Male differentiation in the marine copepod *Oithona nana* reveals the development of a new nervous ganglion linked to Lin12-Notch-Repeat protein-associated proteolysis

Kevin Sugier[1], kevin.sugier@gmail.com

Romuald Laso-Jadart[1], ljromu@gmail.com

Benoit Vacherie[2], bvacheri@genoscope.cns.fr

Jos Käfer[3], jos.kafer@univ-lyon1.fr

Laurie Bertrand[1], lbertran@genoscope.cns.fr

Karine Labadie[2], klabadie@genoscope.cns.fr

Nathalie Martins[1], nmartins@genoscope.cns.fr

Céline Orvain[1], corvain@genoscope.cns.fr

Emmanuelle Petit[2], mpetit@genoscope.cns.fr

Julie Poulain[1], poulain@genoscope.cns.fr

Patrick Wincker[1], pwincker@genoscope.cns.fr

Jean-Louis Jamet[4], jean-louis.jamet@univ-tln.fr

Adriana Alberti[2] , aalberti@genoscope.cns.fr

Mohammed-Amin Madoui[1], amadoui@genoscope.cns.fr, corresponding author

[1]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, Evry, France

[2]Genoscope, Institut de biologie François-Jacob, Commissariat à l'Energie Atomique (CEA), Université Paris-Saclay, F-91000 Evry, France

[3]Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, F-69622 Villeurbanne, France

[4]Université de Toulon, Aix-Marseille Université, CNRS/INSU/IRD, Mediterranean Institute of Oceanography MIO UMR 7294, EMBIO, CS 60584, 83041 Toulon cedex 9, France

Abstract

Background: Copepods are among the most numerous animals, and play an essential role in the marine trophic web and biogeochemical cycles. The genus *Oithona* is described as having the highest density of copepods, and as being the most cosmopolite copepods. The *Oithona* male paradox describes the activity states of males, which are obliged to alternate between immobile and mobile phases for ambush feeding and mate searching, respectively, while the female is typically less mobile and often feeding. To characterize the molecular basis of this sexual dimorphism, we combined immunofluorescence, genomics, transcriptomics, and protein-protein interaction approaches.

Results: Immunofluorescence of β3- and α-tubulin revealed two male-specific nervous ganglia in the lateral first segment of the *Oithona nana* male's prosome. In parallel, transcriptomic analysis showed male-specific enrichment for nervous system development-related transcripts. Twenty-seven Lin12-Notch Repeat domain-containing protein coding genes (LDPGs) of the 75 LDPGs identified in the genome were specifically expressed only in males. Furthermore, most of the LDPGs (27%) coded for proteins having predicted proteolytic activity, and non-LDPG proteolysis-associated transcripts showed a male-specific enrichment. Using yeast double–hybrid assays, we constructed a protein-protein interaction network involving two LDPs with proteases, extracellular matrix proteins, and neurogenesis-related proteins.

Conclusions: For the first time, our study describes the lateral nervous ganglia of *O. nana* males, unique to copepods. We also demonstrated a role of LDPGs and their associated proteolysis in male-specific physiology, and we hypothesize a role of the LDPGs in the development of the lateral ganglia through directed lysis of the extracellular matrix for the growth of neurites and genesis of synapses.

51

52

Background

Copepods are small planktonic crustaceans that represent the most abundant metazoan subclass on Earth, and occupy all ecological aquatic niches [1, 2]. Among them, the genus *Oithona* is described as having the highest numerical density [3], being the most cosmopolitan [4], and playing a key role as a secondary producer in the marine food web and in biogeochemical cycles [5]. Due to its importance and abundance, *Oithona* phylogeography, ecology, behavior, life cycle, anatomy, and genomics are well studied [6–14].

Oithona spp. are active ambush-feeding omnivores; that is, to feed, the individuals remain static, jump on prey that come on their range, and capture them with their buccal appendages [8]. While females spend the majority of time feeding, and are thus mostly immobile, males actively seek females for mating. The mating success of males thus increases by being motile and not feeding. Theoretically, to maximize mating success, males have to alternate feeding and female-searching periods, which constitutes a paradox in the *Oithona* male behavior [8].

We performed a multi-year survey in Little Bay of Toulon, where *O. nana* is the dominant zooplankton throughout the year with no significant seasonal variation, suggesting continuous reproduction [15] as observed in other *Oithona* populations [16]. Under laboratory conditions with the two sexes incubated separately, *O. nana* males have a mean lifetime of 25 days, and females live for approximately 42 days. However, the lifespan *in situ* is unknown, as is the reproduction rate. Nonetheless, in the case of female saturation, *O. davisae* males have a reproduction rate of 0.9 females male^{-1} day^{-1}, depending on the production of spermatophores that are transferred during mating [8]. A strongly biased sex-ratio toward females (male/female ratio < 0.22) was observed in the *O. nana* population of Toulon Little Bay (France) [15]. Several causes could explain this observation; one possibility is the higher male exposure to predators due to its higher motility, considered as a “risky” behavior [17,

18]. However, other possibilities can be proposed, such as environmental sex determination (ESD) that has been observed in other copepods [19], or energy resource depletion as a consequence of male energy consumption during the mate search.

Recently, the *O. nana* genome was sequenced, and its comparison to other genomes showed an explosion of Lin12-Notch-Repeat (LNR) domain-containing protein-coding genes (LDPGs) [9]. Among the LDPGs present in the genome, five were found under natural selection in Mediterranean Sea populations, including notably one point mutation generating an amino acid change within the LNR domain of a male-specific protein [9, 20]. This provided the first evidence of *O. nana* molecular differences between sexes at the transcriptional level, and a potential new repertoire of candidate genes for functional analysis.

To further investigate the molecular basis of *O. nana* sexual differentiation, we used in this study a multi-approach analysis including, (i) *in situ* sex ratio determination over a fifteen-year time series, (ii) sexual system determination by sex-specific polymorphism analysis, (iii) immunofluorescence staining of the nervous system in male and female *O. nana* copepods, (iv) *in silico* analysis of the structure and evolution of the LDPGs, (v) sex-specific gene expression through RNA-seq analysis, and (vi) the identification of an LDP protein-protein interaction network using a yeast two-hybrid system.

Results

A female-biased sex ratio of *Oithona nana*

Between 2002 and 2017, 186 samples were collected in the Toulon Little Bay (Figure 1A), from which *O. nana* female and male adults were isolated (Figure 1B). Across fifteen years of observation, we noted minimum male/female ratios in February (0.11), maxima in September, October, and November (0.17), and a mean sex-ratio of 0.15 ± 0.11 over all years (Figure

1C). This monitoring of the sex-ratio showed a strong bias toward females, with relative stability over the years (ANOVA, $p = 0.87$).

Central nervous system labelling by immunofluorescence

The central nervous system labeling with β3-tubulin in *O. nana* (Figure 2A–B) showed the male-specific presence of a lateral ganglion with a high density of β3-tubulin in the anterior part of the ganglion, and a higher density of nuclei in its posterior part. We also observed the presence of β3-tubulin-rich post-ganglionic nerves that possibly connect the anterior part of the lateral ganglion to the tritocerebrum and/or the subesophageal ganglion. Not all males presented this labeling, and certain males contained one ganglion symmetrically on each side. However, no females presented this ganglion (Figure 2C). The α-tubulin labeling showed about seven to nine parallel afferent nerve fibers on the lateral part of the ganglion (Figure 2D–E) connected to free nerve endings located in the ventral part of the ganglion and in the external environment. Such labeling was absent in *O. nana* females. Together, these assays indicated the presence of a new nerve ganglion present only in *O. nana* males, located on the anterolateral part of the prosome.

Transcriptomic support for *Oithona nana* male homogamety

To identify the most likely cause of this sex-ratio bias between potential environmental sex determination (ESD) and higher male mortality, we used SD-pop on four individual transcriptomes of both sexes to first determine the *O. nana* sexual system. According to SD-pop, the ZW model was preferred (lowest Bayesian information criterion (BIC)) for *O. nana*. This result is unlikely to be due to chance, as in none of the runs on the 69 datasets for which the sex was permuted did the ZW model have the lowest BIC. Eleven genes had a probability

of being sex-linked in *O. nana* greater than 0.8; however, none of the SNPs in these genes showed the typical pattern of a fixed ZW SNP. The four females genotyped were heterozygous, and the four males were homozygous (except for some SNPs, for which one male individual was not genotyped), indicating that the recombination suppression between the gametologs is recent, and that no or few mutations have been fixed independently in both gametolog copies. Annotation of these eleven genes shows that only one shared homology with other metazoan genes, that being *ATP5H*, which codes a subunit of the mitochondrial ATP synthase (Supplementary Notes S5). As in *Drosophila*, *O. nana ATP5H* is encoded in the nucleus [21].

