

Supporting information

A drug discovery approach based on comparative transcriptomics between two potentially toxin-secreting marine annelids: *Glycera alba* and *Hediste diversicolor*

Inês Moutinho Cabral, Carolina Madeira, Ana R. Grosso, Pedro M. Costa

Associate Laboratory i4HB – Institute for Health and Bioeconomy, NOVA School of Science and Technology, NOVA University of Lisbon, 2829-516 Caparica, Portugal

UCIBIO – Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, NOVA University of Lisbon, 2829-516 Caparica, Portugal

Table of Contents

Table S.1. Sampling.	S3
Table S.2. The primers sequences used for PCR and RT-qPCR.	S4
Figure S.1. Relative expression of selected genes in <i>Glycera alba</i> and <i>Hediste diversicolor</i>	S5
Figure S.2. Volcano plot of differentially-expressed genes.	S7
Figure S.3. Heatmaps illustrating differential gene expression profiling.	S8
Table S.3. Top10 overexpressed genes in <i>Glycera alba</i> 's proboscis relative to the skin.	S9
Figure S.4. Aminoacid sequence of a cysteine-rich venom protein isolated from <i>Glycera alba</i>	S11
Figure S.5. Aminoacid sequence of an ovoinhibitor-like protein isolated from <i>Glycera alba</i>	S12
Figure S.6. Aminoacid sequence of a pathogenesis-related protein (a CRISP) isolated from <i>Hediste diversicolor</i>	S14
Figure S.7. Aminoacid sequence of thyrostimulin beta-5 subunit isolated from <i>Hediste diversicolor</i>	S15
Figure S.8. Protein-protein interactions between the potential interactors from the proboscis and skin of <i>Glycera alba</i> and the human proteome retrieved from the HuRI platform.	S16
Figure S.9. Protein-protein interactions between the potential interactors from the glands and proboscis of <i>Hediste diversicolor</i> and the human proteome retrieved from the HuRI platform.	S18
Table S.4. Enriched biological processes affected by the potential interactors from the proboscis of <i>Glycera alba</i> and their human targets.	S20
Table S.5. Enriched biological processes affected by the potential interactors from the skin of <i>Glycera alba</i> and their human targets.	S23
Table S.6. Enriched biological processes affected by the potential interactors from the glands of <i>Hediste diversicolor</i> and their human targets.	S24
Table S.7. Enriched biological processes affected by potential interactors from the proboscis of <i>Hediste diversicolor</i> and their human targets.	S27

Table S.1. Sampling.

	Sample Name	Species	Organ	Number of Individuals	Number of Raw Reads
1	GaPR1	<i>Glycera alba</i>	Proboscis	1	39 749 500
2	GaPR2	<i>Glycera alba</i>	Proboscis	2 (pool)	43 177 154
3	GaPR3l	<i>Glycera alba</i>	Proboscis	3 (pool)	45 489 590
4	GaBW1	<i>Glycera alba</i>	Skin	1	43 073 170
5	GaBW2	<i>Glycera alba</i>	Skin	3 (pool)	46 741 642
6	GaBW3	<i>Glycera alba</i>	Skin	3 (pool)	38 915 178
7	Hd1G	<i>Hediste diversicolor</i>	Glands	15 (pool)	31 099 868
8	Hd2G	<i>Hediste diversicolor</i>	Glands	10 (pool)	123 202 926
9	Hd3G	<i>Hediste diversicolor</i>	Glands	10 (pool)	49 858 858
10	Hd1PR	<i>Hediste diversicolor</i>	Proboscis	15 (pool)	42 056 540
11	Hd2PR	<i>Hediste diversicolor</i>	Proboscis	10 (pool)	94 650 066
12	Hd3PR	<i>Hediste diversicolor</i>	Proboscis	10 (pool)	42 195 514

Table S.2. The primers sequences used for PCR and RT-qPCR. The genes of interest amplified encoded for an ovoinhibitor (Ov), a cysteine-rich venom protein (TX31), a pathogenesis-related protein (a CRISP) (Pat) and thyrostimulin beta-5 subunit (Thy), whereas housekeeping calibrators were 18S and beta-actin (BAct). The primer forward of the Thyrostimulin beta-5 subunit was used for PCR and RT-qPCR.

Target	Primers	PCR or RT-qPCR	Primer Sequences	Amplicon Size
18S	Foward	PCR/RT-qPCR	CGATGGTACGTGATATGCC	176
	Reverse	PCR/RT-qPCR	CGAATGAGTCCCGTATTGT	
BAct	Foward	PCR/RT-qPCR	CGGTATCGTGCTGGATTC	163
	Reverse	PCR/RT-qPCR	CGTGGTGGTGAAGCTGTA	
Ov	Foward_1	PCR	GCTTACATCTTATCATGCTC	639
	Reverse_1	PCR	CTGTATTGCACTCAGGTTC	
	Foward_2	PCR	CTACAAGTCGTGTCATGC	789
	Reverse_2	PCR	CTGCTTCTATTGGTTGGC	
TX31	Foward	PCR	GAATCCTGAACCTGTCTGTG	954
	Reverse	PCR	CAACCTGTTCTTAACATACTCC	
Pat	Foward	PCR	GTGGTCAAGATCAGAACTGC	667
	Reverse	PCR	GGAAACCATTTACAGGAGAGG	
Thy	Foward	PCR/RT-qPCR	GGACAGGCATGAAGGACA	465
	Reverse	PCR	CTCACAACGATTCGCAACT	
Ov	Foward	RT-qPCR	GTGTACCTACGAATACAACC	150
	Reverse	RT-qPCR	CCATCAGATCCGCAAAC	
TX31	Foward	RT-qPCR	AGGCTTCTTGAATGATGCT	172
	Reverse	RT-qPCR	TGTGTTGGACAGTGGTT	
Pat	Foward	RT-qPCR	ACCTCATCATCCTCTTTGC	140
	Reverse	RT-qPCR	TCTGCTGCTCAGTCTTGC	
Thy	Reverse	RT-qPCR	AGGATACAGACGGCAGTG	167

