Figures and figure supplements

Combined transcriptome and proteome profiling reveals specific molecular brain signatures for sex, maturation and circalunar clock phase

Sven Schenk et al
Figure 1. Strategy to capture the impact of maturation, sex and circalunar phase on the head transcriptome and proteome of Platynereis dumerilii. (a) Light regime of the Platynereis culture and its correlation with maturation; daily and nocturnal light conditions are schematised. Eight consecutive
nights with nocturnal light per month provide a simplified ‘full moon’ (FM) stimulus, alternating with dark nights (‘new moon’/NM) (intensity and spectra see Figure 1—figure supplement 2a–d). (b) Representative images of the different animal stages relevant for the study: IM - immature; PM - premature; M - mature. Top row: overview of heads at the respective stage; dotted line: cut site for head sampling. scale bars: 1 mm. Lower rows: representative microscopic images of coelomic cells used to diagnose each stage. Whereas immature animals display essentially only eleocytes (left), premature animals (middle) show eleocytes as well as germ cells (oocytes and spermatogonia, respectively). In mature animals (right), oocytes have reached their full size of ~180 μm in females, and male spermatogonia have differentiated into spermatozoa. Scale bars: 50 μm. (c) Scheme illustrating the two distinct circalunar phases used for sampling; NM – New Moon; FRFM – Free Running Full Moon (Zantke et al., 2014, Zantke et al., 2013). (d) Schematic flow chart of the newly developed protocol, allowing to profile the proteome and corresponding transcriptome of each sample.

DOI: https://doi.org/10.7554/eLife.41556.003
Figure 1—figure supplement 1. Schematic of representation of the strategy for the generation of the maturation, sex and lunar transcript/proteome regulated lists. Step 1 involves the generation of DESeq-EdgeR or ROTS-LIMMA rank sum lists for each of the 14 comparisons as outlined in the Materials and methods section. Step 2 is different for sex biased and circalunar transcripts/proteins on the one hand and for maturation on the other hand. For sex and lunar transcripts/proteins this step consists of subsequently merging the three individual comparisons - PM_NM_F vs. PM_NM_M, M_NM_F vs. M_NM_M and PM_FRFM_F vs. PM_FRFM_M for sex biased transcripts/proteins; and IM_NM vs. IM_FRFM, PM_NM_F vs. PM_FRFM_F and PM_NM_M vs. PM_FRFM_M for the lunar comparison - and keeping all entries in the final lists resulting in the final sex biased and circalunar transcript/proteome lists. For the maturation comparison step two is the generation of ‘unisex’ lists for the different developmental comparisons, for example out of the two IM_NM vs. PM_NM comparisons, IM_NM vs. PM_NM_F and IM_NM vs. PM_NM_M a single IM_NM vs. PM_NM list is build. To achieve this, the respective lists are merged by ID as above, but only IDs present in both initial lists are kept. This eliminates transcripts/proteins called DE in only one sex; however, entries regulated differently between sexes and called DE between two developmental stages are retained (see overlap between maturation and sex bias in Figures 3a and 5a, respectively). Step three only applies to the maturation comparison: In this step the four ‘unisex’ lists, IM_NM vs. PM_NM, IM_NM vs. M_NM, PM_NM vs. M_NM and IM_FRFM vs. PM_FRFM, are merged by ID and all entries are kept in the final list, as described for step two for the sex and lunar list. The resulting list contains all maturation transcripts/proteins.

DOI: https://doi.org/10.7554/eLife.41556.004
Figure 1—figure supplement 2. Ambient light conditions and overview over the multiplexing strategies employed for transcriptomic and proteomic analyses. (a–d) Representative illumination profiles of the worm culture rooms at different lunar phases. (a) Light spectrum of the ‘in phase culture room’.
(b) Light spectrum of the ‘outphase culture room’ during day. (c) Nocturnal light spectrum ‘inphase room’ under Full-moon conditions (FM). (d) Nocturnal light spectrum ‘outphase room’ under FM conditions. (e) Schematic overview of the multiplexing strategy applied for RNA-Seq; (f) respective overview of the multiplexing strategy for proteomics. In (e,f), red arrows denote NM samples run as technical replicates in the FRFM transcriptome sequencing and proteome measurements.

