doi: 10.1387/ijdb.103063bm

THE INTERNATIONAL JOURNAL OF
DEVELOPMENTAL
BIOLOGY
www.intjdevbiol.com

SUPPLEMENTARY MATERIAL

corresponding to:

A polymorphic, thrombospondin domain-containing lectin is an oocyte marker in *Hydractinia*: implications for germ cell specification and sex determination

BRAHIM MALI, R CATHRIONA MILLANE, GÜNTER PLICKERT, MARCUS FROHME and URI FRANK

A

```
CELc        GGTTACTTGTTCTTCCTGCAGTTTTGTATTATCTAGAAAAAAACAAAGATGGTGAATTTG 60
CELd        GGTTACTTGTTCTTCCTGCAGTTTTGTATTATCTAGAAAAAAACAAAGATGGTGAATTTG 60
CELa        GGTTACTTGTTCTTCCTGCAGTTTTGTATTATCTAGAAAAAAACAAAGATGGTGAATTTG 60
CELb        GGTTACTTGTTCTTCCTGCAGTTTTGTATTATCTAGAAAAAAACAAAGATGGTGAATTTG 60
            ************************************************************

CELc        CGATTATCCTTCGTGGTTCTTTTCATTGTTTTCATGTACACGAACTCTGGTTTTGCAAAG 120
CELd        CGATTATCCTTCGTGGTTCTTTTCATTGTTTTCATGTACACGAACTCTGGTTTTGCAAAG 120
CELa        CGATTATCCTTCGTGGTTCTTTTCATTGTTTTCATGTACACGAACTCTGGTTTTGCAAAG 120
CELb        CGATTATCCTTCGTGGTTCTTTTCATTGTTTTCATGTACACGAACTCTGGTTTTGCAAAG 120
            ************************************************************

CELc        AGTTCAAAGCGTGTGTGCGAGAATAAAACAATGCGCATTAGCTGTTCGTCAGGTCGAGTG 180
CELd        AGTTCAAAGCGTGTGTGCGAGAATAAAACAATGCGCATTAGCTGTTCGTCAGGTCGAGTG 180
CELa        AGTTCAAAGCGTGTGTGCGAGAATAAAACAATGCGCATTAGCTGTTCGTCAGGTCGAGTG 180
CELb        AGTTCAAAGCGTGTGTGCGAGAATAAAACAATGCGCATTAGCTGTTCGTCAGGTCGAGTG 180
            ************************************************************

CELc        ATACAAATTCGAAGAGCAATGTATGGCCGATACACAACTAGTTATTGCGGCTGGAATATC 240
CELd        ATACAAATTCGAAGAGCAATGTATGGCCGATACACAACTAGTTATTGCGGCTGGAATATC 240
CELa        ATACAAATTCGAAGAGCAATGTATGGCCGATACACAACTAGTTATTGCGGCTGGAATATC 240
CELb        ATACAAATTCGAAGAGCAATGTATGGCCGATACACAACTAGTTATTGCGGCTGGAATATC 240
            ************************************************************

CELc        TTTCAAGGACGTTGTAACTCTCACTCTAGCAGATCTTTACATGAAGTAAAATTAAGATGC 300
CELd        TTTCAAGGACGTTGTAACTCTCACTCTAGCAGATCTTTACATGAAGTAAAATTAAGATGC 300
CELa        TTTCAAGGACGTTGTAACTCTCACTCTAGCAGATCTTTACATGAAGTAAAATTAAGATGC 300
CELb        TTTCAAGGACGTTGTAACTCTCACTCTAGCAGATCTTTACATGAAGTAAAATTAAGATGC 300
            ************************************************************

CELc        CAGAACAAGAGATCGTGTAGCGTGAATGCTGTTAACGGGGTGTTTGGTGATCCTTGCTTT 360
CELd        CAGAACAAGAGATCGTGTAGCGTGAATGCTGTTAACGGGGTGTTTGGTGATCCTTGCTTT 360
CELa        CAGAACAAGAGATCGTGTAGCGTGAATGCTGTTAACGGGGTGTTTGGTGATCCTTGCTTT 360
CELb        CAGAACAAGAGATCGTGTAGCGTGAATGCTGTTAACGGGGTGTTTGGTGATCCTTGCTTT 360
            ************************************************************

CELc        GGAACGAGAAAATATTTAGAAGTTGATTATTACTGCAGAAATTTACCAGTTGATGGAGGA 420
CELd        GGAACGAGAAAATATTTAGAAGTTGATTATTACTGCAGAAATTTACCAGTTGATGGAGGA 420
CELa        GGAACGAGAAAATATTTAGAAGTTGATTATTACTGCAGATATTTACCAGTTGATGGAGGA 420
CELb        GGAACGAGAAAATATTTAGAAGTTGATTATTACTGCAGATATTTACCAGTTGATGGAGGA 420
            *************************************** ********************

CELc        TACACAAAGTGGAATATCCAAAGGTGTTCTCGCACCTGTGGTGGAGGGACAAAAACATTT 480
CELd        TACACAAAGTGGAATATCCAAAGGTGTTCTCGCACCTGTGGTGGAGGGACAAAAACATTT 480
CELa        TACACAAAGTGGAATATCCATAGGTGTTCTCGCTCTTGTGGTGGAGGGACAAAAACATTT 480
CELb        TACACAAAGTGGAATATCCATAGGTGTTCTCGCTCTTGTGGTGGAGGGACAAAAACATTT 480
            ******************** ************ * ************************

CELc        ACTAGGACATGCACAAACCCTTCTCCAAAATTTGGAGGAAGGAACTGCGAAATGCTTGGT 540
CELd        ACTAGGACATGCACAAACCCTTCTCCAAAATTTGGAGGAAGGAACTGCGAAATGCTTGGT 540
CELa        ACTAGGGCATGCACAAACCCTTCTCCAAAATTTGGAGGAAGGAACTGCGAAATGCTTGGT 540
CELb        ACTAGGGCATGCACAAACCCTTCTCCAAAATTTGGAGGAAGGAACTGCGAAATGCTTGGT 540