LNR domains burst in the *O. nana* proteome

To identify LDPs, we developed a HMM dedicated to *O. nana* LNR identification based on 31 conserved amino acid residues. In the *O. nana* proteome, 178 LNR and LNR-like domains were detected, encoded by 75 LDPGs, while a maximum of eight domains coded by six LDPGs were detected in four other copepods (Figure 3A–B). Among the 178 *O. nana* domains, 22 were canonical LNR and 156 were LNR-like domains (Figure 3C). By comparing the structure of Notch, LNR, and LNR-like domains, we observed the loss of two cysteines (Figure 3C) in the LNR-like domains. Among the 75 LDPs, we identified nine different protein structure patterns (Figure 3D), including notably 47 LNR-only proteins, 12 trypsin-associated LDPs, and eight metallopeptidase-associated LDPs. Overall, LDPs were predicted to contain a maximum of 5 LNR domains and 13 LNR-like domains.

Forty-nine LDPs were predicted to be secreted (eLDP), six membranous (mLDPs), and twenty intracellular (iLDPs) (Supplementary Notes S6). Among the iLDPs, two were associated with proteolytic domains, three associated with sugar-protein or protein-protein interaction domains (PAN/Appel, lectin, and ankyrin domains), and 13 (65%) were LNR-only

proteins. Among the eLDPs, 18 (37%) contained proteolytic domains corresponding to a significant enrichment of proteolysis in eLDPs (hypergeometric test, p = 2.13e − 17); other eLDPs corresponded to LNR-only proteins (63%). The mLDPs were represented by one Notch protein, two proteins with LNR domains associated with lectin or thrombospondin domains, respectively, and three LNR-only proteins.

In phylogenetic trees based on nucleic acid sequences of the LNR and LNR-like domains (Figure 3E), only 17% of the nodes had support over 90%. Twenty-seven branch splits corresponded to tandem duplications involving 15 LDPGs, including Notch and a cluster of five trypsin-associated LDPGs coding three eLDPs and two iLDPs.

***Oithona nana* male gene expression**

Among the 15,399 genes predicted in the *O. nana* reference genome, 1,233 (~8%) were significantly differentially expressed in at least one of the five developmental stages. Among them, 619 genes were specifically upregulated in one stage, with 53 genes upregulated in the egg, 19 in nauplii, 75 in copepodids, 27 in adult females, and 445 in adult males (Figure 4A). The male-upregulated genes were categorized based on their functional annotation (Figure 4B).

Upregulation of LNR-coding and proteolytic genes in males

The 1,233 differentially expressed genes contained 27 LDPGs (36% of total LDPGs) (Figure 4C). Of these 27 genes, 18 were specifically upregulated in adult males, producing a significant and robust enrichment of LDPGs in the adult male transcriptomes (fold change > 8; p = 2.95e − 12) (Figure 4C). Among the 445 male-specific genes, 27 were predicted to play a role in proteolysis, including 16 trypsins with three trypsin-associated LDPGs, showing

174 significant enrichment of trypsin-coding genes in males (p = 1.73e – 05), as well as three
175 metalloproteinases and five proteases inhibitors.

176

177 *Upregulation of nervous system-associated genes in adult males*

178 Forty-eight upregulated genes in males had predicted functions in the nervous system
179 (Supplementary Notes S7). These included 36 genes related to neuropeptides and hormones,
180 either through their metabolism (10 genes, seven of which encode enzymes involved in
181 neuropeptide maturation and one of which is an allatostatin), through their transport and
182 release (9 genes), or through neuropeptide or hormone receptors (17 genes, seven of which
183 are FMRF amide receptors). Six genes were predicted to be involved in neuron polarization,
184 four in the organization and growth guidance of axons and dendrites (including homologs to
185 *B4GAT1* and zig-like genes), one in the development and maintenance of sensory and motor
186 neurons (*IMPL2*), and one in synapse formation (*SYG-2*, futsch-like).

187

188 *Upregulation of amino-acid conversion into neurotransmitters in male adults*

189 We observed 10 upregulated genes in males predicted to play a role in amino acid metabolism
190 (Figure 4D). This includes five enzymes that directly convert lysine, tyrosine, and glutamine
191 into glutamate through the activity of one α-aminoadipic semialdehyde synthase (AASS), one
192 tyrosine aminotransferase (TyrAT), and three glutaminases, respectively. Three other
193 upregulated enzymes play a role in the formation of pyruvate: one alanine dehydrogenase
194 (AlaDH), one serine dehydrogenase (SDH), and, indirectly, one phosphoglycerate mutase
195 (PGM). Furthermore, two other upregulated enzymes are involved in the formation of glycine,
196 one sarcosine dehydrogenase (SARDH) and one betaine-homocysteine methyltransferase
197 (BHMT) (Figure 3D).

198

199 *Downregulation of food uptake regulation in male adult*

200 Three genes with predicted functions in food uptake regulation showed specific patterns in
201 males. These included an increase of the mRNA encoding allatostatin, a neuropeptide known
202 in arthropods to reduce food uptake, but also three male under-expressed genes: a crustacean
203 cardioactive peptide (CCAP, a neuropeptide that triggers digestive enzymes activation), and
204 the two bursicon protein subunits, which encode hormones known to be involved in intestinal
205 and metabolic homeostasis.

206

207 **Protein-protein interaction network involving LDPs and IGFBP**

208 In order to further characterize the role of LDPs, we studied their potential protein interactions
209 using a yeast two-hybrid system (Y2H). To this end, we selected 11 genes: 7 male-
210 overexpressed LDPGs, and 4 potential IGFBPs. The choice of the IGFBPs was made with the
211 hypothesis that potential insulin-like androgenic gland hormone partners would be found in
212 decapods [22].

213 We performed Y2H analysis using two different approaches (Supplementary Notes
214 S3): the first was a matrix-based screen with pairwise interaction assays, and the second
215 aimed to identify potential interactors in the entire *O. nana* proteome by a random library
216 screen. This latter screening approach was more time-consuming, and was applied to only a
217 subset of four genes (two LDPGs and two IGFBPs) used as bait proteins against a Y2H
218 library constructed from *O. nana* cDNAs. Together, these two approaches allowed the
219 reconstruction of a protein network containing 17 proteins, including two LDPs and one
220 IGFBP used as baits (Figure 5A) (Supplementary Notes S4), and 14 interacting partners, of

which six have orthologs in other metazoans and five have no orthologs, but at least one of which was detected using the InterProScan domain (Figure 5B).

On_LDP1, a putative extracellular trypsin-containing LDP, was found to form a homodimer and interact with a trypsin, two extracellular matrix proteins, and also an insulin-like growth factor binding protein (On_IGFBP7) that contained a trypsin inhibitor kazal domain. Based on its phylogeny, this protein is homologous to IGFBP7, also present in vertebrates (Supplementary Notes S8).

On_IGFBP7 formed a homodimer and interacted with three other proteins: one spondin-1 like protein (On_Spon1-like) containing a kazal domain, one thrombospondin domain-containing protein, and one vitellogenin 2-like protein (On_Vtg2).

On_LDP2, coded by a gene upregulated in males (Figure 5C), interacted with nine proteins: one vitellogenin 2-like protein, three uncharacterized proteins, one homolog to somatomedin-B thrombospondin type 1 domain-containing protein, one homolog to the neuroendocrine protein 7b2 that contains a secretogranin V-like domain, one wnt5-like protein, one laminin 1 subunit β, and one furin-like 1 protein. No PPIs with IGF were detected, and no homologs to insulin-like androgenic gland hormone [22] were found in the *O. nana* proteome.

Discussion

ZW sexual system and high mortality rate of *O. nana* males

Over 15 years of sampling, we observed a stable and strongly female-biased sex-ratio (~1:9) in the Toulon Little Bay. A similar observation on a smaller time scale was previously made in another *O. nana* population [16] and in 132 other Oithonidae populations [18]. Two main factors could lead to this sex-ratio: a higher mortality of males, or environmentally-induced

sex determination. As we showed in our study, *O. nana* seems to have recently evolved genetic sex determination of the ZW type; thus, a 1:1 sex-ratio is expected in eggs. Therefore, a higher mortality rate of males is more likely to explain our *in situ* observations. Moreover, these results are in accordance with the previously described risky behavior of *Oithona* males, that is, frequent motion to find females, thus making them more vulnerable to predators than immobile females [17].