DOI: https://doi.org/10.7554/eLife.41556.005
Figure 2. Processing and basic features of the Platynereis head transcriptome. The primary assembly consisting of 407,172 transcripts with >200 bp was successively processed by: i) collapsing multi-copy transcripts to the longest transcripts of each cluster, ii) removing transcripts shorter than 500 bp, iii) removing transcripts arising from common model organisms or potential food sources, iv) substituting less complete transcripts with previously published sequences from NCBI and the transcriptome published by Conzelmann et al. (2013), and v) adding missing sequences from NCBI. In total 35,176 Pdu_HeadRef_TS_v5 sequences have BLASTx hits against the NCBI nr- database.

DOI: https://doi.org/10.7554/eLife.41556.008
Figure 3. Maturation, sex and circalunar phase have distinct transcriptomic signatures in Platynereis heads. (a) Venn diagram showing the overall numbers of genes identified as significantly regulated within and across the three different biological processes. Percent in brackets indicate the percentage of DETs for each process. (b) Heatmaps showing the expression patterns of genes in the different clusters for each process. (c) Bar charts showing the percentage of DETs for different functional categories.
Figure 3 continued

respective fraction compared to total number of analyzed IDs. (b) Soft clustering of differentially expressed genes using the Mfuzz algorithm. Maturation (yellow) and sexual differences (light red) yielded five clusters; circalunar phase-regulated (light blue) expression was best represented by four clusters. Compare to Figure 5b for similarity to regulated protein clusters. (c) Results from the GO-term enrichment analysis using the GOStats package on the differentially expressed genes in each comparison; the ten most abundant terms in the Molecular Function category are displayed (for more GO-results see Figure 3—figure supplement 1—9). Green boxes: actual percentage of differentially expressed transcripts for each term; grey bars: expected number of transcripts per category. Maturation: protein binding: GO:0005515, catalytic activity, acting on a protein: GO:0140096, inorganic molecular entity transmembrane transporter activity: GO:0015318, calcium ion binding: GO:0005509, kinase activity: GO:0016301, substrate-specific channel activity: GO:0022838, passive transmembrane transporter activity: GO:0022803, oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen: GO:0016705, heme binding: GO:0020037, metal ion transmembrane transporter activity: GO:0046873. Sex bias: catalytic activity: GO:0003824, transporter activity: GO:0005215, peptidase activity: GO:0008233, calcium ion binding: GO:0005509, inorganic molecular entity transmembrane transporter activity: GO:0015318, endopeptidase activity: GO:0004175, cation transmembrane transporter activity: GO:0008324, ion gated channel activity: GO:0022839, cysteine-type peptidase activity: GO:0008234. Circalunar phase: phosphorus-oxygen lyase activity: GO:0016849, protein kinase activity: GO:0004672, insulin receptor binding: GO:0005158, transposase activity: GO:0004803. Statistical significance was tested with a hypergeometric G-test implemented in the GOStats package. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, *****p < 0.00001. DOI: https://doi.org/10.7554/eLife.41556.013
Figure 3—figure supplement 1. Analyses of over-represented GO-terms support the existence of distinct molecular signatures for maturation, sex and circalunar clock phase. (a–c) top 10 enriched terms in the Biological Process category; blue boxes: percentage of differentially expressed transcripts for

Figure 3—figure supplement 1 continued on next page

DOI: https://doi.org/10.7554/eLife.41556.014
Analyses of under-represented GO-terms support the existence of distinct molecular signatures for maturation, sex and circalunar clock phase. (a’’–b’’’) corresponding analyses focusing on top under-represented GO-terms; (a’’’,b’’’) top under-represented GO-terms in Figure 3—figure supplement 2 continued on next page.

