos* expression in activated parts of PaG measured using quantitative PCR (RT-qPCR; Fig. 4B; Table 7). The selective stimulation via the anal nerve elicited clear

and specific hybridization patterns with the *c-fos/fosl2* probe in the ganglia (Fig. 4C). Comparisons of neurons projecting to the anal nerve identified by the neurobiotin backfill with *c-fos/fosl2* hybridization patterns further showed the accurate identification of these neurons by the probe in the PdG and Par, and revealed additional neurons possibly representing the secondary activation response. These results show the potential of identified IEGs to reveal gene activation patterns in the snail nervous system.

To further assess the expression of c-Fos protein homolog in snails, we conducted IHC experiments in the PaG containing the two giant premotor interneurons (Pa2/3) and PdG using the mouse monoclonal antibody against amino acids of c-Fos of human origin (Table 6). The IHC showed increased c-Fos expression in the Pa2/3 nuclei after the *in vivo* behavioral training using taste aversion paradigm (Fig. 4D–F), and after the unilateral stimulation of the lip by quinine in the semi-intact snail CNS preparation (Fig. 4E,F). The activation of c-Fos in Pa2 neurons after the behavioral training was not directly linked to associative learning (Fig. 4G,H), consistent with studies conducted in mice and other vertebrates (Herdgen and Leah, 1998; Kovacs, 1998; Kaczmarek et al., 1999; Burnham et al., 2010; Lopez et al., 2018; Fredericksen et al., 2019; Möser et al., 2019). Intriguingly, a group of serotonergic neurons located in PdG showed greater c-Fos upregulation after paired CS and US presentation compared to unpaired ones (Fig. 4G,H). This suggests possible roles of these neurons in the formation of learning and memory in snails.

Discussion

Gastropods, including the sea slug *A. californica*, the terrestrial slug *L. valentianus*, the pond snail *L. stagnalis*, and the terrestrial snail *H. lucorum* are widely used in studies of nervous system organization and function due

to their simpler organization and the presence of large neurons and axons, facilitating electrophysiological readings (Clarac and Pearlstein, 2007). These studies produced fundamental insights into the basic mechanisms of learning and memory formation. The potential of these model systems was not fully used, however, due to insufficient knowledge of the molecular mechanisms accompanying the neuronal activation.

In the absence of the genome sequence or its close homologues, the *de novo* transcriptome assembly represents a useful approach to identify transcripts and quantify their expression. In our study, we conducted the assembly of the snail neuronal transcriptome using >943 million reads with lengths from 100 to 251 nt cumulatively covering >98 billion positions. The resulting transcriptome contained gaps, with an average of three contigs mapping to the same protein. It is recognized that *de novo* transcriptome assembly intrinsically generates fewer complete transcripts than reference genome-based methods (Martin and Wang, 2011; Bushmanova et al., 2018). Moreover, the extent of fully-constructed transcripts drops drastically when the number of isoforms per gene increases (Chang et al., 2014). This represents a common issue of *de novo* transcriptome reconstruction for non-model organisms, which is exacerbated by factors such as homologous or repetitive genomic regions (Treangen and Salzberg, 2011), varying read coverage depths along a transcript (Chow et al., 2014), low-expression transcripts (Zhao et al., 2011) and sequencing artifacts. Taking into account the number of reads generated in our study ($n = 943$ millions), further improvement in the assembly quality can be achieved by the use of alternative protocols involving longer read lengths. Nonetheless, the assembly quality (N50 = 865 bp) was substantially better than that reported for another land snail *Cornu aspersum* (N50 = 365 bp; Parmakelis et al., 2017), and comparable to the transcriptome assembly quality of the sea snail *Dicathais orbita* (N50 = 608 bp; Baten et al., 2016), the land snail *Aegista chejuensis* (N50 = 788 bp; Kang et al., 2016), and the pond snail *L. stagnalis* (N50 ranged from 564 to 1614 bp depending on the assembly algorithm; Sadamoto et al., 2012).