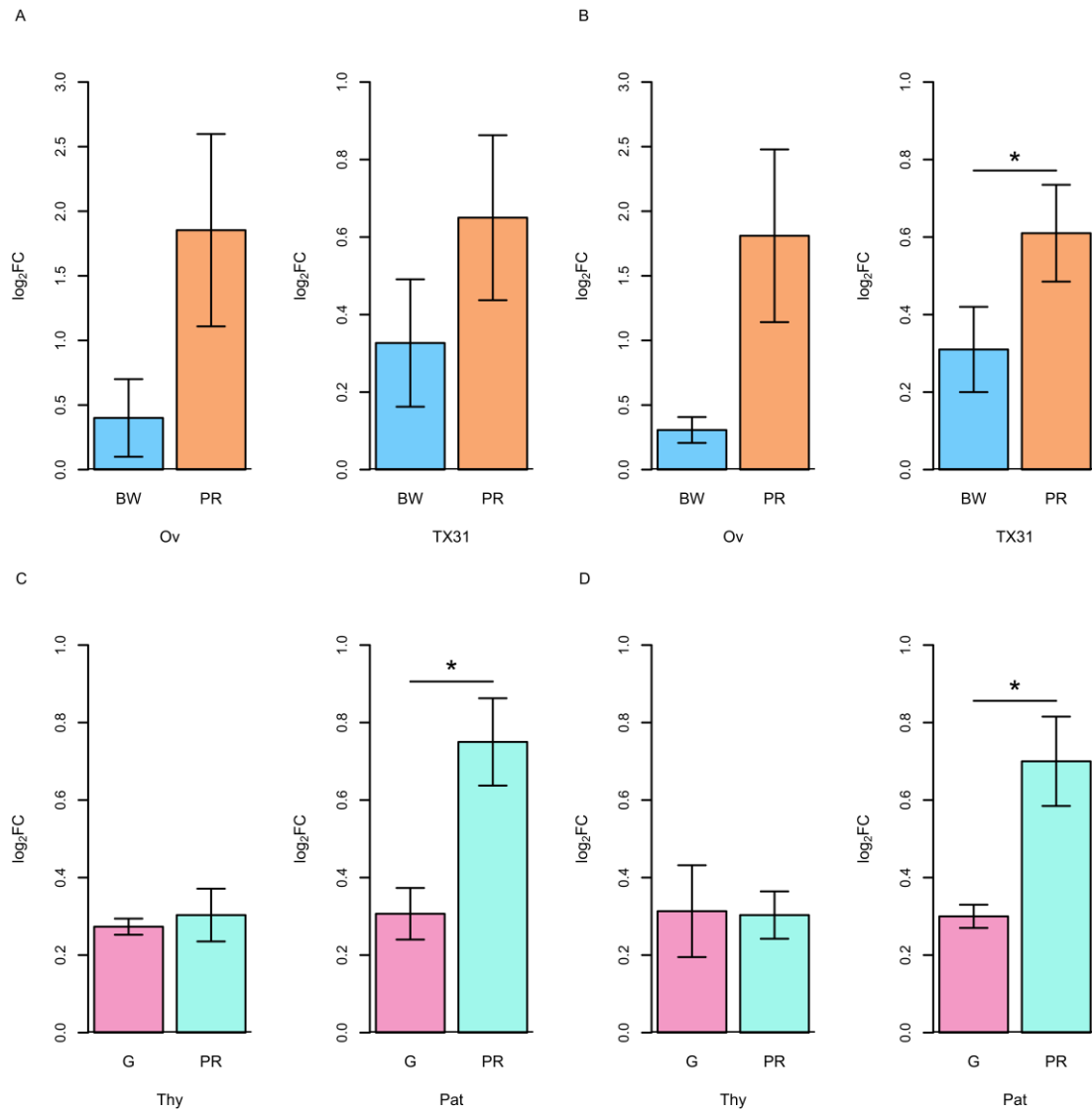


Figure S.1. Relative expression of selected genes in *Glycera alba* and *Hediste diversicolor*. The expressed sequence tags (EST) detected by RT-qPCR were in mRNAs that encoded for an ovoinhibitor-like protein (Ov), a cysteine-rich venom protein (TX31), a pathogenesis-related protein (a CRISP) (Pat) and thyrostimulin beta-5 subunit (Thy). Relative levels of expression of Ov and TX31 were compared between *Glycera alba*'s proboscis (PR) and skin (BW) (A and B), while, for Thy and Pat, the comparison was between *Hediste diversicolor*'s glands (G) and proboscis (PR) (C and D). In A) and C), the levels of expression were normalized with the house-keeping gene 18S, while in B) and D) were with beta-actin. *indicates significant differences between organs (Student's

t-test, $p < 0.05$). Although the relative expression of gene that encoded the ovoinhibitor-like protein was not found to be significantly different, likely due to relatively high error and low n , the trend is similar to RNA-Seq-based quantification, i.e., a higher expression in the proboscis relative to the skin. The relative expression of the gene that encodes thyrostimulin beta-5 subunit was not either found to be significantly different between *H. diversicolor*'s organs. This might be due to the $\log_2FC$ value, which was 2.33.

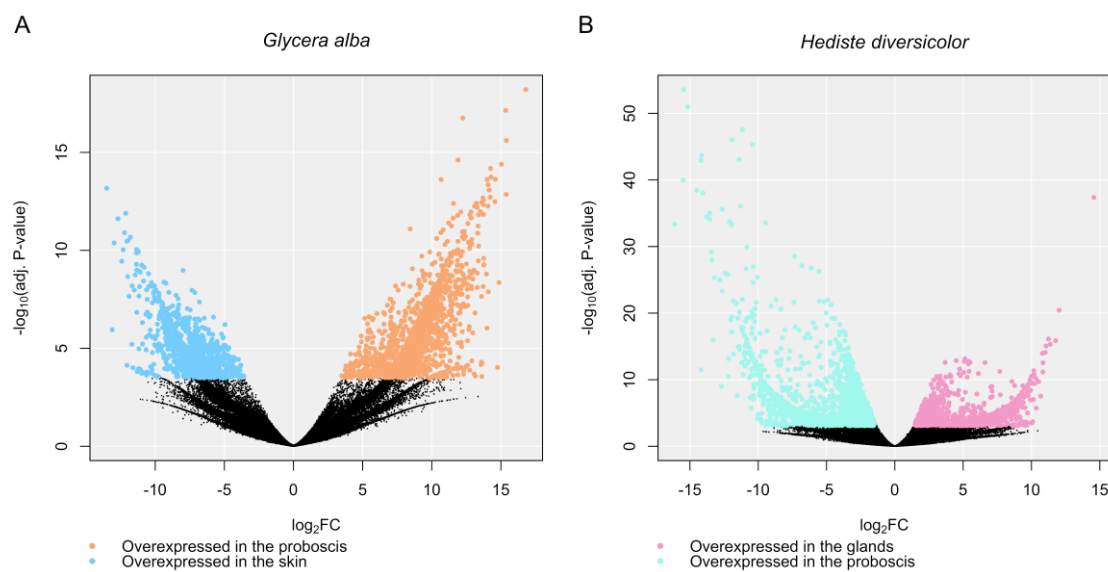


Figure S.2. Volcano plot of differentially-expressed genes. A) between *Glycera alba*'s proboscis and skin. B) between *Hediste diversicolor*'s glands and proboscis. Black dots represent genes that are not differentially-expressed between organs. The cut-off for differential expression was set at $|\log_2\text{FC}| > 1.5$ and FDR-adjusted $p < 0.05$.