The development of a male-specific nervous ganglion is supported by nervous system-gene expression

Immunofluorescence labeling of the *O. nana* nervous system demonstrated the existence of a ganglion in the anterolateral part of the male prosome. α-tubulin labeling showed nervous termination in the external environment, which suggests a sensory role to external cues. These observations constitute a new nervous and anatomical dimorphism between *O. nana* males and females [23][24]. The male-specific overexpression of numerous genes involved in the development of the nervous system supports the development of this male-specific ganglion at the molecular level, including axon guidance and synapse formation. From the upregulation of neuronal developmental genes in adult males, especially syg-2 and zig-8 normally expressed during the larval phase [25], we may infer ongoing formation of new axons and/or dendrites and synapses associated with the male lateral ganglion (Figure 6). On_LDPG2, a male-overexpressed gene under natural selection in the Mediterranean Sea populations, has been shown to code an eLDP. The latter interacts with two proteins involved in the nervous system development, On_Wnt5 and On_Lamβ1 (Figure 6), notably important in axon guidance [26, 27]. Therefore, through its protein-protein interactions, On_LDP2 may modulate neurogenesis in males and participate in the sexual dimorphism of the *O. nana* nervous system. It has been demonstrated that *O. davisae* males have a preference for virgins

[28]; we could also speculate that the lateral ganglion may help in virgin-sensing of *O. nana* through the recognition of virgin-specific chemical cues.

LDP-driven proteolysis is potentially linked to the male ganglion formation

The explosion of LDPGs in the *O. nana* genome is unique in metazoans, and is associated with the formation of new protein structures containing proteolytic domains. Owing to the small size of the LNR domain (~40 amino acids) and the substantial polymorphism within LNR domain sequences, the deep branches of the tree are weakly supported, and the evolutionary scenario resulting in the amplification of the domain remains undetermined. Two previous studies on *O. nana* population genomics [9, 20] identified five LDPGs under natural selection, with point mutations within an LNR domain. These results reinforce the idea of an ongoing evolution of these domains, especially in *O. nana* males, and support an important role for these genes in the male biology.

In metazoans, LNR domains are known to be involved in extracellular PPI [29] and cleavage site-accessibility modulation [30]. Half of the eLDPs contain LNR domains only, and the other half associates with peptidases (trypsin and metalloproteinase). From the PPI network, we showed that On_LDP1 interacts with different types of extracellular proteins involved notably in tissue structure, energy storage, and extracellular proteolysis. On the other hand, transcriptomic analyses showed an enrichment of extracellular trypsins in adult males (Figure 4C). This information supports the lysis of extracellular proteins in males involving complex protein interactions between eLDPs and trypsins. From our transcriptomic results on the expression of LDPs and nervous system genes, we hypothesize that the LDPs may play a role in the development of the male-specific ganglion by guiding proteolysis of the extracellular matrix around the ganglion. This lysis may help and guide the development of new neurites and the formation of synapses by preserving other anatomical structures.

However, our study lacks a comprehensive or precise view of the function of the LDPs, due to the labor and time needed to optimize protocols for functional analysis on non-model species. Indeed, several RNAi assays were performed during this study without success due to the death of transfected individuals. However further investigations and methodological developments may help to validate our hypothesis of the role of LDPs in the neural development of *O. nana* male-specific lateral ganglion.

Conclusions

The hyper-motility of *O. nana* ZZ males and their faculty to find females could be one of the factors of the ecological success of *O. nana*, and explain the observed female-dominated sex ratio over a fifteen year course in the Toulon Little Bay. The male-specific nervous ganglion may play a role in its sexual behavior, the development of which was linked at the transcriptomic and protein levels to LDPGs, which seem to play an important role in the male-specific neurogenesis and proteolysis. The presence of LDPGs in other *Oithona* species should be investigated to better understand the evolution and role of these genes in other cyclopoid species.

Abbreviations

BIC: Bayesian information criterion
CNS: Central Nervous SystemHMM: Hidden Markov Model
LDP: LNR Domain-containing Protein
LDPG: LNR domain-containing Protein coding Gene
LNR: Lin12 Notch Repeat
ORF: Open Reading Frame
Y2H: Yeast two hydrid
PPI: Protein-protein interaction
IGFBP7: Insulin-like Growth Factor Binding Protein 7

Methods

Sex ratio in the Toulon Little Bay

Oithona nana specimens were sampled on the East side of the Toulon Little Bay, France (Lat. 43° 06' 52.1" N, Long. 05° 55' 42.7" E) which is not a protected area and does not require any permission to catch plankton according to local laws and regulations http://www.var.gouv.fr/IMG/pdf/2017_16_ap_reglementant_la_navigation_le_mouillage_la_baignade_et_la_plongee_dans_les_eaux_maritimes_de_la_rade_de_toulon.pdf. The samples were collected from the upper water layer (0–10 m) using zooplankton nets with a mesh of 90 and 200 µm. Samples were preserved in 5% formaldehyde. The monitoring of *O. nana* in Toulon Little Bay was performed monthly from 2002 to 2017. Individuals of both sexes were identified and counted under a stereomicroscope.

Immunofluorescence analysis

Adult individuals of both sexes collected in June, 2017 were isolated under a stereomicroscope and washed in 0.5X PBST, and incubated for 72 h at room temperature with rabbit β3-tubulin antibody (ab179513, 1:100) and 100 µl normal donkey serum (d9663). After washing with PBST, the individuals were incubated for 72 h with cy3 goat anti-rabbit conjugated Ig-antibody (AB2338000, 1:200, Jackson ImmunoResearch Laboratory) and normal donkey serum. Individuals were then washed with PBST and mounted on slides and microslides with Citifluor mounting solution (AF87). A similar protocol was used to label other individuals for α-tubulin using rabbit α-tubulin antibody (ab15446, 1:100). Observations were made on an Olympus BX43 fluorescence microscope, and images were taken with Toupview, then labeled with Inkscape 0.92.4.

Biological materials and RNA-seq experiments

Plankton sampled in November, 2015 and November, 2016 in the Little Bay of Toulon, France were preserved in 70% ethanol and stored at −20°C. The copepods were isolated under the stereomicroscope as previously described. We selected *O. nana* individuals from five different development stages: five pairs of egg-sacs, four nauplii (larvae), four copepodites (juveniles), four adult females and four adult males. All individuals were isolated from the November, 2015 sampling, except for the eggs. Each individual was isolated, then crushed with a tissue grinder (Axygen) into a 1.5 ml Eppendorf tube. Total mRNAs were extracted with the NucleoSpin RNA XS kit (Macherey-Nagel) following the manufacturer's instructions, then quantified on a Qubit 2.0 with the RNA HS Assay kit (ThermoFisher Scientific); quality was assessed on a Bioanalyzer 2100 with the RNA 6000 Pico Assay kit (Agilent). cDNAs were constructed using the SMARTer v4 Ultra low Input RNA kit (Takara). After cDNA shearing using a Covaris E210 instrument, Illumina libraries were constructed using the NEBNext Ultra II kit (New England Biolabs) and sequenced on an Illumina HiSeq2500. A minimum of $9.7e^{6}$ reads pairs were produced from each individual (Supplementary Notes S1).

Sex-determination system identification by RNA-seq

RNA-seq reads from both sexes (four females and four males) were aligned against *O. nana* genes. Reads having an alignment length lower than 80% and nucleic identity lower than 97% were removed. The variant calling step was performed with the "*samtools mpileup*" and "*bcftools call*" commands, with default parameters [31], and only bi-allelic sites were kept.

To identify the most likely sexual system in *O. nana*, we used *SD-pop* [32]. Just like its predecessor *SEX-DETector* [33], *SD-pop* calculates the likelihood of three sexual models (the absence of sex chromosomes, the XY system, or the ZW system), which can be compared using the Bayesian information criterion. *SD-pop* is based on population genetics

(i.e., Hardy-Weinberg equilibrium for autosomal genes, and different equilibria for sex-linked genes) instead of Mendelian transmission from parents to offspring, and thus can be used without the requirement of a controlled cross.

The number of individuals used (four for each sex) is close to the lower limit for *SD-pop*, where the robustness of the method is weakening. To test whether the model preferred by *SD-pop* could have been preferred purely by chance, we permuted the sex of the individuals, with the constraint of keeping four females and four males ((8!/(4!*4!) - 1 = 69 permuted datasets). As the XY model is strictly equivalent to the ZW model with the sexes of all individuals changed, two *SD-pop* models (no sex chromosomes, and ZW) were run on all possible permutations of the data, and the BIC of each model was calculated. The genes inferred as sex-linked based on their posterior probability (>0.8) were manually annotated.