The differential expression analysis revealed the substantial gene expression activation after the serotonin stimulation both within a specific ganglion of the snail nervous system (E1) and in a dissected medial part of a ganglion containing identified interneurons (E2). Notably, genes activated in the snail contained a number of homologues to well-characterized mammalian IEGs including *c-jun/jund*, *C/EBP*, *c-fos/fosl2*, *Egr1*, *Ier5l*, *Socs2*, and *Dusp10*. Some of these genes were characterized in other invertebrates with a similar nervous system organization. Specifically, *c-jun* was shown to modulate the synaptogenesis between sensory neurons and motor neurons in *A. californica* (Sung et al., 2006). Similarly, *Egr1*, a member of the *Egr* family, was upregulated after long-time sensitization training in *A. californica* (Cyriac et al., 2013). Another well-characterized mammalian IEG also found in the snail in our study, *C/EBP*, was detected in all 11 ganglia (Hatakeyama et al., 2004) and undergone expres-

sion changes during early memory consolidation in the pond snail *L. stagnalis* (Hatakeyama et al., 2006). Moreover, *C/EBP* was shown to be involved in the consolidation phase of long-term facilitation in *A. californica* (Alberini et al., 1994).

Besides known mammalian IEGs' homologues, we found a number of not yet annotated putative snail IEGs. Genes activated in our experiments were particularly overrepresented in immune or stimulus response terms, indicating the potential activation of these systems by the serotonin stimulation procedures. Notably, the number of contigs activated after the stimulation of the PaG network exceeded the number detected in a dissected part of the ganglion containing two giant premotor interneurons by more than 3-fold. This observation was not caused by the difference in numbers of contigs detected in the two experiments, and matched reports indicating the variability of IEG repertoire among different neuron types (Guo et al., 2014; Chen et al., 2016).

IEGs are often used as neuronal activity markers to visualize the activity patterns within the nervous system (Wilson et al., 2002; Bepari et al., 2012; Fujita et al., 2013). Our *in situ* hybridization experiments conducted using the probe specific to the snail homolog of mammalian *fos* family IEGs *c-fos* and *fosl2*, and fly IEG *kayak* indicated the potential of identified snail transcripts to visualize activated neurons. Designs of antibodies specific to snail IEGs would be the next step to facilitate studies of the gene activation networks in a low-complexity model system.

References

- Alberini CM, Ghirardi M, Metz R, Kandel ER (1994) C/EBP is an immediate-early gene required for the consolidation of long-term facilitation in *Aplysia*. *Cell* 76:1099–1114.
- Balaban PM (1980) Sensitization and habituation in command neurons of the defensive reflex in *Helix lucorum*. *Neurosci Behav Physiol* 10:340–345.
- Balaban PM (2002) Cellular mechanisms of behavioral plasticity in terrestrial snail. *Neurosci Biobehav Rev* 26:597–630.
- Balaban P, Chase R (1989) Self-stimulation in snails. *Neurosci Res Commun* 4:139–146.
- Balaban PM, Poteryaev DA, Zakharov IS, Uvarov P, Malyshev A, Belyavsky AV (2001) Up- and down-regulation of *Helix* command-specific 2 (HCS2) gene expression in the nervous system of terrestrial snail *Helix lucorum*. *Neuroscience* 103:551–559.
- Balaban PM, Roshchin M, Timoshenko A, Zuzina AB, Lemak M, Ierusalimsky VN, Aseyev NA, Malyshev AY (2015) Homolog of protein kinase M ζ maintains context aversive memory and underlying long-term facilitation in terrestrial snail *Helix*. *Front Cell Neurosci* 9:222.
- Barry DN, Commins S (2017) Temporal dynamics of immediate early gene expression during cellular consolidation of spatial memory. *Behav Brain Res* 327:44–53.
- Baten A, Ngangbam AK, Waters DL, Benkendorff K (2016) Transcriptome of the Australian mollusc *Dicathais orbita* provides insights into the biosynthesis of indoles and choline esters. *Mar Drugs* 14:E135.
- Bepari AK, Watanabe K, Yamaguchi M, Tamamaki N, Takebayashi H (2012) Visualization of odor-induced neuronal activity by immediate early gene expression. *BMC Neurosci* 13:140.
- Bojovic O, Panja D, Bittins M, Bramham CR, Tjølsen A (2015) Time course of immediate early gene protein expression in the spinal