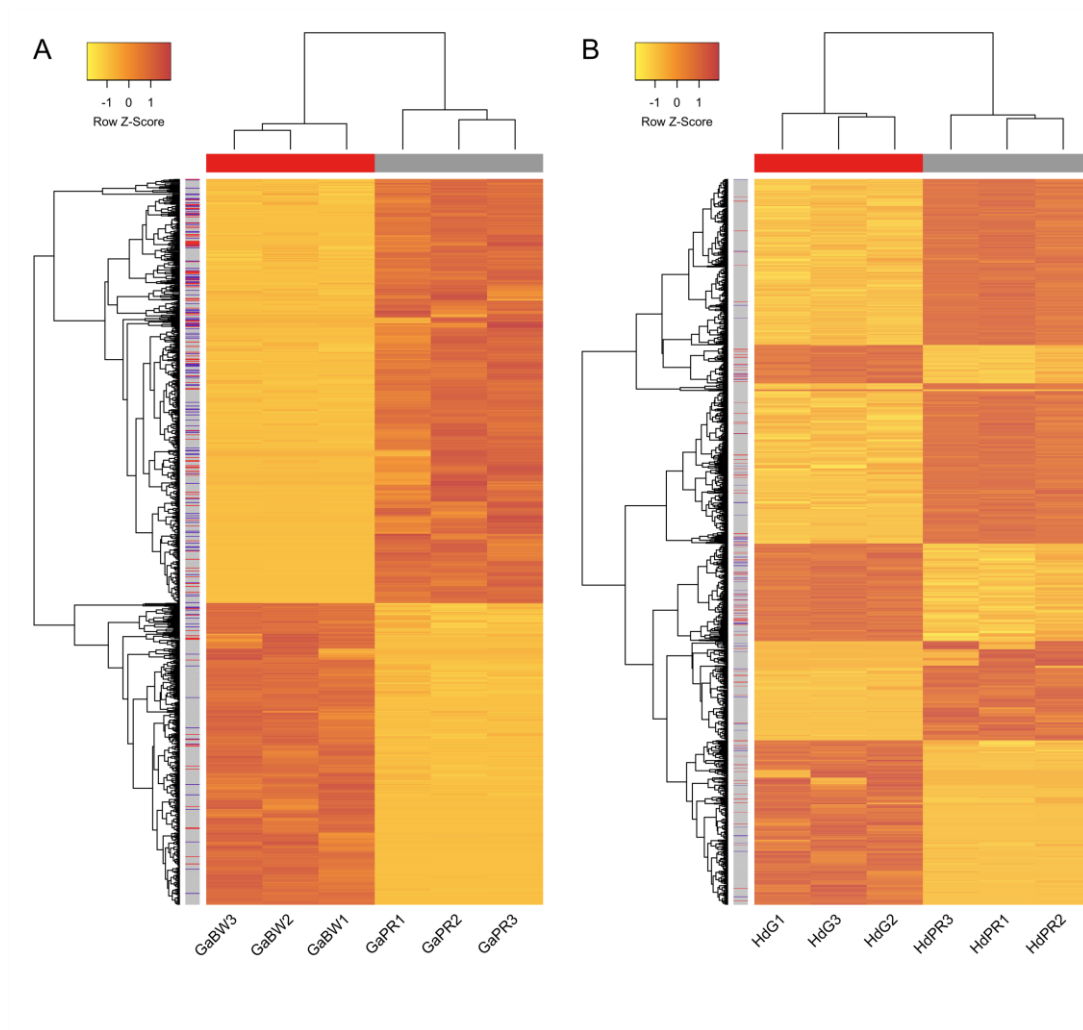


Figure S.3. Heatmaps illustrating differential gene expression profiling. A) between *Glycera alba*'s proboscis (PR) and skin (BW). B) between *Hediste diversicolor*'s glands (G) and proboscis (PR). A $|\log_2FC| > 1.5$ and an FDR adjusted $p < 0.05$ were the cut-offs set for differential expression. The horizontal dendrogram illustrates the association between the three independent replicates, whereas the vertical dendrogram represents the association between the expression profiling of different genes. The metric and function of the cluster analysis are Euclidian distances and complete linkage, respectively. Side bars indicate coding regions with homology-matching with potential toxins (blue) and permeabilizing or diffusing agents (red).

Table S.3. Top10 overexpressed genes in *Glycera alba*'s proboscis relative to the skin. The cut-offs established were $\log_2FC > 1.5$ and an FDR-adjusted $p < 0.05$.

log₂FC	log₂CPM	FDR_p	Protein	Accession	%ID	e-value	Organism
14.595	6.903	2.37E-14	Cysteine-rich venom protein (CRVP) (Substrate-specific endoprotease Tex31)	Q7YT83	35.567	1.59E-22	<i>Conus textile</i>
14.566	6.875	3.27E-13	Pancreatic triacylglycerol lipase (Fragment)	Q64425	34.921	1.75E-47	<i>Myocastor coypus</i>
14.093	6.402	1.31E-08	Plasma kallikrein (Plasma prekallikrein) (PKK) [Cleaved into: Plasma kallikrein heavy chain; Plasma kallikrein light chain]	P03952	39.382	3.43E-45	<i>Homo sapiens</i>
14.061	6.369	5.51E-13	Tenascin-R (Neural recognition molecule J1-160/180)	Q8BYI9	35.859	8.64E-26	<i>Mus musculus</i>
13.950	6.259	1.62E-09	Cysteine-rich venom protein (CRVP) (Substrate-specific endoprotease Tex31)	Q7YT83	37.008	6.51E-39	<i>Conus textile</i>
13.759	6.068	5.63E-10	Tenascin-R (Restrictin)	Q05546	36.869	2.78E-26	<i>Rattus norvegicus</i>
13.756	6.065	2.58E-10	Pro-low-density lipoprotein receptor-related protein 1 [Cleaved into: Low-density lipoprotein receptor-related protein 1 85 kDa subunit (LRP-85); Low-density lipoprotein receptor-related protein 1 515 kDa subunit (LRP-515); Low-density lipoprotein receptor-related protein 1 intracellular domain (LRPICD)]	G3V928	34.848	2.02E-46	<i>Rattus norvegicus</i>
13.717	6.027	7.02E-09	Glycerotoxin (Fragment)	A0A1U9VX95	31.017	2.37E-172	<i>Glycera tridactyla</i>
13.613	5.923	2.73E-04	Ovoinhibitor (Serine protease inhibitor Kazal-type 5) (allergen Gal d OIH)	P10184	29.216	1.34E-30	<i>Gallus gallus</i>

log₂FC	log₂CPM	FDR_p	Protein	Accession	%ID	<i>e</i>-value	Organism
13.595	5.905	2.24E-08	A disintegrin and metalloproteinase with thrombospondin motifs 7 (ADAM-TS 7)	Q1EHB3	30.000	1.50E-19	<i>Rattus norvegicus</i>

Predicted	1	MLP-----ILMLSVALIGFLNDAVMGRQIHPPFHRQFHHKRVSSGCTVSGVAYPSGHTM	53
OL606744	1	MLP-----ILMLSVALIGFLNDAVMGRQIHPPFHRQFHHKRVSSGCTVSGVAYPSGHTM	53
Q7YT83	1	MLSTMQTVGAVLMLSIVLVA-----GRKRHHCDISKYYE-----LTPAHTM	40
Predicted	54	CLSPASNMTPKPLSNTDQNAVVDKHNSYRSDVSPKASDMMKMYWDDSIAAVAQAWAETCP	113
OL606744	54	CLSPASNMTXPPLSNTDQNAVVDKHNSYRSDVSPKASNMMKMYWDDSIAAVAQAWAETCP	113
Q7YT83	41	CLTDKPNNAVAVPLTQETEHEILEMHNKIRADVT-DAANMLKMEWDERLATVAQKWAMQC-	98
Predicted	114	TTSFPHDTD-RSVPAYGISIGQNGAF--GQTDYTAAVASWHGEVTDFTYGAANALEDVGH	170
OL606744	114	TTSFPHDTD-RSVPAYGISIGQNGAF--GQTDYTAAVASWHGEVTDFTYGAANALEDVGH	170
Q7YT83	99	--ILGHDSGRRGEPDLPGSVGQNVAVSSGDLTFLGAVQMWADEIVDFQYGVWT--DGTGH	154
Predicted	171	YTQVVAGPAVAIGCGAANCPDNSYPDVFFFCNYAYGQSDFDNPNVSDINGCTNTCGSNVCN	230
OL606744	171	YTQVVAGPAVAIGCGAANCPDNSYPDVFFFCNYAYGQSDFDTPYVSDTXGCTNTCGSNVCN	230
Q7YT83	155	YIQQVFAGASRIGCGQSACGNKY---FVCNYYKGTMG-DEPY--QLGRPCSQCRSSCQH	208
Predicted	231	-----NLCDC--GNKICRNGGTLNLSCKCDCLSIYSGDTCQTTNCPDE--ETLCW	277
OL606744	231	-----NLCDC--GNKICLNGGTLNLSCKCDCLSVYSGDTCQTTNCPDE--ETLCW	277
Q7YT83	209	IRGSQGRWGS LCDCTNGPDACFNNGIFNINTCQCECSGIWGGADCQEKHCPNEDFDDMC-	267
Predicted	278	RWGSS-----PMCSWGPTVVDCPYTCGVC-----	301
OL606744	278	RWGSS-----PMCSWGPTVVDCPYTCGVC-----	301
Q7YT83	268	RYPDALRRPQHWCQYDNFQSDCPILCGYCPNPN	300

Figure S.4. Aminoacid sequence of a cysteine-rich venom protein isolated from *Glycera alba*.

The validated sequence (GenBank accession OL606744) was aligned with the original translated transcript as reconstructed by Trinity (predicted) along with the best hit produced by homology-matching against Swiss-Prot (cysteine-rich venom protein from *Conus textile*, Q7YT83).