Arthropoda **phylogenetic tree**

The ribosomal 18S sequences from seven arthropods, including five copepods (*O. nana, Lepeophtheirus salmonis, Tigriopus californicus, Eurytemora affinis, Calanus glacialis, Daphnia pulex,* and *Drosophila melanogaster*) were downloaded from NCBI. The sequences were aligned with MAFFT [34] using default parameters. The nucleotide blocks conserved among the seven species were selected by Gblock on Seaview [35] and manually curated. The maximum-likelihood phylogenetic tree was constructed using PhyML 3.0 with the General Time Reversible (GTR) model and branch supports computed by the approximate likelihood ratio test (aLRT) [36].

Gene annotation

The functional annotation of genes was updated from the previous genome annotation [9] using InterProScan v5.8-49.0 [37], BlastKOALA v2.1 [38], and by alignment on the NCBI non-redundant protein database using Diamond [39]. Furthermore, a list of *O. nana* genes under natural selection in the Mediterranean Sea was added based on previous population genomic analyses [20]. We further considered the annotation provided by either (i) Pfam [40] or SMART [41] protein domains, (ii) GO terms (molecular function, biological process, or cellular component) [42], (iii) KEGG pathways [43], and (iv) the presence of loci under natural selection. These four gene features were used to identify specific enrichment in a given set of genes using a hypergeometric test that estimated the significance of the intersection between a specific gene list and one of the four global annotation lists.

HMM search for LDPGs identification

From the InterProScan annotation of the *O. nana* proteome, 25 LNR domain sequences were detected ($p \leq 10^{-6}$), extracted, and aligned with MAFFT using default parameters [34]. A Hidden Markov Model (HMM) was generated from the aligned sequences using the "*hmmbuild*" function of the HMMER tool version 3.1b1 [44]. The *O. nana* proteome was scanned by "*hmmsearch*" using the LNR HMM profile. Detected domains were considered canonical LNR domains for *E-value*, *c-E-value*, and *i-E-value* $< 10^{-6}$ and containing at least six cysteines, or considered as LNR-like domains for *E-value*, *c-E-value*, and *i-E-value* between 10^{-6} and 10^{-1} and containing at least four cysteines. A weblogo [45] was generated to represent conserved residues for three LNRs of the Notch protein, and the LNRs and LNR-like proteins detected by HMM. The LNR and LNR-like domain-containing proteins constituting the final LDP set was used for further analysis. Deep-Loc (online execution) [46] was used to determinate the cellular localization of LDPs. To detect signal peptides and

membrane protein topology, we used the online services of SignalP 5.0 [47] and TOPCONS [48], respectively.

Phylogeny tree of *O. nana* LNR domains

The *O. nana* nucleotide sequences of the LNR and LNR-like domains were aligned using MAFFT with default parameters. The maximum-likelihood phylogenetic tree was constructed using PhyML.3.0 with a model designed by the online execution of Smart Model Selection v.1.8.1 (Lefort, Longueville, and Gascuel 2017), and with branch supports computed by the aLRT method. The GTR model was used with an estimated discrete gamma distribution (a = 1.418) and a proportion of fixed invariable sites (I = 0.25). The tree was visualized using MEGA-X (Kumar et al. 2018).

Differential expression analysis

RNA-seq reads from 20 libraries were mapped independently against the *O. nana* genes using "*bwa-mem*" (v. 0.7.15-r1140) with default parameters [49]. Read counts were extracted from the 20 BAM files with samtools (v. 1.4) [31]. Each set of reads was validated by a pairwise MA-plot to ensure a global representation of the *O. nana* transcriptome in each sample (Supplementary Notes S2). One nauplius sampled showing a biased read count distribution was discarded. Read counts from valid replicates were used as input data for the DESeq R package [50] to identify differential gene expression between the five development stages through pairwise comparisons of each developmental stage. Genes having a Benjamini-Hochberg corrected *p*-value ≤ 0.05 in one of the pairwise comparisons were considered significantly differentially expressed. To identify stage-specific genes, we selected those that were at least twice as highly expressed based on the normalized read count mean

443 (log_2(foldChange) > 1) in one development stage compared to the four others. Upregulated
444 stage-specific genes were represented by a heatmap. The same method was used to
445 determined downregulated stage-specific genes (with log2(foldChange) < −1).

446

447 **Protein-protein interaction assays by yeast two-hybrid screening**

448 Yeast two-hybrid experiments were performed using the Matchmaker Gold Yeast Two-
449 Hybrid System (Takara). The coding sequences were first cloned into the entry vector
450 pDONR/Zeo (ThermoFisher), and the correct ORF sequences were verified by Sanger
451 sequencing. To this end, LDPGs were PCR-amplified with Gateway-compatible primers
452 (Supplementary Notes S3, S4) using cDNAs of pooled male individuals as template. In the
453 case of secreted proteins, the amplified ORF lacked the signal peptide. Then, the cloned ORFs
454 were reamplified by a two step-PCR protocol allowing the creation of a recombination
455 cassette containing the ORF flanked by 40-nucleotide tails homologous to the ends of the
456 pGBKT7 bait vector at the cloning site. Linearized bait vectors and ORF cassettes were co-
457 transformed into the Y2HGold yeast strain, and ORF cloning was performed by homologous
458 recombination directly in yeast. Y2H screening for potential interacting partners of the bait
459 proteins was performed via two methods: first, by directly testing pairs of candidates, and
460 second, by testing the candidates as baits against a cDNA library obtained from a pool of total
461 mRNAs from 100 *O. nana* male individuals cloned into pGAD-AD prey vectors.

462 Before screening, the self-activity of each bait clone was tested by mating with the
463 Y187 strain harboring an empty pGADT7-AD vector, and then plating on SD/-His/-Leu/-Trp
464 medium supplemented with 0, 1, 3, 5, or 10 mM 3-amino 1,2,4-triazole (3-AT). Each bait
465 clone was then mated with the prey library containing approximately 4 x 10^6 individual
466 clones, and plated on low-stringency agar plates (SD/-Trp/-Leu/-His) supplemented with the
467 optimal concentration of 3-AT based on the results of the self-activity test. To decrease the

false positive rate, after five days of growth at 30°C, isolated colonies were spotted on high-stringency agar plates (SD/-Leu/-Trp/-Ade/-His) supplemented with 3-AT and allowed to grow another five days. Colony PCR on positive clones growing on high stringency medium was performed with primers flanking the cDNA insert on the pGAD-AD vector, and PCR products were directly Sanger-sequenced.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

The *O. nana* RNA-seq data are available at ENA (*Supplementary Notes* S1) and *O. nana* samples are available upon request.

Competing interests

We declare that none of the authors have any competing of interests.

Funding

The CEA-Genoscope and France Génomique (grant ANR-10-INBS-09) financed the wet laboratory work, the RNA-seq sequencing and the computational cost of the bioinformatic analyses. The French Ministry of Research financed the PhD grant of Kevin Sugier.

Authors' contributions

KS, JP, and JLJ collected the samples; JLJ generated the sex-ratio data; KS, KL, MAM, EP, and JP generated the molecular data; RLJ, BV, and MAM performed the immunofluorescence analysis. JK performed sexual system analysis; KS, AA, LB, SK, NM, and CO performed the yeast two-hybrid analysis; AA designed the yeast two-hybrid method; KS and MAM per-

formed the analyses; PW provided technical facilities support; KS, AA, and MAM wrote the manuscript; MAM supervised the study. All authors have read and approved the manuscript.

Acknowledgements

We acknowledge the Commissariat à l'Energie Atomique et aux Energie Alternatives and the Agence Nationale de la Recherche

References

1. Huys R, Boxshall GA. Copepod evolution. Ray Society; 1991.

2. Kiørboe T. What makes pelagic copepods so successful? J Plankton Res. 2011;33:677–85.

3. Gallienne CP, Robins DB. Is Oithona the most important copepod in the world's oceans? J Plankton Res. 2001;23:1421–32.

4. Nishida S. Taxonomy and distribution of the family oithonidae (Copepoda, Cyclopoida) in the Pacific and Indian. Bull Ocean Res Inst. 1985;20:1–167.

5. Steinberg DK, Landry MR. Zooplankton and the Ocean Carbon Cycle. Ann Rev Mar Sci. 2017;9:413–44.

6. Cornils A, Wend-Heckmann B, Held C. Global phylogeography of Oithona similis s.l. (Crustacea, Copepoda, Oithonidae) – A cosmopolitan plankton species or a complex of cryptic lineages? Mol Phylogenet Evol. 2017;107:473–85.

7. Dvoretsky VG, Dvoretsky AG. Life cycle of Oithona similis (Copepoda: Cyclopoida) in Kola Bay (Barents Sea). Mar Biol. 2009;156:1433–46.

8. Kiørboe T. Mate finding, mating, and population dynamics in a planktonic copepod Oithona davisae: There are too few males. Limnol Oceanogr. 2007;52:1511–22.