DOI: https://doi.org/10.7554/eLife.41556.015
Figure 3—figure supplement 3. Heat map displaying under-represented GO-terms in all categories. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase (left), sex (middle) and/...
Figure 3—figure supplement 3 continued

or maturation (right), colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.016
Figure 3—figure supplement 4. Heat map displaying under-represented GO-terms in the Biological Process category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase (left), continued on next page.
Figure 3—figure supplement 4 continued

sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.017
Figure 3—figure supplement 5. Heat map displaying under—represented GO-terms in the Cellular Compartment category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar
Figure 3—figure supplement 5 continued

clock phase (left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.018
Figure 3—figure supplement 6. Heat map displaying under-represented GO-terms in the Molecular Function category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase (left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.019
Figure 3—figure supplement 7. Heat map displaying over-represented GO-terms in all categories. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase (left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.020
Figure 3—figure supplement 8. Heat map displaying over-represented GO-terms in the Biological Process category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase.

Figure 3—figure supplement 8 continued on next page
Figure 3—figure supplement 8 continued
(left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a
hypergeometric test implemented in the GOStats package.
DOI: https://doi.org/10.7554/eLife.41556.021
Figure 3—figure supplement 9. Heat map displaying over-represented GO-terms in the Cellular Compartment category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase (left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.
DOI: https://doi.org/10.7554/eLife.41556.022
Figure 3—figure supplement 10. Heat map displaying over-represented GO-terms in the Molecular Function category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase (left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.023
Figure 4. Exemplary validation of candidate genes by qRT-PCR and expression domains in the head. (a’-a’’, b’-b’’) Maturation markers fatty acid binding protein and qpeptin. (c’-c’’, d’-d’’) Sexual differentiation markers dmrt and sten. (a-d): qRT-PCR validation of the candidates relative to the arithmetic mean of the reference genes cd5 and sams. All graphs show the arithmetic mean with standard deviation, and the individual data points. (a’-d’’) Whole mount in situ hybridisations of the respective genes, comparing premature animals (PM, (a’,b’)) with mature animals (M, a’’,b’’), and males (c’,d’,d’’) with females (c’’,d’’, d’’’), respectively. All images except d’’’ and d’’’’: dorsal views, anterior to the top, scale bar: 250 μm. d’’ and d’’’: ventral views, anterior top, scale bar: 200 μm. Also see Figure 4—figure supplement 1 for further validations. a.e.: anterior eye; p.e.: posterior eye, an: antenna, pa: palp.
Figure 4—figure supplement 1. Exemplary validation of additional transcripts differentially expressed during maturation, between sexes or between circalunar clock phases. (a,b) yellow- identified as maturation regulated; (c) light red- identified as sex-biased regulated; (d,e) light blue- identified as...
circalunar phase regulated. Note that for circalunar phase regulated transcripts verification by qPCR needs a higher cDNA amount input than for the tested maturation and sex-specific regulated transcripts. The graphs show the arithmetic mean with standard deviation, together with individual data points. Further details see Figure 4 (legend) and Materials and methods.
DOI: https://doi.org/10.7554/eLife.41556.040
**Figure 5.** The *Platynereis* head proteome is differentially affected by maturation, sex and circalunar phase. (a) Venn diagram showing the proteins significantly regulated within and across the three different biological processes in the three major comparisons made (brackets indicate respective
fraction of total comparable proteins. (b) Soft clustering of differentially expressed proteins using the Mfuzz algorithm. The expression data for maturation and circalunar phase are best represented by five and for sexually dimorphic protein expression by three clusters. Compare to Figure 3b for similar dynamics in the regulated transcript clusters. (c) GO-term enrichment analysis using the GOStats package on the differentially expressed proteins in each comparison showing the ten most abundant terms in the Molecular Function category (further GO-term analyses: see Figure 5—figure supplement 2–6). Green boxes show the percentage of differentially expressed proteins for each term, the grey bars depict the expected amount of proteins per category. Maturation: binding: GO:0005488, calcium ion binding: GO:0005509, structural molecule activity: GO:0005198, isomerase activity: GO:0016853, cysteine-type peptidase activity: GO:0008234, actin binding: GO:0003779, proton transmembrane transporter activity: GO:0015078, unfolded protein binding: GO:0051082, motor activity, GO:0003774, protein heterodimerization activity, GO:0046982. Sex bias: oxygen binding: GO:0019825, heme binding: GO:0020037, ferric iron binding: GO:0008199, oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced pteridine as one donor, and incorporation of one atom of oxygen (oxidoreductase activity, acting on paired...): GO:0016714, lipid transporter activity: GO:0005319, carbonate dehydratase activity: GO:0004089. Circalunar phase: nucleic acid binding: GO:0003676, calcium ion binding: GO:0005509, structural constituent of ribosome: GO:0003735, proton transmembrane transporter activity: GO:0015078, protein heterodimerization activity: GO:0046982, peptidyl-prolyl cis-trans isomerase activity: GO:0003755, oxygen binding: GO:0019825, phosphotransferase activity, phosphate group as acceptor: GO:0016776, nucleobase-containing compound kinase activity: GO:0019205, protein dimerization activity: GO:0046983. Statistical significance was tested with a hypergeometric test implemented in the GOStats package. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.
Figure 5—figure supplement 1. Overall correlation of protein and transcript abundances detected in the same biological samples. The log10 transformed normalised mean protein abundances of all 2,290 detected proteins (x axis) were correlated with the log10 transformed DESeq-normalised RNA transcript abundances (y axis). Coloured dots correspond to proteins significantly regulated in the indicated analyses (maturation, sex, circalunar). DOI: https://doi.org/10.7554/eLife.41556.043
Analyses of over-represented GO-terms support the existence of distinct molecular signatures for maturation, sex and circalunar clock phase. GO-term enrichment analyses confirm the distinct regulation of the head proteome during the different tested biological phases.