Predicted	1	-----MLPT---FRLAILVVVG VASSYDECLRAC PFI-----YAPVCD AEG	39
OL606745	1	-----MLPT---FRLAILVVVG VASSYDECLRAC PFI-----YAPVCD AEG	39
P10184	1	MTDWVLHHKVGPLDMTTRYIFLLPLPFLPHSESKRAVCAPRCSAMRTARQFVQVALALC	60
Predicted	40	NFYDNACMMQADAC-----IGDRKVVPAMCTMEYKPVC GSDGRTYANECQL---NARGCL	91
OL606745	40	NFYDNACMMQADAC-----IGDRKVVPAMCTMEYKPVC GSDGRTYANECQL---NARGCL	91
P10184	61	CFADIAFGIEVNC SLYASGIGKDGTSWVACPRNLKPVC GTDGSTYSNECGICLYNREHGA	120
Predicted	92	GVMKMSDGECP SRSGRRSL-----GECLTICTMEYNPVC GTNGKMYSNFCQLQV--	140
OL606745	92	GVMKMSDGECP SRSGRRSL-----GECLTICTMEYNPVC GTNGKMYSNFCQLQV--	140
P10184	121	NVEKEYDGE CRPKHVTIDCSPYLQVVRDGN TMVACPRILKPVC GSDSF TYDNECGICAYN	180
Predicted	141	-----VSCLSNGEIGMP-----ACPYNYPVC GSDGITYGNLC	173
OL606745	141	-----VSCLSNGEIGMP-----ACPYNYPVC GSDGITYGNLC	173
P10184	181	AEHHTNISK LHDGECKLEIGSVDCSKYPSTVSKDGR TLVACPRILSPVC GTDGF TYDNEC	240
Predicted	174	SLESQSCFNVTYVS---DGECEEASKAADDEDEPECNTA-----CTYEYNPVC GSDGV	223
OL606745	174	SLESQSCFNVTYVS---DGECEEASKAADDEDEPECNTA-----CTYEYNPVC GSDGV	223
P10184	241	GICAHNAEQ RTHVSKKHDGKCRQEIPEIDCDQYPT RKTGGKLLVRCPRILLPVC GTDGF	300
Predicted	224	KYANPCVLGVASCQ SNGAI-----SMPM-----CTLEYQPVC	255
OL606745	224	KYANPCVLGVASCQ SNGAI-----SMPM-----CTLEYQPVC	255
P10184	301	TYDNECGICA HNAQHGT EVKKSHDGRCKERSTPLDCTQYLSNTQNGEAITACPFILQEV C	360
Predicted	256	GSDGQTYGNQCMLDAQKC---LNVTKVSEGEC-----EAAALMSGGEKQEKCDTTCTY	305
OL606745	256	GSDGQTYGNQCMLDAQKC---LNVTKVSEGEC-----EAAALMSGGEKQEKCDTTCTY	305
P10184	361	GTDGVTYSNDCSLCAHNIELGTSVAKKHDGRCEEVPELDCSKYKTSTLKDGRQVVACTM	420
Predicted	306	EYNPVC GSDGVKYANPCVLKVAS-----CQSGGAISMP ICTM	342
OL606745	306	EYNPVC GSDGVKYANPCVLKVAS-----CQSGGAISMP ICTM	342
P10184	421	IYDPVCATNGVTYASECTLCAHNLEQRTNLGKRKNGRCEE DITKEHCRE FQKVS-PICTM	479
Predicted	343	EYRPVC GSDGLMYGNRCMLNAQRCM GVERADWSKCGTVQKKRKLVDLLRAFANQ	396
OL606745	343	EYRPVC GSDGLMYGNRCMLNAQRCM GVERADWSKCGTVQKKRKLVDLLRAFANQ	396
P10184	480	EYVPHCGSDGVTYSNRCFF-----CNAYVQSNRTLNLVSMAAC-	517

Figure S.5. Aminoacid sequence of an ovoinhibitor-like protein isolated from *Glycera alba*.

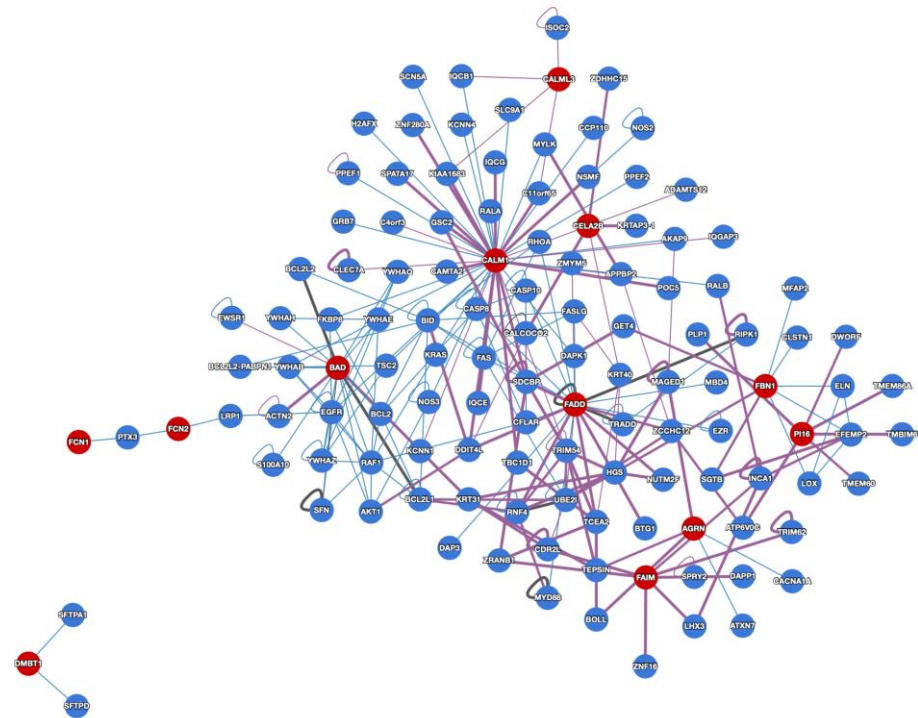
The validated sequence (GenBank accession OL606745) was aligned with the original

translated transcript as reconstructed by Trinity (predicted) along with the best hit produced by homology-matching against Swiss-Prot (ovoinhibitor from *Gallus gallus*, P10184).

Predicted	1	MTNLIILFACFAFAVATPIELINPDANNYEEMMSKRGTCGAQLARLSSRSATNFLNAHNT	60
OL606746	1	MTNLIILFACFAFAVATPIELINPDANNYEEMMSKRGTCGAQLARLSSRSATNFLNAHNT	60
P09042	1	-MEFVLFSQMSSFFLVSTLLL-----FLIISHSCHAQ-----NSQQDYLDAHNT	43
Predicted	61	KRQQEGQGLSGLTWDSDLAARAQELANKCVFNHGLATDCN----GKSCGQNIYYSGGSSF	116
OL606746	61	KRQQEGQGLSGLTWDSDLAARAQELANKCVFNHGLATDCN----GKSCGQNIYYSGGSSF	116
P09042	44	ARAD--VGVEPLTWDDQVAAYAQNYSQ-----LAADCNLVHSHGQYGENLAWGSGDFL	95
Predicted	117	SAAKVVDSWYSEKNDFTYSSNSCASGKACGHYTQIVWKSTQKVGCADCTGKVMGYSPE	176
OL606746	117	SAAKVVDSWYSEKNDFTYSSNSCASGKACGHYTQIVWKSTQKVGCADCTGKVMGYSPE	176
P09042	96	TAAKAVEMWVNEKQYYAHSNTCAQGQVCGHYTQVVWRNSVRVGCARVQCN-----NGG	149
Predicted	177	YIAVCNYPFPGNYIGQKPY	195
OL606746	177	YIAVCNYPFPGNYIGQKPY	195
P09042	150	YIVSCNYDPPGNVIGKSPY	168

Figure S.6. Aminoacid sequence of a pathogenesis-related protein (a CRISP) isolated from *Hediste diversicolor*. The validated sequence (GenBank accession OL606746) was aligned with the original translated transcript as reconstructed by Trinity (predicted) along with the best hit produced by homology-matching against Swiss-Prot (pathogenesis-related protein 1C from *Nicotiana tabacum*, P09042).