9. Madoui M-A, Poulain J, Sugier K, Wessner M, Noel B, Berline L, et al. New insights into global biogeography, population structure and natural selection from the genome of the epipelagic copepod *Oithona*. Mol Ecol. 2017;26:4467–82.

10. Mironova E, Pasternak A. Female gonad morphology of small copepods Oithona similis and Microsetella norvegica. Polar Biol. 2017;40:685–96.

11. Paffenhöfer G-A. On the ecology of marine cyclopoid copepods (Crustacea, Copepoda). J Plankton Res. 1993;15:37–55.

12. Sugier K, Vacherie B, Cornils A, Wincker P, Jamet J-L, Madoui M-A. Chitin distribution in the Oithona digestive and reproductive systems revealed by fluorescence microscopy. PeerJ. 2018;6:e4685.

13. Zamora-Terol S, Kjellerup S, Swalethorp R, Saiz E, Nielsen TG. Population dynamics and production of the small copepod Oithona spp. in a subarctic fjord of West Greenland. Polar Biol. 2014;37:953–65.

14. Zamora Terol S. Ecology of the marine copepod genus Oithona. Universitat Politècnica de Catalunya; 2013.

15. Richard S, Jamet J. An Unusual Distribution of Oithona nana GIESBRECHT (1892) (Crustacea : Cyclopoida) in a Bay : The Case of Toulon Bay (France, Miditerranean Sea). J Coast Res. 2001; April 2015.

16. Temperoni B, Vinas MD, Diovisalvi N, Negri R. Seasonal production of Oithona nana Giesbrecht, 1893 (Copepoda: Cyclopoida) in temperate coastal waters off Argentina. J Plankton Res. 2011;33:729–40.

17. Hirst AG, Bonnet D, Conway DVPP, Kiørboe T. Does predation control adult sex ratios and longevities in marine pelagic copepods? Limnol Oceanogr. 2010;55:2193–206.

18. Kiørboe T. Sex, sex-ratios, and the dynamics of pelagic copepod populations. Oecologia. 2006;148:40–50.

19. Voordouw MJ, Anholt BR. Environmental sex determination in a splash pool copepod. Biol J Linn Soc. 2002;76:511–20.

20. Arif M, Gauthier J, Sugier K, Iudicone D, Jaillon O, Wincker P, et al. Discovering Millions of

549 Plankton Genomic Markers from the Atlantic Ocean and the Mediterranean Sea. Mol Ecol Resour.
550 2019;19:0–3.

551 21. Liao TSV, Call GB, Guptan P, Cespedes A, Marshall J, Yackle K, et al. An efficient genetic screen
552 in Drosophila to identify nuclear-encoded genes with mitochondrial function. Genetics.
553 2006;174:525–33.

554 22. Ventura T, Rosen O, Sagi A. From the discovery of the crustacean androgenic gland to the insulin-
555 like hormone in six decades. Gen Comp Endocrinol. 2011;173:381–8.

556 23. Frase T, Richter S. The brain and the corresponding sense organs in calanoid copepods – Evidence
557 of vestiges of compound eyes. Arthropod Struct Dev. 2020;54:100902.

558 24. Andrew DR, Brown SM, Strausfeld NJ. The minute brain of the copepod Tigriopus californicus
559 supports a complex ancestral ground pattern of the tetraconate cerebral nervous systems. J Comp
560 Neurol. 2012;520:3446–70. doi:https://doi.org/10.1002/cne.23099.

561 25. Shen K, Fetter RD, Bargmann CI. Synaptic Specificity Is Generated by the Synaptic Guidepost
562 Protein SYG-2 and Its Receptor, SYG-1. Cell. 2004;116:869–81.

563 26. Zou Y. Wnt signaling in axon guidance. Trends Neurosci. 2004;27:528–32.

564 27. Randlett O, Poggi L, Zolessi FR, Harris WA. The oriented emergence of axons from retinal
565 ganglion cells is directed by laminin contact in vivo. Neuron. 2011;70:266–80.

566 28. Heuschele J, Kiørboe T. The smell of virgins: mating status of females affects male swimming
567 behaviour in Oithona davisae. J Plankton Res. 2012;34:929–35. doi:10.1093/plankt/fbs054.

568 29. Boldt HB, Conover CA. Pregnancy-associated plasma protein-A (PAPP-A): A local regulator of
569 IGF bioavailability through cleavage of IGFBPs. Growth Horm IGF Res. 2007;17:10–8.

570 30. Sanchez-Irizarry C, Carpenter AC, Weng AP, Pear WS, Aster JC, Blacklow SC. Notch Subunit
571 Heterodimerization and Prevention of Ligand-Independent Proteolytic Activation Depend,
572 Respectively, on a Novel Domain and the LNR Repeats. Mol Cell Biol. 2004;24:9265–73.

573 31. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map
574 format and SAMtools. Bioinformatics. 2009;25:2078–9.

575 32. Käfer J, Lartillot N, Marais GAB, Picard F. Detecting sex-linked genes using genotyped
576 individuals sampled in natural populations. bioRxiv. 2020;:2020.01.31.928291.
577 doi:10.1101/2020.01.31.928291.

578 33. Muyle A, Käfer J, Zemp N, Mousset S, Picard F, Marais GAB. Sex-detector: A probabilistic
579 approach to study sex chromosomes in non-model organisms. Genome Biol Evol. 2016;8:2530–43.

580 34. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements
581 in performance and usability. Mol Biol Evol. 2013;30:772–80.

582 35. Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User Interface
583 for Sequence Alignment and Phylogenetic Tree Building. Mol Biol Evol. 2010;27:221–4.

584 36. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and
585 Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0.
586 Syst Biol. 2010;59:307–21.

587 37. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale
588 protein function classification. Bioinformatics. 2014;30:1236–40.

589 38. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for Functional
590 Characterization of Genome and Metagenome Sequences. J Mol Biol. 2016;428:726–31.

591 39. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat
592 Methods. 2015;12:59–60.

593 40. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein

families database. Nucleic Acids Res. 2014;42.

41. Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018;46:D493–6.

42. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25:25–9.

43. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 2012;40 Database issue:D109-14.

44. Eddy SR. Accelerated Profile HMM Searches. PLoS Comput Biol. 2011;7:e1002195.

45. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: A Sequence Logo Generator. Genome Res. 2004;14:1188–90.

46. Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics. 2017;33:3387–95.

47. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019;37:420–3.

48. Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res. 2015;43:W401-7.

49. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013.

50. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11:R106.

Figure Legends

Figure 1. Life cycle and sex-ratio of the copepod *Oithona nana* in the Toulon Little Bay. A. Map of the sampling site in the Toulon Little Bay created with the "maps" R package 3.3.0 and modified with Inkscape 0.92.4. **B.** The life cycle of *O. nana*. **C.** Sex ratio of *O. nana* from 2002 to 2017 in the Toulon Little Bay. Black circles represent the monthly mean, the blue line represents the 15-year mean (0.15).

Figure 2. Immunofluorescence of the *Oithona nana* nervous system. A. Male labeled with β3-tubulin and DAPI. **B.** Male labeled with β3-tubulin and DAPI. **C.** Female labeled with β3-tubulin and DAPI. **D.** Male labeled with α-tubulin and DAPI. **E.** Male labeled with α-tubulin and DAPI. **F.** Female labeled with α-tubulin and DAPI. P: protocerebrum. D: deutocerebrum. T: tritocerebrum. SEG: suboesophageal ganglion. VNS: ventral nervous system. LG: lateral ganglion. PGF: post-ganglion fibers. a: β3-tubulin-rich area. b: nuclei-rich area. ANF: afferent nerve fibers. FNE: free nerve ending

Figure 3. Lin-12 Notch Repeat (LNR) protein domain burst and domain associations in the *Oithona nana* proteome. A. Phylogeny of five copepod species and two other arthropod species based on 18S ribosomal sequences. The numbers at internal branches show aLRT branch support. The scale bar represents the nucleotide substitution rate. **B.** LNR domain occurrences in seven Arthropoda proteomes, detected by HMM. The number at the front of each bar corresponds to the number of detected genes. **C.** Consensus sequences of the *O. nana* Notch-LNR, LNR, and LNR-like domains generated by WebLogo. The asterisks represent the conserved sites. **D.** Schemata of the *O. nana* LNR and LNR-like protein structure. Numbers under each domain represent the possible occurrence range. The barplot represents the occurrence of the nine structures. **E.** Phylogenetic tree of the *O. nana* LNR and LNR-like domains. Bold branches have aLRT support ≥ 0.90. The red circles represent tandem duplication.