- cord following conditioning stimulation of the sciatic nerve in rats. *PLoS One* 10:e0123604.
- Bonnick K, Bayas K, Belchenko D, Cyriac A, Dove M, Lass J, McBride B, Calin-Jageman IE, Calin-Jageman RJ (2012) Transcriptional changes following long-term sensitization training and in vivo serotonin exposure in *Aplysia californica*. *PLoS One* 7:e47378.
- Burnham KE, Bannerman DM, Dawson LA, Southam E, Sharp T, Baxter MG (2010) Fos expression in the brains of rats performing an attentional set-shifting task. *Neuroscience* 171:485–495.
- Bushmanova E, Antipov D, Lapidus A, Przhibelskiy AD (2018) rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *bioRxiv*. 420208. Advance online publication. Retrieved September 18, 2018. doi:10.1101/420208
- Chang Z, Wang Z, Li G (2014) The impacts of read length and transcriptome complexity for de novo assembly: a simulation study. *PLoS One* 9:e94825.
- Chen X, Rahman R, Guo F, Rosbash M (2016) Genome-wide identification of neuronal activity-regulated genes in *Drosophila*. *Elife* 5:e19942.
- Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162:156–159.
- Chow KS, Ghazali AK, Hoh CC, Mohd-Zainuddin Z (2014) RNA sequencing read depth requirement for optimal transcriptome coverage in *Hevea brasiliensis*. *BMC Res Notes* 7:69.
- Clarac F, Pearlstein E (2007) Invertebrate preparations and their contribution to neurobiology in the second half of the 20th century. *Brain Res Rev* 54:113–161.
- Cyriac A, Holmes G, Lass J, Belchenko D, Calin-Jageman RJ, Calin-Jageman IE (2013) An *Aplysia* Egr homolog is rapidly and persistently regulated by long-term sensitization training. *Neurobiol Learn Mem* 102:43–51.
- Dahmen N, Fehr C, Reuss S, Hiemke C (1997) Stimulation of immediate early gene expression by desipramine in rat brain. *Biol Psychiatry* 42:317–323.
- Fowler T, Sen R, Roy AL (2011) Regulation of primary response genes. *Mol Cell* 44:348–360.
- Fredericksen KE, McQueen KA, Samuelsen CL (2019) Experience-dependent c-Fos expression in the mediodorsal thalamus varies with chemosensory modality. *Chem Senses* 44:41–49.
- Freytag V, Probst S, Hadziselimovic N, Boglari C, Hauser Y, Peter F, Gabor Fenyves B, Milnik A, Demougis P, Vukojevic V, de Quervain DJ, Papassotiropoulos A, Stetak A (2017) Genome-wide temporal expression profiling in *Caenorhabditis elegans* identifies a core gene set related to long-term memory. *J Neurosci* 37:6661–6672.
- Fujita N, Nagata Y, Nishiuchi T, Sato M, Iwami M, Kiya T (2013) Visualization of neural activity in insect brains using a conserved immediate early gene, Hr38. *Curr Biol* 23:2063–2070.
- Fukunaga S, Matsuo R, Hoshino S, Kirino Y (2006) Novel kruppel-like factor is induced by neuronal activity and by sensory input in the central nervous system of the terrestrial slug *Limax valentianus*. *J Neurobiol* 66:169–181.
- Gerstner JR, Koberstein JN, Watson AJ, Zaperio N, Risso D, Speed TP, Frank MG, Peixoto L (2016) Removal of unwanted variation reveals novel patterns of gene expression linked to sleep homeostasis in murine cortex. *BMC Genomics* 17:727.
- Guan Z, Saraswati S, Adolfsen B, Littleton JT (2005) Genome-wide transcriptional changes associated with enhanced activity in the *Drosophila* nervous system. *Neuron* 48:91–107.
- Guo F, Cerullo I, Chen X, Rosbash M (2014) PDF neuron firing phase-shifts key circadian activity neurons in *Drosophila*. *Elife* 3:e02780.
- Hatakeyama D, Fujito Y, Sakakibara M, Ito E (2004) Expression and distribution of transcription factor CCAAT/enhancer-binding protein in the central nervous system of *Lymnaea stagnalis*. *Cell Tissue Res* 318:631–641.