A



B

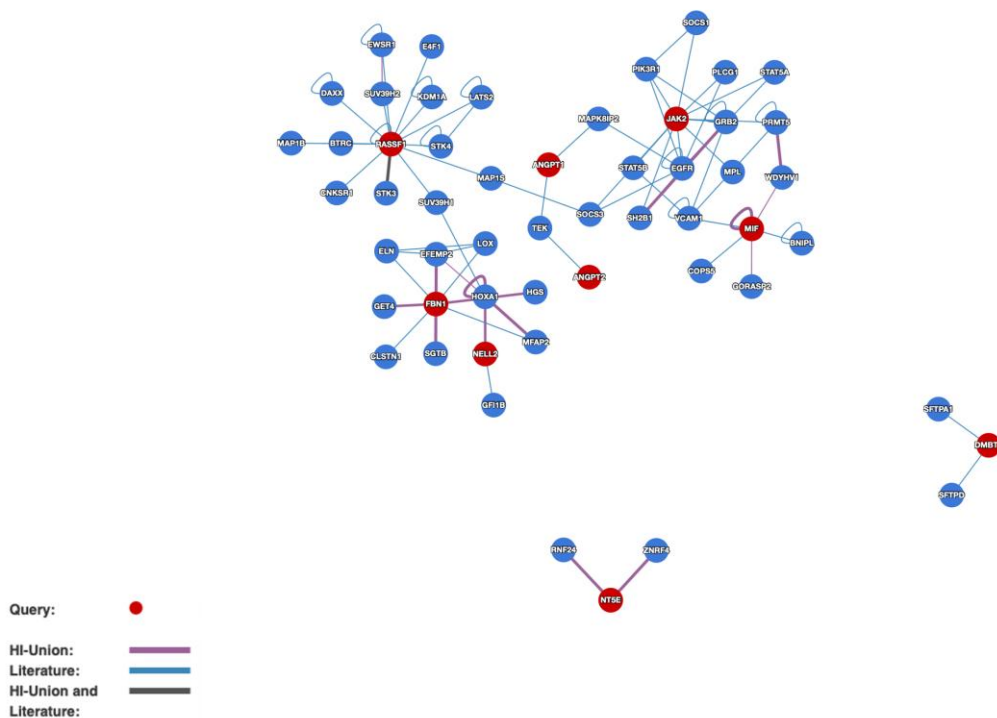
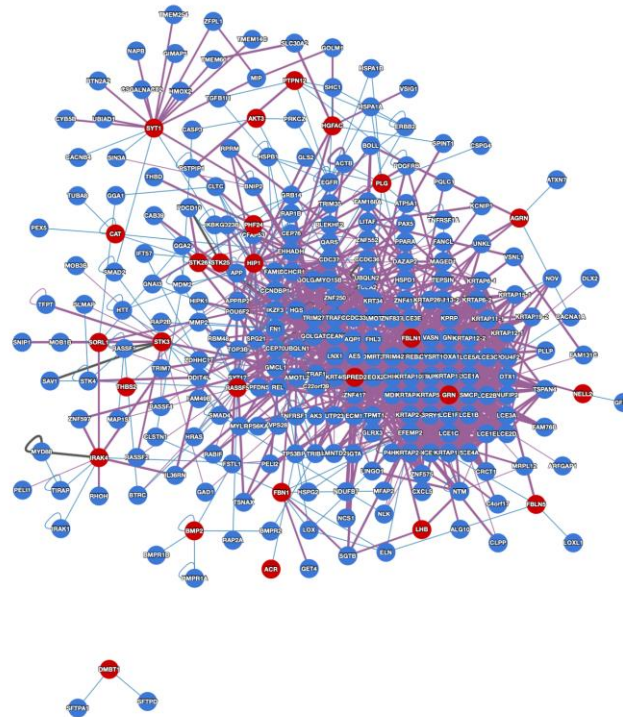


Figure S.8. Protein-protein interactions between the potential interactors from the proboscis and skin of *Glycera alba* and the human proteome retrieved from the HuRI platform. Interactors

were the human homologs of putative proteins up-regulated in the proboscis (A) and in the skin (B).

A



B

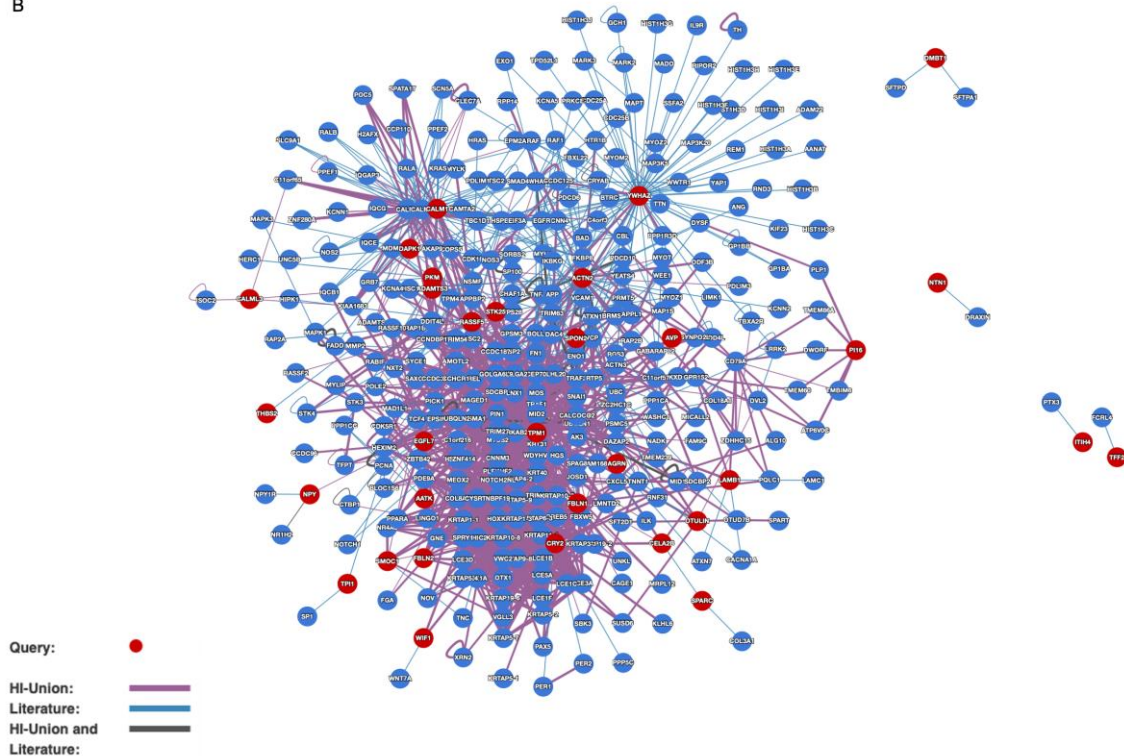


Figure S.9. Protein-protein interactions between the potential interactors from the glands and proboscis of *Hediste diversicolor* and the human proteome retrieved from the HuRI platform.

Interactors were the human homologs of putative proteins up-regulated in the glands (A) and in the proboscis (B).

Table S.4. Enriched biological processes affected by the potential interactors from the proboscis of *Glycera alba* and their human targets. The interactors were the human homologs of the proteins up-regulated in the proboscis. The FDR-adjusted $p < 0.05$ was set as the cut-off. The GO enrichment analysis was performed in the Database for Annotation, Visualization and Integrated Discovery (DAVID).

	GO Terms	%	FDR p	Human targets	Interactor
1	positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway [GO:1900740]	8.065	1.60E-10	BCL2, BID, CASP8, SFN, YWHAB, YWHAE, YWHAH, YWHAZ, YWHAQ	BAD
2	negative regulation of extrinsic apoptotic signaling pathway via death domain receptors [GO:1902042]	8.065	2.12E-10	FAS, FASLG, CASP8, DAPK1, NOS3, RAF1, TRADD, RIPK1, CFLAR	FADD
3	extrinsic apoptotic signaling pathway via death domain receptors [GO:0008625]	7.258	2.98E-08	FAS, FASLG, BCL2, BID, CASP10, DAPK1, TRADD	BAD, FADD
4	regulation of extrinsic apoptotic signaling pathway via death domain receptors [GO:1902041]	5.645	1.78E-07	FAS, FASLG, CASP8, TRADD, RIPK1, CFLAR	FADD
5	membrane organization [GO:0061024]	6.452	2.14E-07	SFN, RALA, YWHAB, YWHAE, YWHAH, YWHAZ, YWHAQ, TBC1D1	
6	activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway [GO:0097296]	4.839	2.08E-06	FAS, FASLG, CASP8, TRADD, RIPK1	FADD
7	negative regulation of apoptotic process [GO:0043066]	13.710	4.05E-06	AKT1, FAS, BCL2, BCL2L1, BCL2L2, EGFR, MYD88, RAF1, SLC9A1, TM6IM6, YWHAZ, ZNF16, LHX3, CFLAR, SPRY2, FKBP8	FAIM