Figure 4. Differential expression analysis of the *Oithona nana* transcriptome. A. Heatmap of the 1,233 significantly differentially expressed genes in at least one of the five developmental stages. **B.** Functional annotation distribution of 445 genes explicitly overexpressed in male adults. **C.** Heatmap of the 27 significantly differentially expressed LDPGs and the composition of their protein domains. **D.** Amino acid conversion to neurotransmitters in *O. nana* males. Overexpressed enzymes in males are indicated in red, amino acids in blue, and neurotransmitter amino acids in green.

Figure 5. Protein-protein interaction of LNR-containing proteins in the *O. nana* male proteome. A. Structure and expression of the PPI candidates. The red arrows represent PCR primers. **B.** PPI network of LDPGs obtained by Yeast two-hybrid assays. Lamβ1: Laminin subunit beta 1. Vtg2: Vitellogenin 2. PUF: protein of unknown function. **C.** RPKM-normalized expression in the five developmental stages. From left to right: egg (e), larva (l), juvenile (j), adult female (f), adult male (m).

Figure 6. Hypothesis of the sensory role of the lateral ganglion in female search and the role of LDP-driven proteolysis. Red arrows represent protein-protein interactions found in this study. Blue arrows indicate enzymatic activities. Green circles represent overexpressed genes in males involved in the nervous system that were detected in this study. Dotted arrows indicate signaling pathways.

Supplementary Notes S1: Transcriptomic data.

Supplementary Notes S2: Transcriptomic data quality.

Supplementary Notes S3: Experimental design of the protein-protein interaction (PPI) analysis.

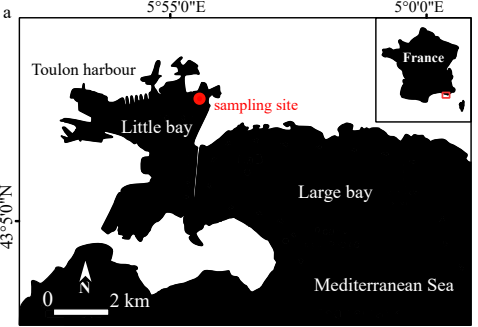
Supplementary Notes S4: Primers used for PCR amplification.

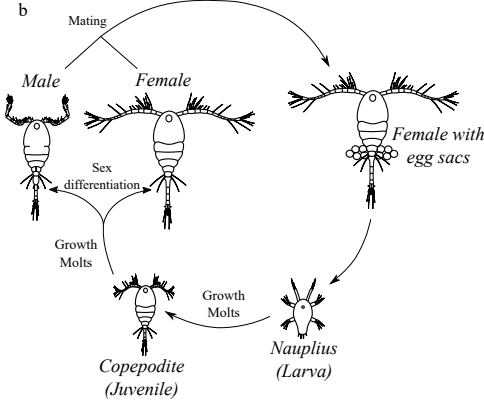
Supplementary Notes S5. Gene annotation of the sex-determination system-associated genes of *Oithona nana*.

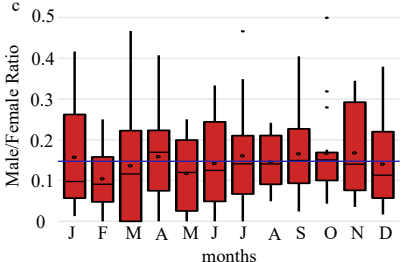
Supplementary Notes S6: Structure and localization of the Oithona *nana* LDPs.

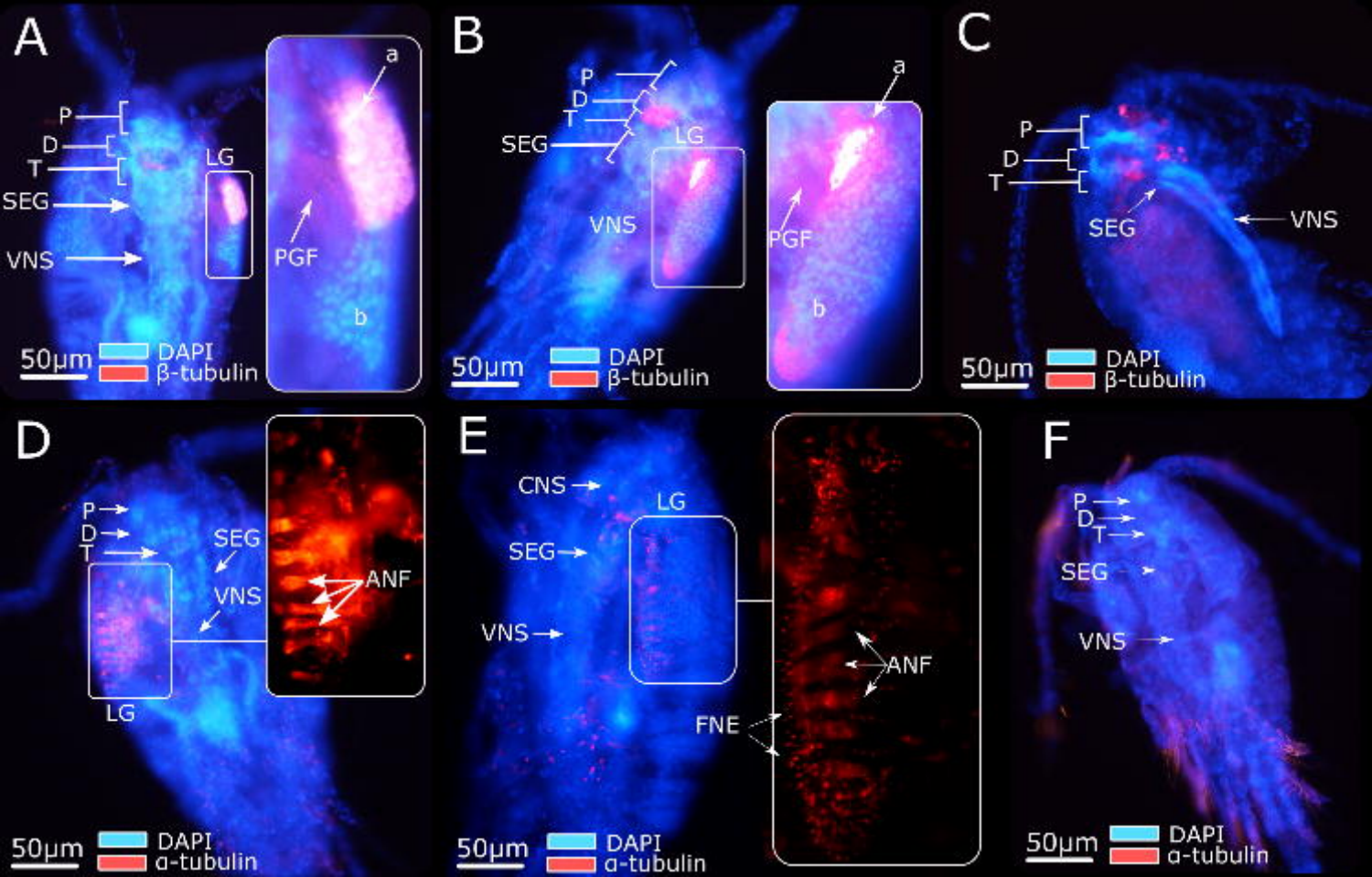
Supplementary Notes S7: Functional annotation of *O. nana* genes overexpressed in males.