Figure 5—figure supplement 2 continued on next page

DOI: https://doi.org/10.7554/eLife.41556.044
Figure 5—figure supplement 3. Heat map displaying over-represented GO-terms in all categories. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase (left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric G test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.045
Figure 5—figure supplement 4. Heat map displaying over-represented GO-terms in the Biological Process Category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase.
Figure 5—figure supplement 4 continued

(left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a
hypergeometric test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.046
Figure 5—figure supplement 5. Heat map displaying over-represented GO-terms in the Cellular Compartment Category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase.
Figure 5—figure supplement 5 continued
(left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.
DOI: https://doi.org/10.7554/eLife.41556.047
Figure 5—figure supplement 6. Heat map displaying over-represented GO-terms in the Cellular Compartment Category. The heat map is subdivided (columns) by the regulation of each GO-term by circalunar clock phase (left), sex (middle) and/or maturation (right); colours are indicative of the enrichment p-value as calculated by a hypergeometric test implemented in the GOStats package.

DOI: https://doi.org/10.7554/eLife.41556.048
Figure 5—figure supplement 7. The normalisation procedure applied to the raw protein abundance values leads to evenly spread and centred abundance profiles. (a) Raw intensity values for each of the proteome measurements, excluding missing data. (b) Corresponding abundance values after normalisation.

Schenk et al. eLife 2019;8:e41556. DOI: https://doi.org/10.7554/eLife.41556

Figure 5—figure supplement 7 continued on next page
normalisation. Annotations on the x axis in (b) also correspond to the respective columns in (a). Yellow: NM samples BR1, orange: NM samples BR2, red: NM samples BR3, dark blue: FRFM samples BR1, blue: FRFM samples BR2, light blue: FRFM samples BR3. Suffices TMT_XYZ denote the label, used for labelling the respective sample, that is TMT126 is labelled with a tandem mass tag with the reporter mass of 126 Da.

DOI: https://doi.org/10.7554/eLife.41556.049
Figure 6. Members of a lineage-specific expansion of Ependymin-related proteins are subject to circalunar regulation on RNA and protein level. (a) Unrooted maximum likelihood phylogenetic tree of Ependymin-related proteins based on the analysis published by Suárez-Castillo and García-Schenk et al. eLife 2019;8:e41556. DOI: https://doi.org/10.7554/eLife.41556