	GO Terms	%	FDR_p	Human targets	Interactor
8	positive regulation of I-kappaB kinase/NF-kappaB signaling [GO:0043123]	8.871	9.95E-06	FASLG, RHOA, CASP8, CASP10, MYD88, UBE2I, TRADD, RIPK1, CFLAR, TRIM62	FADD
9	regulation of apoptotic process [GO:0042981]	9.677	1.14E-05	ACTN2, FAS, BCL2L1, BCL2L2, BID, CASP8, CASP10, DAPK1, RAF1, CFLAR, MAGED1	FADD
10	death-inducing signaling complex assembly [GO:0071550]	4.032	1.14E-05	CASP8, RAF1, TRADD, RIPK1	FADD
11	extrinsic apoptotic signaling pathway [GO:0097191]	5.645	2.41E-05	FAS, FASLG, CASP8, TRADD, RIPK1	BAD, FADD
12	cellular response to mechanical stimulus [GO:0071260]	6.452	2.66E-05	AKT1, FAS, CASP8, EGFR, MYD88, SLC9A1	BAD, FADD
13	apoptotic signaling pathway [GO:0097190]	6.452	2.66E-05	FAS, FASLG, CASP8, CASP10, DAPK1, DAP3, RIPK1	FADD
14	apoptotic process [GO:0006915]	13.710	4.06E-05	FAS, FASLG, BCL2, CASP8, CASP10, DAPK1, MYD88, RAF1, RALB, DAP3, TRADD, RIPK1, CFLAR, FKBP8	BAD, FADD, FAIM
15	activation of cysteine-type endopeptidase activity involved in apoptotic process [GO:0006919]	6.452	6.78E-05	FAS, FASLG, BID, CASP8, TRADD, RIPK1	BAD, FADD
16	extrinsic apoptotic signaling pathway in absence of ligand [GO:0097192]	4.839	1.52E-04	FAS, BCL2, BCL2L1, BCL2L2	BAD, FADD
17	necroptotic signaling pathway [GO:0097527]	3.226	3.15E-04	FAS, FASLG, RIPK1	FADD
18	release of cytochrome c from mitochondria [GO:0001836]	4.032	7.47E-04	BCL2, BCL2L1, BID, SFN	BAD
19	substantia nigra development [GO:0021762]	4.839	8.14E-04	RHOA, PLP1, YWHAH, YWHAQ	CALM1

	GO Terms	%	FDR_p	Human targets	Interactor
20	viral process [GO:0016032]	8.871	1.08E-03	RHOA, ATP6V0C, H2AFX, RALA, TSC2, UBE2I, YWHAB, YWHAE, CFLAR, CALCOCO2, FKBP8	
21	signal transduction [GO:0007165]	16.129	5.74E-03	AKT1, FAS, FASLG, DAPK1, EGFR, SFN, GRB7, MYD88, RAF1, RALA, RALB, YWHAZ, TRADD, HGS, AKAP9, DAPP1, TRIM54, IQGAP3	FADD, AGRN, AGRN
22	positive regulation of apoptotic process [GO:0043065]	8.065	5.98E-03	AKT1, FAS, FASLG, BCL2L1, BID, SLC9A1, TRADD, RIPK1	BAD, FADD
23	positive regulation of cell growth [GO:0030307]	4.839	9.14E-03	AKT1, BCL2, EGFR, SFN, SDCBP, SLC9A1	
24	intrinsic apoptotic signaling pathway in response to DNA damage [GO:0008630]	4.032	1.00E-02	BCL2, BCL2L1, BCL2L2, SFN	BAD
25	regulation of nitric-oxide synthase activity [GO:0050999]	3.226	2.54E-02	AKT1, EGFR, NOS3	CALM1
26	protein heterooligomerization [GO:0051291]	4.032	3.62E-02	CASP8, YWHAB, TRADD, RIPK1	FADD
27	positive regulation of peptidyl-serine phosphorylation [GO:0033138]	4.032	3.96E-02	AKT1, BCL2, RAF1, AKAP9, SPRY2	
28	Ras protein signal transduction [GO:0007265]	4.032	3.96E-02	KRAS, RALA, RALB, SDCBP, IQGAP3	
29	positive regulation of intrinsic apoptotic signaling pathway [GO:2001244]	3.226	4.45E-02	BCL2, BCL2L1, BID	BAD

FDR_p – False discovery rate adjusted p-value; % – Percentage of proteins

Table S.5. Enriched biological processes affected by the potential interactors from the skin of *Glycera alba* and their human targets. The interactors were the human homologs of the proteins up-regulated in the skin. The FDR-adjusted $p < 0.05$ was set as the cut-off. The GO enrichment analysis was performed in the Database for Annotation, Visualization and Integrated Discovery (DAVID).

GO Terms	%	FDR p	Human targets	Interactor
1 leukocyte migration [GO:0050900]	12.963	7.88E-04	GRB2, PIK3R1, PLCG1, TEK	ANGPT1, ANGPT2, MIF, MIF
2 JAK-STAT cascade [GO:0007259]	9.259	7.88E-04	STAT5A, STAT5B, SOCS1, SOCS3	JAK2
3 Tie signaling pathway [GO:0048014]	5.556	4.84E-03	TEK	ANGPT1, ANGPT2
4 viral process [GO:0016032]	14.815	4.84E-03	DAXX, E4F1, GRB2, PIK3R1, PLCG1, SUV39H1, VCAM1, BTRC	
5 glomerulus vasculature development [GO:0072012]	5.556	6.68E-03	TEK	ANGPT1, ANGPT2
6 cell proliferation [GO:0008283]	14.815	1.15E-02	E4F1, EGFR, ELN, MPL, GFI1B, PRMT5, KDM1A	MIF, MIF

FDR p – False discovery rate adjusted p-value; % – Percentage of proteins

Table S.6. Enriched biological processes affected by the potential interactors from the glands of *Hediste diversicolor* and their human targets. The interactors were the human homologs of the proteins up-regulated in the glands. The FDR-adjusted $p < 0.05$ was set as the cut-off. The GO enrichment analysis was performed in the Database for Annotation, Visualization and Integrated Discovery (DAVID).

	GO Terms	%	FDR p	Target	Interactor
1	peptide cross-linking [GO:0018149]	5.495	3.80E-12	FN1, LCE2B, CRCT1, LCE4A, LCE5A, LCE1A, LCE1B, LCE1C, LCE1D, LCE1E, LCE1F, LCE2D, LCE3A, LCE3C, LCE3E	
2	keratinocyte differentiation [GO:0030216]	6.227	3.80E-12	CASP3, STK4, LCE2B, CRCT1, SAV1, LCE4A, LCE5A, LCE1A, LCE1B, LCE1C, LCE1D, LCE1E, LCE1F, LCE2D, LCE3A, LCE3C, LCE3E	
3	keratinization [GO:0031424]	4.762	7.75E-10	LCE2B, LCE4A, LCE5A, LCE1A, LCE1B, LCE1C, LCE1D, LCE1E, LCE1F, LCE2D, LCE3A, LCE3C, LCE3E	
4	hippo signaling [GO:0035329]	2.564	5.52E-04	CASP3, STK4, AMOTL2, SAV1, MOB1B, AMOT	STK3, STK3
5	signal transduction by protein phosphorylation [GO:0023014]	2.930	6.33E-04	BMPR1A, BMPR1B, BMPR2, STK4, CAB39	STK3, STK3, STK25, STK26, STK26
6	protein stabilization [GO:0050821]	4.396	7.07E-04	HSPA1A, HSPA1B, HSPD1, STK4, RASSF2, CDC37, RASSF1, PDCD10, NLK, SAV1	HIP1, STK3, STK3
7	positive regulation of I-kappaB kinase/NF-kappaB signaling [GO:0043123]	4.396	3.07E-03	ECM1, IRAK1, MYD88, REL, TNFRSF1A, IKBKG, LITAF, ZDHHC17, PELI2, PELI1, TIRAP	IRAK4