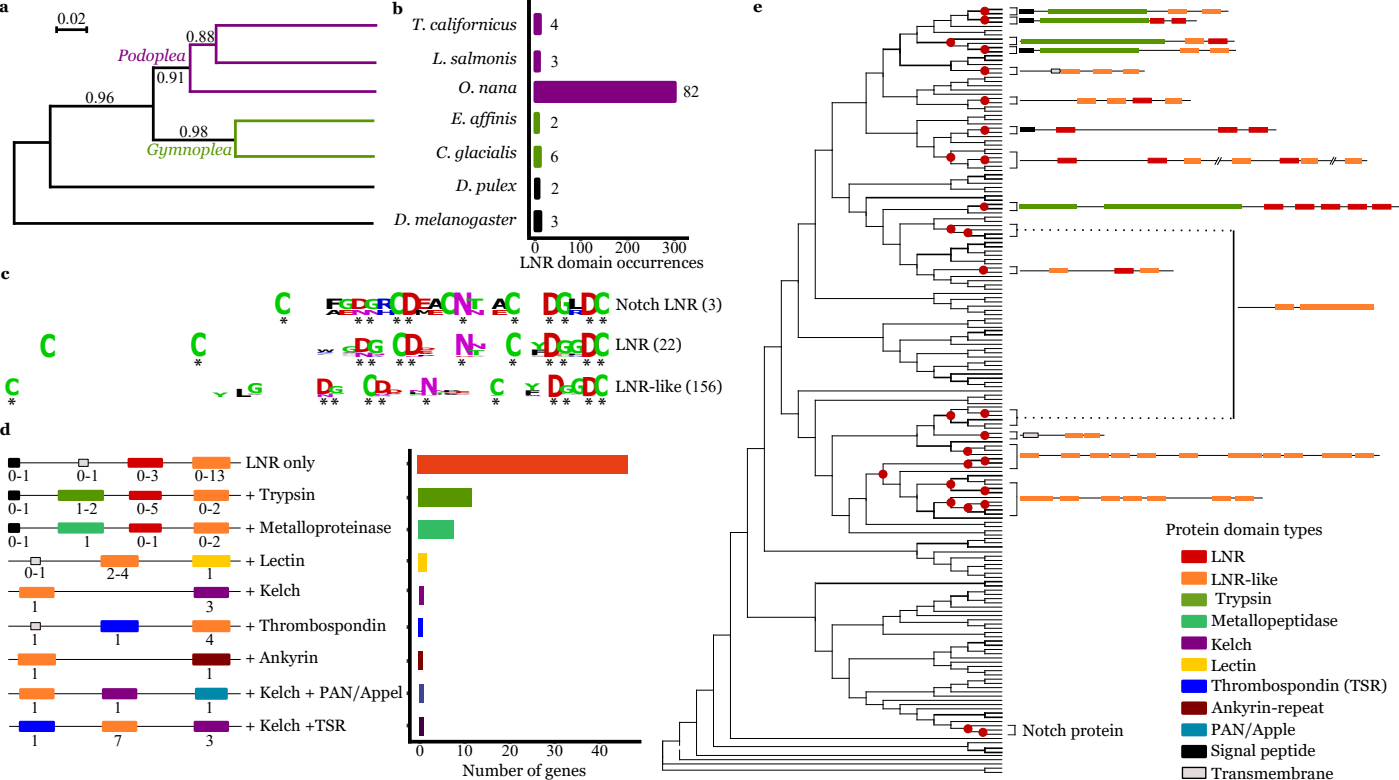
Supplementary Notes S8: Phylogenetic tree of On_IGFBP7.

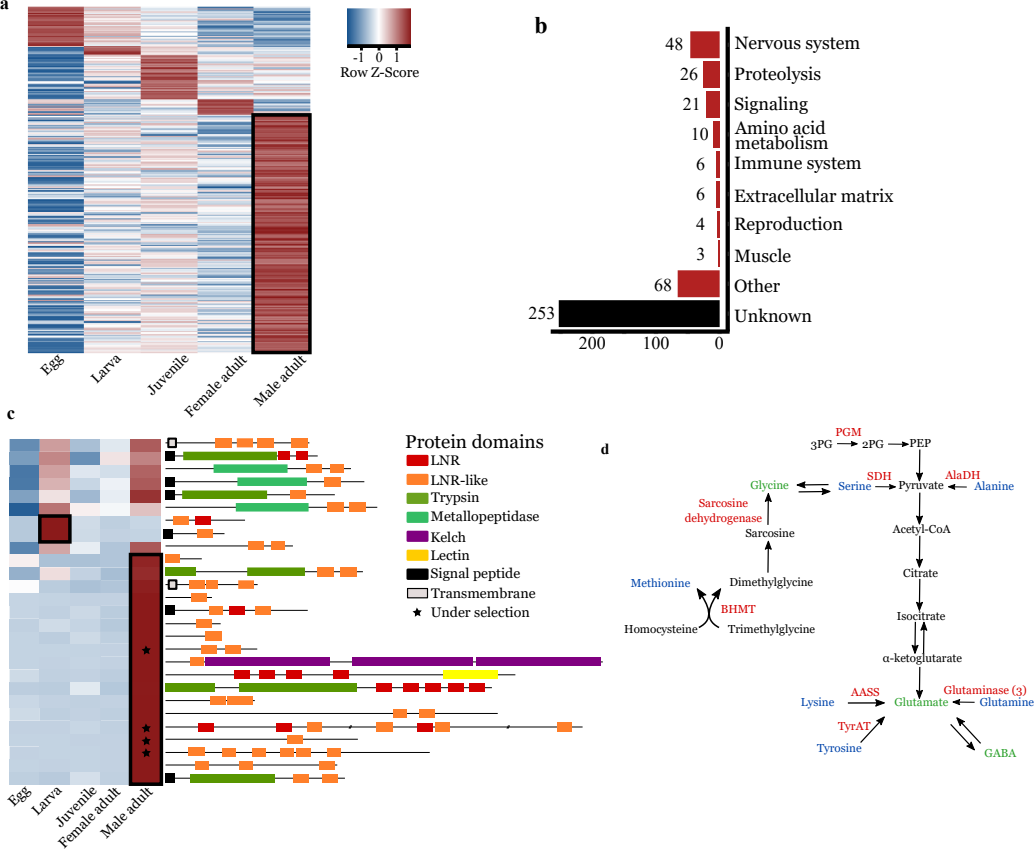

a
5°5'0"E
5°0'0"E
43°5'0"N
Toulon harbour
sampling site
Little bay
Large bay
Mediterranean Sea
France
N
0
2 km

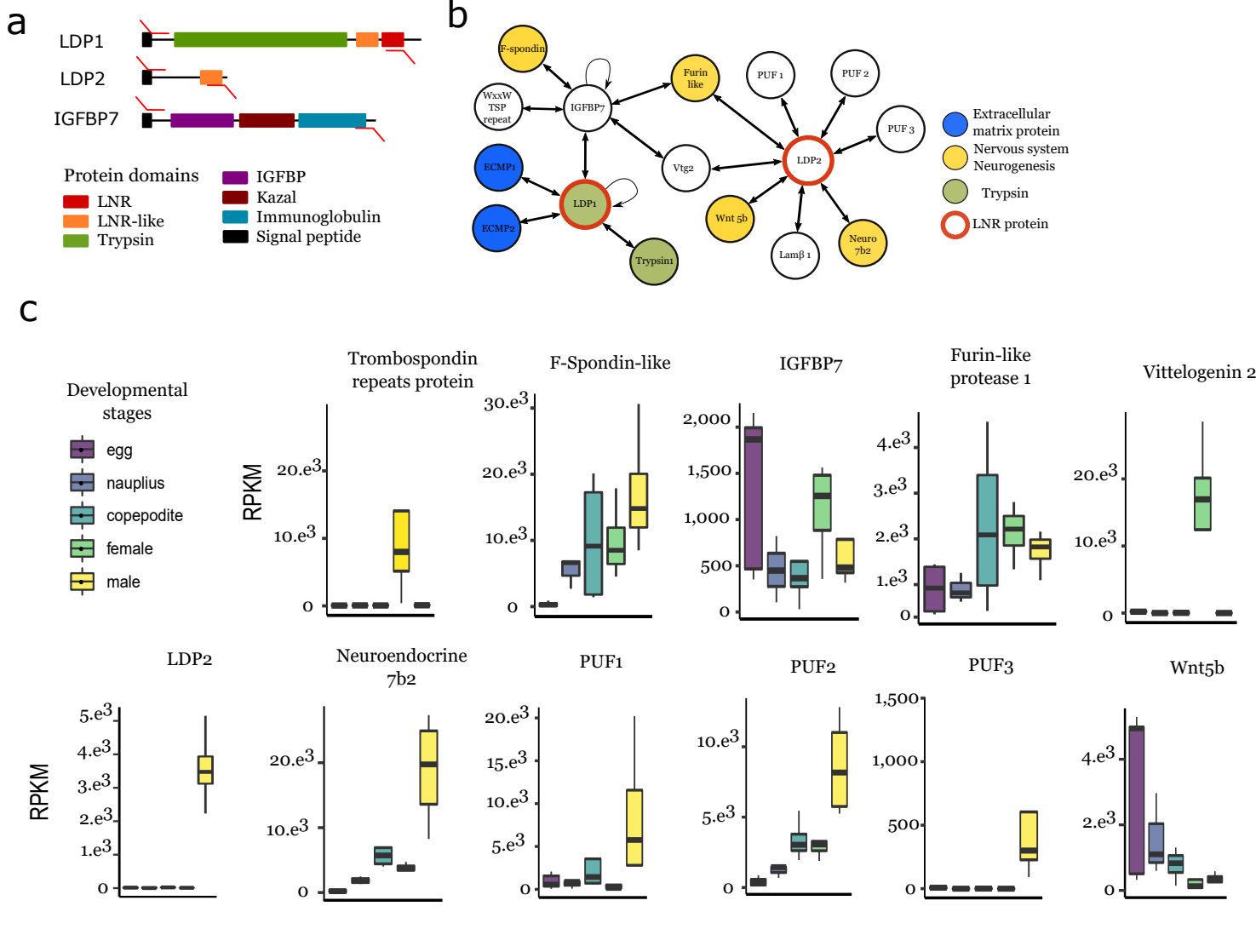

b
Mating
Male
Female
Female with
egg sacs
Sex
differentiation
Growth
Molts
Growth
Molts
Copepodite
(Juvenile)
Nauplius
(Larva)