Tools and resources
Ecology | Neuroscience
Figure 6 continued

**Arrarás, 2007** including eleven of 15 **Platynereis** Ependymin-related proteins (ERP4.1–4.10; ERP3; blue boxes) and various additional invertebrate Ependymin sequences. The phylogeny reveals four distinct groups of Ependymin-related proteins (ERP) that are termed ERP1, ERP2, ERP3/4 and fish-specific ERPs; group-specific expansions of ERPs is observed at several places in the phylogeny. All identified **Platynereis** Ependymins (blue) fall into the ERP3/4 clade. A full tree with un-collapsed nodes, the alignment, NCBI accession numbers and the **Platynereis** sequence identifiers are provided in **Figure 6—figure supplement 1** and **Figure 6—source datas 1 and 2**. (b,b') Protein and RNA-Seq expression profiles validate **Platynereis** ERP4.9 as a new target of circalunar phase on RNA and protein level. (b) Normalised protein expression profile; (b') DESeq2 normalised RNA expression profile. The graphs show the arithmetic mean with standard deviation, together with individual data points. (c–e) RNA expression pattern of erp4.9 in immature adult **Platynereis** heads. a.e.: anterior eye; p.e.: posterior eye; n.o.: nuchal organ. (c) Dorsal view, anterior to the top. Characteristic expression is observed around the dorsal blood vessel (central arrow), anterior of the nuchal organ (arrowheads) and next to the anterior eyes (arrows). Scale bar: 250 μm. (d) Dorsal view of the base of a prostomial cirrus, scale bar: 50 μm. (e) Ventral view, anterior to the top. Expression is observed in a plate-like structure at the bottom of the head, as well as around the mouth opening (arrows) and the palps (arrow heads). Scale bar: 250 μm. For additional circalunar-regulated ERPs see **Figure 6—figure supplement 2** and **Figure 5—source datas 16 and 17**.

DOI: https://doi.org/10.7554/eLife.41556.071
Figure 6—figure supplement 1. Maximum likelihood phylogeny of Ependymin-related proteins using the same alignment and settings as in Figure 6a. For better visibility, the tree is shown as rooted at the split between ERP group one and ERP group 2. Platyneuris sequences are labelled
Figure 6—figure supplement 1 continued

with a systematic nomenclature following the apparent phylogeny, all other entries are labelled with genus_and Accession-Number. For cases in which translated (tr) genomic regions were initially used, two accession numbers are given.

DOI: https://doi.org/10.7554/eLife.41556.072
Figure 6—figure supplement 2. ERP4.10 is regulated in both protein and RNA expression by circalunar phase. (a,b) Normalised values for abundance of ERP4.10 on protein (a) and RNASeq (b) level. (c–c’’) Results of a whole mount in-situ hybridisation on immature Platynereis heads. a.e.: anterior adult eye; p.e.: posterior adult eye; n.o.: nuchal organ. (c) Dorsal view, anterior to the top. Characteristic expression is observed around the dorsal blood vessel (central arrow), anterior of the nuchal organ (arrow heads) and next to the anterior adult eyes (arrows). Scale bar: 250 μm. (c’) Ventral view, anterior to the top. Expression is observed in a plate-like structure at the bottom of the head. Scale bar: 250 μm. (c’’) Dorsal view of cirrus; Scale bar: 50 μm.

DOI: https://doi.org/10.7554/eLife.41556.073
Figure 7. Proteome analyses reveal additional molecular markers for circalunar phase with distinct expression profiles and proposed functions. (a–a‴) The neuropeptide Whitnin, a′- peristomeal cirrus, a′′- head, a′‴- indicated head expression in a′ with different focus. (b–b″) The oxygen storage protein Haemerythrin, b″- head, b′- lateral head aspects below the focal plane of b″. (c–c″) The iron storage protein Ferritin. c′- peristomeal cirrus. (a–c) compares the normalised protein expression profiles (top, ‘proteomics’) with the corresponding DESeq2 normalised mRNA expression profile (bottom, ‘RNA-Seq’). All graphs show the arithmetic mean with standard deviation, and the individual data points. (a′–c″) Expression domains as characterised by whole-mount in situ hybridisation; (a′″, b′″, c′″) dorsal views of immature Platyneris heads stained with riboprobes of the corresponding transcript; All images are oriented with the anterior side to the top; scale bar: 250 μm. a.e.: anterior eye, p.e.: posterior eye.

DOI: https://doi.org/10.7554/eLife.41556.076
Figure 7—figure supplement 1. Platynereis Whitnin is part of a Proctolin/Whitnin family of neuropeptides. Maximum likelihood phylogeny of Proctolin/Whitnin family, see Figure 7—source datas 1 and 2 for alignment and Accession numbers.

DOI: https://doi.org/10.7554/eLife.41556.077