	GO Terms	%	FDR _p	Target	Interactor
8	signal transduction [GO:0007165]	12.821	4.36E-03	ECM1, EGFR, ERBB2, GRB14, HRAS, IRAK1, MYD88, PDGFRB, PRKCZ, RAP2B, CXCL5, SHC1, STK4, THBD, TNFRSF1A, TRAF1, TRAF2, BTRC, PSTPIP1, HGS, LITAF, ZDHHC17, RPS6KA6, SAV1, RASSF4, LINGO1, TIRAP	GRN, GRN, LHB, SORL1, STK3, STK3, AKT3, STK25, IRAK4, AGRN, AGRN, AGRN, AGRN
9	cellular response to BMP stimulus [GO:0071773]	2.198	8.45E-03	BMPR1A, BMPR1B, BMPR2, SMAD4, SPINT1	BMP2
10	response to hydrogen peroxide [GO:0042542]	2.564	1.02E-02	CASP3, HSPD1, PDGFRB, PDCD10	CAT, STK25, STK26, STK26
11	extracellular matrix organization [GO:0030198]	4.396	1.17E-02	APP, ELN, FN1, HSPG2, LOX, LOXL1, MFAP2, SPINT1	FBLN1, FBN1, FBN1, FBLN5, AGRN, AGRN, AGRN, AGRN
12	positive regulation of NF-kappaB transcription factor activity [GO:0051092]	3.663	1.17E-02	HSPA1A, HSPA1B, IRAK1, MYD88, PRKCZ, TRAF1, TRAF2, IKBKG, TIRAP	CAT
13	positive regulation of MAP kinase activity [GO:0043406]	2.564	1.79E-02	EGFR, ERBB2, HRAS, IRAK1, PDGFRB, MAGED1, PDCD10	
14	protein phosphorylation [GO:0006468]	6.593	2.13E-02	APP, BMPR1A, BMPR1B, ERBB2, IRAK1, SMAD2, PRKCZ, STK4, RASSF2, RPS6KA6, NLK, TRIB3, HIPK1	BMP2, STK3, STK3, AKT3, STK25, STK26, STK26

	GO Terms	%	FDR _p	Target	Interactor
15	cellular protein metabolic process [GO:0044267]	3.297	2.26E-02	APP, HSPG2, MMP2, SFTPD, GGA2, GGA1, SFTPA1	DMBT1, DMBT1, DMBT1, DMBT1, DMBT1, PLG, PLG

FDR_p – False discovery rate adjusted p-value; % – Percentage of proteins

Table S.7. Enriched biological processes affected by potential interactors from the proboscis of *Hediste diversicolor* and their human targets.

The interactors were the human homologs of the proteins up-regulated in the proboscis. The FDR-adjusted $p < 0.05$ was set as the cut-off.

The GO enrichment analysis was performed in the Database for Annotation, Visualization and Integrated Discovery (DAVID).

	GO Terms	%	FDR p	Target	Interactor
1	regulation of gene silencing [GO:0060968]	2.8169	2.58E-11	HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F	
2	DNA replication-dependent nucleosome assembly [GO:0006335]	3.0986	1.97E-07	HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F, CHAF1A	
3	cellular protein metabolic process [GO:0044267]	4.7887	4.00E-07	APP, FGA, MMP2, SFTPD, UBC, HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F, SFTPA1	DMBT1, DMBT1
4	telomere organization [GO:0032200]	2.8169	4.25E-07	HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F	
5	chromatin silencing at rDNA [GO:0000183]	2.8169	7.66E-06	HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F	
6	negative regulation of gene expression, epigenetic [GO:0045814]	3.0986	7.79E-06	TRIM27, HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F	

	GO Terms	%	FDR _p	Target	Interactor
7	regulation of tumor necrosis factor-mediated signaling pathway [GO:0010803]	2.5352	1.57E-05	TNFAIP3, TRAF1, TRAF2, UBC, IKBKG, MADD, RNF31, HIPK1	OTULIN, OTULIN, OTULIN, OTULIN
8	protein heterotetramerization [GO:0051290]	2.8169	1.58E-05	HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F	
9	MAPK cascade [GO:0000165]	5.6338	1.10E-04	ARAF, CALM2, CALM3, EGFR, HRAS, KRAS, MARK3, MAP3K5, MOS, PPP5C, MAPK1, MAPK3, PSMA1, PSMC5, RAF1, UBC, AKAP9, LRRK2	ACTN2, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1
10	muscle contraction [GO:0006936]	3.6620	1.45E-04	ACTN3, CALM2, CALM3, CRYAB, MYLK, TPM4, TTN, DYSF, MYOM2, MYOT, TRIM63	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, TPM1
11	response to calcium ion [GO:0051592]	2.8169	2.12E-04	AANAT, BAD, CALM2, CALM3, EGFR, FGA, TTN, PDCD6	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, SPARC, SPARC

	GO Terms	%	FDR _p	Target	Interactor
12	positive regulation of gene expression, epigenetic [GO:0045815]	2.8169	3.46E-04	HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F	
13	positive regulation of protein serine/threonine kinase activity [GO:0071902]	2.2535	4.88E-04	CALM2, CALM3, RALB, STK3, STK4, CDK5R1, PDCD10	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1
14	cell-cell adhesion [GO:0098609]	5.3521	4.88E-04	CBL, MARK2, ENO1, GOLGA2, SDCBP, YWHAE, HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F, PDLIM1	PKM, PKM, YWHAZ
15	peptide cross-linking [GO:0018149]	2.5352	4.88E-04	COL3A1, FN1, LCE3D, LCE5A, LCE1A, LCE1B, LCE1C, LCE1F, LCE3A	
16	protein phosphorylation [GO:0006468]	7.0423	6.33E-04	APP, CDC25B, CTBP1, MARK2, ILK, LIMK1, MAP3K5, MYLK, CDK16, PRKAB2, PRKCE, MAPK1, MAPK3, RAF1, STK3, STK4, PICK1, RASSF2, MAP3K20, LRRK2, HIPK1, SBK3	DAPK1, AATK, AATK, STK25
17	hippo signaling [GO:0035329]	1.9718	1.01E-03	DVL2, STK3, STK4, YWHAE, YAP1, WWTR1, AMOTL2	

	GO Terms	%	FDR _p	Target	Interactor
18	blood coagulation [GO:0007596]	4.2254	1.01E-03	FGA, GP1BA, GP1BB, HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F, BLOC1S6, AK3	
19	G2/M transition of mitotic cell cycle [GO:0000086]	3.6620	1.01E-03	CALM2, CALM3, CDC25A, CDC25B, TPD52L1, UBC, WEE1, YWHAE, BTRC, CCP110, AKAP9, CEP70	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1
20	circadian regulation of gene expression [GO:0032922]	2.5352	1.01E-03	PER1, PPARA, PPP1CA, PPP1CC, PER2, USP2, MAGED1, PRMT5	CRY2
21	keratinocyte differentiation [GO:0030216]	2.8169	1.11E-03	NOTCH1, STK4, YAP1, LCE3D, LCE5A, LCE1A, LCE1B, LCE1C, LCE1F, LCE3A	
22	wound healing [GO:0042060]	2.8169	1.61E-03	COL3A1, EGFR, FN1, TNC, MAP3K5, PPARA, RAF1	SPARC, SPARC, TFF2, TPM1
23	negative regulation of extrinsic apoptotic signaling pathway via death domain receptors [GO:1902042]	1.9718	2.56E-03	FGA, NOS3, RAF1, TNFAIP3, TRAF2, FADD	DAPK1
24	regulation of circadian rhythm [GO:0042752]	2.2535	2.71E-03	PER1, PPARA, PPP1CA, PPP1CC, PER2, BTRC, MAGED1	CRY2
25	positive regulation of protein autophosphorylation [GO:0031954]	1.6901	2.83E-03	CALM2, CALM3, RAP2A, RAP2B, RASSF2	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1