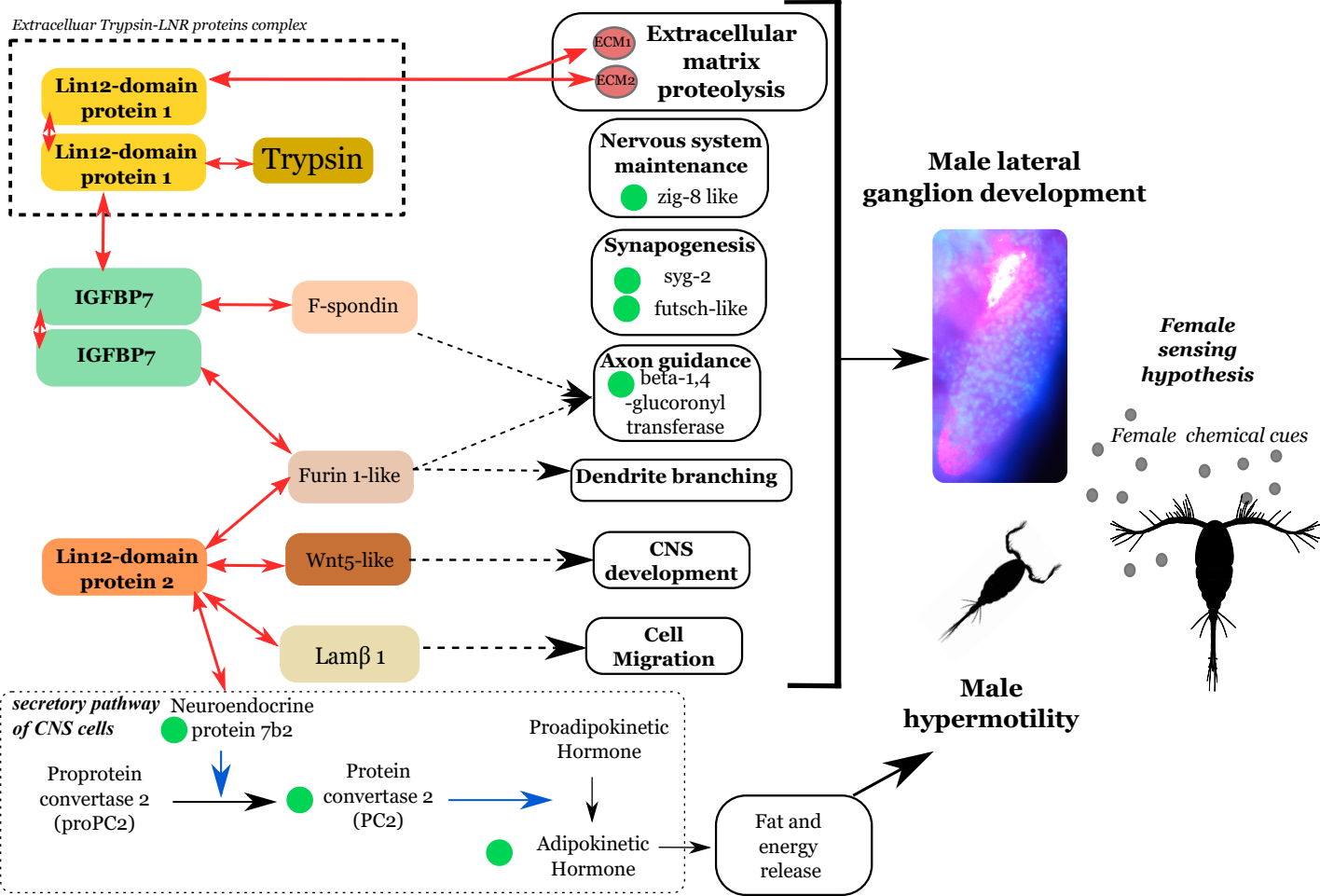

c
Male/Female Ratio
0.5
0.4
0.3
0.2
0.1
0
J
F
M
A
M
J
J
A
S
O
N
D
months


A
P
D
T
SEG
VNS
LG
a
PGF
b
50µm
DAPI
β-tubulin
B
P
D
T
SEG
LG
VNS
a
PGF
b
50µm
DAPI
β-tubulin
C
P
D
T
SEG
VNS
50µm
DAPI
β-tubulin
D
P
D
T
SEG
VNS
ANF
LG
50µm
DAPI
α-tubulin
E
CNS
LG
SEG
VNS
ANF
FNE
50µm
DAPI
α-tubulin
F
P
D
T
SEG
VNS
50µm
DAPI
α-tubulin


a
0.02
0.88
Podoplea
0.91
0.96
0.98
Gymnoplea
b
T. californicus
L. salmonis
O. nana
E. affinis
C. glacialis
D. pulex
D. melanogaster
4
3
82
2
6
2
3
0
100
200
300
LNR domain occurrences
c
Notch LNR (3)
LNR (22)
LNR-like (156)
d
LNR only
+ Trypsin
+ Metalloproteinase
+ Lectin
+ Kelch
+ Thrombospondin
+ Ankyrin
+ Kelch + PAN/Appel
+ Kelch +TSR
0-1 0-1 0-3 0-13
0-1 1-2 0-5 0-2
0-1 1 0-1 0-2
0-1 2-4 1
1 3
1 1 4
1 1
1 1 1
1 7 3
0
10
20
30
40
Number of genes
e
Notch protein
Protein domain types
LNR
LNR-like
Trypsin
Metallopeptidase
Kelch
Lectin
Thrombospondin (TSR)
Ankyrin-repeat
PAN/Apple
Signal peptide
Transmembrane


a
-1 0 1
Row Z-Score
Egg
Larva
Juvenile
Female adult
Male adult
b
48 Nervous system
26 Proteolysis
21 Signaling
10 Amino acid metabolism
6 Immune system
6 Extracellular matrix
4 Reproduction
3 Muscle
68 Other
253 Unknown
200 100 0
c
Protein domains
LNR
LNR-like
Trypsin
Metallopeptidase
Kelch
Lectin
Signal peptide
Transmembrane
Under selection
Egg
Larva
Juvenile
Female adult
Male adult
d
PGM
3PG 2PG PEP
Glycine
SDH
Serine Pyruvate
AlaDH
Alanine
Sarcosine dehydrogenase
Sarcosine
Acetyl-CoA
Citrate
Isocitrate
α-ketoglutarate
Methionine
Dimethylglycine
BHMT
Homocysteine
Trimethylglycine
AASS
Lysine
Glutamate
Glutaminase (3)
Glutamine
TyrAT
Tyrosine
GABA


a
LDP1
LDP2
IGFBP7
Protein domains
LNR
LNR-like
Trypsin
IGFBP
Kazal
Immunoglobulin
Signal peptide
b
F-spondin
WxxW TSP repeat
IGFBP7
Furin like
PUF 1
PUF 2
PUF 3
ECMP1
ECMP2
LDP1
Vtg2
LDP2
Wnt 5b
Lamβ 1
Neuro 7b2
Trypsin1
Extracellular matrix protein
Nervous system Neurogenesis
Trypsin
LNR protein
c
Developmental stages
egg
nauplius
copepodite
female
male
RPKM
Trombospondin repeats protein
F-Spondin-like
IGFBP7
Furin-like protease 1
Vittelogenin 2
LDP2
Neuroendocrine 7b2
PUF1
PUF2
PUF3
Wnt5b


Extracelluar Trypsin-LNR proteins complex
Lin12-domain protein 1
Lin12-domain protein 1
Trypsin
IGFBP7
IGFBP7
F-spondin
Furin 1-like
Lin12-domain protein 2
Wnt5-like
Lamβ 1
ECM1
ECM2
Extracellular matrix proteolysis
Nervous system maintenance
zig-8 like
Synapogenesis
syg-2
futsch-like
Axon guidance
beta-1,4 -glucoronyl transferase
Dendrite branching
CNS development
Cell Migration
Male lateral ganglion development
Female sensing hypothesis
Female chemical cues
Male hypermotility
secretory pathway of CNS cells
Neuroendocrine protein 7b2
Proprotein convertase 2 (proPC2)
Protein convertase 2 (PC2)
Proadipokinetic Hormone
Adipokinetic Hormone
Fat and energy release