- Hatakeyama D, Sadamoto H, Watanabe T, Wagatsuma A, Kobayashi S, Fujito Y, Yamashita M, Sakakibara M, Kemenes G, Ito E (2006) Requirement of new protein synthesis of a transcription factor for memory consolidation: paradoxical changes in mRNA and protein levels of C/EBP. *J Mol Biol* 356:569–577.
- Herdegen T, Leah JD (1998) Inducible and constitutive transcription factors in the mammalian nervous system: control of gene expression by Jun, Fos and Krox, and CREB/ATF proteins. *Brain Res Brain Res Rev* 28:370–490.
- Kaczmarek L, Zangenehpour S, Chaudhuri A (1999) Sensory regulation of immediate-early genes c-fos and zif268 in monkey visual cortex at birth and throughout the critical period. *Cereb Cortex* 9:179–187.
- Kang SW, Patnaik BB, Hwang HJ, Park SY, Wang TH, Park EB, Chung JM, Song DK, Patnaik HH, Lee JB, Kim C, Kim S, Park HS, Lee JS, Han YS, Lee YS (2016) De novo transcriptome generation and annotation for two Korean endemic land snails, *Aegista chejuensis* and *Aegista quepartensis*, using illumina paired-end sequencing technology. *Int J Mol Sci* 17:379.
- Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465:182–187.
- Knudsen B, Kohn AB, Nahir B, McFadden CS, Moroz LL (2006) Complete DNA sequence of the mitochondrial genome of the sea-slug, *Aplysia californica*: conservation of the gene order in Euthyneura. *Mol Phylogenet Evol* 38:459–469.
- Kovacs KJ (1998) c-Fos as a transcription factor: a stressful (re)view from a functional map. *Neurochem Int* 33:287–297.
- Lacar B, Linker SB, Jaeger BN, Krishnaswami S, Barron J, Kelder M, Parylak S, Paquola A, Venepally P, Novotny M, O'Connor C, Fitzpatrick C, Erwin J, Hsu JY, Husband D, McConnell MJ, Lasken R, Gage FH (2016) Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat Commun* 7:11022.
- Lakhina V, Arey RN, Kaletsky R, Kauffman A, Stein G, Keyes W, Xu D, Murphy CT (2015) Genome-wide functional analysis of CREB/long-term memory-dependent transcription reveals distinct basal and memory gene expression programs. *Neuron* 85:330–345.
- Lanahan A, Worley P (1998) Immediate-early genes and synaptic function. *Neurobiol Learn Mem* 70:37–43.
- Lee YS, Choi SL, Kim TH, Lee JA, Kim HK, Kim H, Jang DJ, Lee JJ, Lee S, Sin GS, Kim CB, Suzuki Y, Sugano S, Kubo T, Moroz LL, Kandel ER, Bhak J, Kaang BK (2008) Transcriptome analysis and identification of regulators for long-term plasticity in *Aplysia kuro-dai*. *Proc Natl Acad Sci USA* 105:18602–18607.
- Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods* 25:402–408.
- Lopez J, Gamache K, Milo C, Nader K (2018) Differential role of the anterior and intralaminar/lateral thalamic nuclei in systems consolidation and reconsolidation. *Brain Struct Funct* 223:63–76.
- Malyshev AY, Balaban PM (2002) Identification of mechanoafferent neurons in terrestrial snail: response properties and synaptic connections. *J Neurophysiol* 87:2364–2371.
- Martin JA, Wang Z (2011) Next-generation transcriptome assembly. *Nat Rev Genet* 12:671–682.
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17:10–12.
- Möser CV, Möller M, Fleck SC, Thomas D, Geisslinger G, Niederberger E (2019) Inhibition of the protein kinase IKKepsilon attenuates neuropathic pain in mice. *Neuropharmacology* 146:198–211.
- Parmakelis A, Kotsakiozi P, Kontos CK, Adamopoulos PG, Scorilas A (2017) The transcriptome of a “sleeping” invader: de novo assembly and annotation of the transcriptome of aestivating *Cornu aspersum*. *BMC Genomics* 18:491.

- Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.
- Sadamoto H, Sato H, Kobayashi S, Murakami J, Aonuma H, Ando H, Fujito Y, Hamano K, Awaji M, Lukowiak K, Urano A, Ito E (2004) CREB in the pond snail *Lymnaea stagnalis*: cloning, gene expression, and function in identifiable neurons of the central nervous system. *J Neurobiol* 58:455–466.
- Sadamoto H, Kitahashi T, Fujito Y, Ito E (2010) Learning-dependent gene expression of CREB1 isoforms in the molluscan brain. *Front Behav Neurosci* 4:25.
- Sadamoto H, Takahashi H, Okada T, Kenmoku H, Toyota M, Asakawa Y (2012) De novo sequencing and transcriptome analysis of the central nervous system of mollusc *Lymnaea stagnalis* by deep RNA sequencing. *PLoS One* 7:e42546.
- Saha RN, Wissink EM, Bailey ER, Zhao M, Fargo DC, Hwang JY, Daigle KR, Fenn JD, Adelman K, Dudek SM (2011) Rapid activity-induced transcription of *Arc* and other IEGs relies on poised RNA polymerase II. *Nat Neurosci* 14:848–856.
- Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S (2016) TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome Res* 26:1134–1144.
- Spiegel I, Mardinly AR, Gabel HW, Bazinet JE, Couch CH, Tzeng CP, Harmin DA, Greenberg ME (2014) *Npas4* regulates excitatory-inhibitory balance within neural circuits through cell-type-specific gene programs. *Cell* 157:1216–1229.
- Sung YJ, Wu F, Schacher S, Ambron RT (2006) Synaptogenesis regulates axotomy-induced activation of c-Jun-activator protein-1 transcription. *J Neurosci* 26:6439–6449.
- Thompson CL, Wisor JP, Lee CK, Pathak SD, Gerashchenko D, Smith KA, Fischer SR, Kuan CL, Sunkin SM, Ng LL, Lau C, Hawrylycz M, Jones AR, Kilduff TS, Lein ES (2010) Molecular and anatomical signatures of sleep deprivation in the mouse brain. *Front Neurosci* 4:165.
- Treangen TJ, Salzberg SL (2011) Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* 13:36–46.
- West AE, Greenberg ME (2011) Neuronal activity-regulated gene transcription in synapse development and cognitive function. *Cold Spring Harb Perspect Biol* 3:a005744.
- Wilson Y, Nag N, Davern P, Oldfield BJ, McKinley MJ, Greferath U, Murphy M (2002) Visualization of functionally activated circuitry in the brain. *Proc Natl Acad Sci USA* 99:3252–3257.
- Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, Huang W, He G, Gu S, Li S, Zhou X, Lam TW, Li Y, Xu X, Wong GK, Wang J (2014) SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* 30:1660–1666.
- Zangenehpour S, Chaudhuri A (2002) Differential induction and decay curves of *c-fos* and *zif268* revealed through dual activity maps. *Brain Res Mol Brain Res* 109:221–225.
- Zhao QY, Wang Y, Kong YM, Luo D, Li X, Hao P (2011) Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study. *BMC Bioinformatics* 12 [Suppl 14]:S2.