GO Terms	%	FDR _p	Target	Interactor
				CALM1, CALM1, CALM1
26 nucleosome assembly [GO:0006334]	3.0986	6.01E-03	H2AFX, HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F	
27 response to hypoxia [GO:0001666]	3.6620	7.02E-03	ANG, BAD, CRYAB, KCNA5, SMAD4, MMP2, NOS2, PPARA, RAF1, TH, VCAM1, PDLIM1	PKM, PKM
28 regulation of nitric-oxide synthase activity [GO:0050999]	1.6901	7.56E-03	CALM2, CALM3, EGFR, GCH1, NOS3	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1
29 positive regulation of MAP kinase activity [GO:0043406]	2.2535	7.60E-03	EGFR, HRAS, ILK, KRAS, TPD52L1, MAGED1, PDCD10, LRRK2	
30 positive regulation of epithelial cell proliferation [GO:0050679]	2.2535	8.18E-03	BAD, EGFR, HRAS, LAMC1, NOTCH1, SCN5A, NR4A3	LAMB1, LAMB1, LAMB1, LAMB1

	GO Terms	%	FDR _p	Target	Interactor
31	platelet degranulation [GO:0002576]	2.8169	8.53E-03	APP, CALM2, CALM3, FGA, FN1, TTN	ACTN2, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, ITIH4, SPARC, SPARC
32	positive regulation of cell migration [GO:0030335]	3.6620	1.12E-02	EGFR, GRB7, HRAS, ILK, MYLK, NOTCH1, MAPK1, SDCBP, SNAI1, PDCD10, COL18A1, WASHC1	LAMB1, LAMB1, LAMB1, LAMB1
33	positive regulation of axon extension [GO:0045773]	1.6901	1.31E-02	FN1, ILK, LIMK1, MAPT, DISC1	NTN1, NTN1
34	gene silencing by RNA [GO:0031047]	2.8169	1.38E-02	HIST1H3A, HIST1H3D, HIST1H3C, HIST1H3E, HIST1H3I, HIST1H3G, HIST1H3J, HIST1H3H, HIST1H3B, HIST1H3F	
35	keratinization [GO:0031424]	1.9718	1.47E-02	LCE3D, LCE5A, LCE1A, LCE1B, LCE1C, LCE1F, LCE3A	
36	angiogenesis [GO:0001525]	3.9437	1.61E-02	ANG, COL8A1, FN1, MEOX2, MMP2, MYH9, NOS3, NOV, EGFL7 WNT7A, PDCD6, PDCD10, COL18A1, UNC5B	
37	response to corticosterone [GO:0051412]	1.4085	1.61E-02	AANAT, CALM2, CALM3, TH	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1

	GO Terms	%	FDR _p	Target	Interactor
38	platelet activation [GO:0030168]	2.8169	1.61E-02	COL3A1, FGA, GP1BA, GP1BB, PRKCE, MAPK1, MAPK3, RAF1, RAP2B	YWHAZ
39	extracellular matrix organization [GO:0030198]	3.6620	1.65E-02	APP, COL3A1, COL8A1, FGA, FN1, TNC, LAMC1, VCAM1, COL18A1	FBLN1, LAMB1, LAMB1, LAMB1, LAMB1, SPARC, SPARC, AGRN, AGRN, AGRN, AGRN
40	microtubule cytoskeleton organization [GO:0000226]	2.2535	1.78E-02	MARK2, MAPT, MARK3, MID1, ATXN7, WEE1, DISC1, MAP1S	
41	entrainment of circadian clock by photoperiod [GO:0043153]	1.4085	2.26E-02	PER1, PPP1CA, PPP1CC, USP2	CRY2
42	viral process [GO:0016032]	4.5070	2.42E-02	ATP6V0C, H2AFX, MDM2, MAP3K5, MAPK1, MAPK3, RALA, SP1, SP100, TSC2, VCAM1, YWHAE, BTRC, CALCOCO2, FKBP8	FBLN1
43	Fc-epsilon receptor signaling pathway [GO:0038095]	3.3803	2.42E-02	CALM2, CALM3, HRAS, KRAS, MAPK1, MAPK3, PSMA1, PSMC5, UBC, IKBKG, BTRC	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1

	GO Terms	%	FDR _p	Target	Interactor
44	regulation of cell cycle [GO:0051726]	2.8169	2.42E-02	CDC25A, CTBP1, CDK16, TSC2, WEE1, MADD, PER2, BTRC, COPS5, CCNDBP1	
45	positive regulation of protein dephosphorylation [GO:0035307]	1.4085	2.50E-02	CALM2, CALM3, PIN1, NSMF	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1
46	cerebral cortex development [GO:0021987]	1.9718	2.62E-02	BAD, COL3A1, H2AFX, PAX5, TH, YWHAE	NPY
47	positive regulation of protein phosphorylation [GO:0001934]	2.8169	2.66E-02	DVL2, EGFR, HRAS, KRAS, PIN1, MAPK3, RALB, RAP2A, STK4, LRRK2	
48	negative regulation of protein binding [GO:0032091]	1.9718	2.66E-02	GOLGA2, PIN1, PPARA, PPP1CA, RALB, TMBIM6, LRRK2	
49	muscle filament sliding [GO:0030049]	1.6901	2.66E-02	ACTN3, TNNT1, TPM4, TTN	ACTN2, TPM1
50	positive regulation of nitric-oxide synthase activity [GO:0051000]	1.4085	2.66E-02	CALM2, CALM3, GCH1, KRAS	CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1, CALM1
51	regulation of cytokinesis [GO:0032465]	1.4085	2.66E-02	CALM2, CALM3, PIN1, CCP110	CALM1, CALM1, CALM1, CALM1,

GO Terms	%	FDR _p	Target	Interactor
				CALM1, CALM1, CALM1, CALM1, CALM1
52 stimulatory C-type lectin receptor signaling pathway [GO:0002223]	2.5352	2.97E-02	HRAS, KRAS, PSMA1, PSMC5, RAF1, UBC, IKBKG, BTRC, CLEC7A	
53 response to ethanol [GO:0045471]	2.5352	2.97E-02	BAD, CBL, HTR1B, TNC, TBXA2R, TH, VCAM1	AVP, SPARC, SPARC
54 axon guidance [GO:0007411]	3.0986	3.03E-02	HRAS, KRAS, SMAD4, MAPK1, MAPK3, NR4A3, CDK5R1, ZNF280A, DRAXIN	NTN1, NTN1, SPON2
55 negative regulation of neuron death [GO:1901215]	1.6901	3.11E-02	PPARA, PPP5C, REL, TRAF2, IKBKG, LRRK2	
56 platelet aggregation [GO:0070527]	1.6901	3.35E-02	FGA, GPIBA, HSPB1, ILK, MYH9, RAP2B	
57 signal transduction [GO:0007165]	10.7042	3.35E-02	EGFR, FGA, GRB7, HRAS, IL9R, LIMK1, PDE9A, PRKAB2, PRKCE, MAPK1, RAF1, RALA, RALB, RAP2B, CXCL5, STK3, STK4, TRAF1, TRAF2, FADD, BTRC, HGS, AKAP9, APPL1, TRIM54, TRIM63, LINGO1, IQGAP3, UNC5B, RASSF10	AVP, DAPK1, SPARC, SPARC, YWHAZ, STK25, WIF1, SMOC1, SMOC1, SMOC1, AGRN, AGRN, AGRN, AGRN
58 neuron death [GO:0070997]	1.1268	3.35E-02	CBL, MEOX2, SLC9A1, LRRK2	
59 endocardium development [GO:0003157]	0.8451	3.51E-02	NOTCH1, STK3, STK4	
60 positive regulation of cyclic nucleotide metabolic process [GO:0030801]	0.8451	3.51E-02	CALM2, CALM3	CALM1, CALM1, CALM1, CALM1,

GO Terms	%	FDR _p	Target	Interactor
				CALM1, CALM1, CALM1, CALM1, CALM1
61 positive regulation of cell proliferation [GO:0008284]	5.6338	4.06E-02	CDC25B, EGFR, FN1, HRAS, TNC, ILK, KRAS, MDM2, NOTCH1, MAPK1, CXCL5, SDCBP, HDAC4, YAP1, PDCD10, WWTR1, COL18A1, HIPK1	AVP, NTN1, NTN1
62 positive regulation of JNK cascade [GO:0046330]	1.9718	4.30E-02	HRAS, MAP3K5, SDCBP, STK3, TPD52L1, WNT7A, RASSF2	
FDR _p – False discovery rate adjusted p-value; % – Percentage of proteins				