            ****** *****************************************************

CELc        CCAGCAACATTTACAGAAGCATGCAATTCTGAACCATGCCCAATCAATGGTGGGTTCAGC 600
CELd        CCAGCAACATTTACAGAAGCATGCAATTCTGAACCATGCCCAATCAATGGTGGGTTCAGC 600
CELa        CCAGCAACATTTACAGAAGCATGCAATTCTGAACCATGCCCAATCAATGGTGGGTTCAGC 600
CELb        CCAGCAACATTTACAGAAGCATGCAATTCTGAACCATGCCCAATCAATGGTGGGTTCAGC 600
            ************************************************************

CELc        AAGTGGGGAAAGTATGGCGTTTGCTCTGTTAAATGTGGAGGTGGAATGCAGTCACGAAGA 660
CELd        AAGTGGGGAAAGTATGGCGTTTGCTCTGTTAAATGTGGAGGTGGAATGCAGTCACGAAGA 660
CELa        AAGTGGGGAAAGTATGGCGTTTGCTCTGTTAAATGTGGAGGTGGAATGCAGTCACGTAGA 660
CELb        AAGTGGGGAAAGTATGGCGTTTGCTCTGTTAAATGTGGAGGTGGAATGCAGTCACGTAGA 660
            ******************************************************** ***

CELc        AGAACGTGTACAAATCCTTCGCCTCAATATGGAGGATTAGATTGTACACGACTTGGACCT 720
CELd        AGAACGTGTACAAATCCTTCGCCTCAATATGGAGGATTAGATTGTACACGACTTGGACCT 720
CELa        AGAACGTGTACAAATCCTTC---------------------------------------- 680
CELb        AGAACGTGTACAAATCCTTC---------------------------------------- 680
            ********************

CELc        AATTCATCTACACGAGCTTGCAATGAACATGCCTGTCCA--------------------- 759
CELd        AATTCATCTACACGAGCTTGCAATCAACATGCCTGTCCAATCAATGGTGGATTCCGCAAG 780
CELa        ------------------------------------------------------------
CELb        ------------------------------------------------------------

CELc        ------------------------------------------------------------
CELd        TGGGGAAAGTATGGCGTTTGCTCTGTTAAATGTGGAGGTGGAATGCAGTCACGTAGAAGA 840
CELa        ------------------------------------------------------------
CELb        ------------------------------------------------------------

CELc        ------------------------------------------------------------
CELd        ACGTGTACAAATCCTTCGCCTCAATATGGAGGATTAGATTGTACACGACTTGGACCTAAT 900
CELa        ------------------------------------------------------------
CELb        ------------------------------------------------------------

CELc        ------------------------------------ATCAATGGTGGATTCAGCAAGTGG 783
CELd        TCATCTACACGAGCTTGCAATCAACATGCCTGTCCAATCAATGGTGGATTCAGCAAGTGG 960
CELa        ------------------------------------------------------------
CELb        ------------------------------------------------------------

CELc        GGAAAGTATGGGGTTTGCTCTGTTAAATGTGGAAGTGGAATGCAGTTACGAAGAAGAACG 843
CELd        GGAAAGTATGGGGTTTACTCTGATAAATGTGGAGGTGGAATGCAGTTACGAAGAAGAACG 1020
CELa        ------------------------------------------------------------
CELb        ------------------------------------------------------------

CELc        TGTACAAATCCTTCTCCTCAATATGGAGGATTAGATTGTACACGACTTGGACCTAATTCA 903
CELd        TGTACAAATCCTTCTCCTCAATATGGAGGATTAGATTGTACACGACTTGGACCTAATTCA 1080
CELa        --------------TCCTCAACATGGAGGATTATATTGTACACGACTTGGACCTAATTCA 726
CELb        --------------TCCTCAACATGGAGGATTATATTGTACACGACTTGGACCTAATTCA 726
                          ******* *********** **************************

CELc        TCTACACGAGCTTGCAATGAACATGCCTGTCCAATCAATGGTGGATTCAGCAAGTGGGGA 963
CELd        TCTACACGAGCTTGCAATGAACATGCCTGTCCAATCAATGGTGGATTCAGCAAGTGGGGA 1140
CELa        TCTACACGAGCTTGCAATCAACATGCCTGTCCAAT------------------------- 761
CELb        TCTACACGAGCTTGCAATCAACATGCCTGTCCAATCAATGGTGGGTTCAGCAAGTGGCGA 786
            ****************** ****************

CELc        AAGTATGGGGTTTGCTCTGTTAAATGTGGAGGTGGAATGCAGTTACGAAAAAGAACGTGT 1023
CELd        AAGTATGGGGTTTGCTCTGTTAAATGTGGAGGTGGAATGCAGTTACGAAAAAGAACGTGT 1200
CELa        ------------------------------------------------------------
CELb        AAGTATGGCGTTTGCTCTGTTAAATGTGGAGGTGGAATGCAGTCACGAAGAAGAACGTGT 846

CELc        ACAAATCCTTCTCCTCAATATGGAGGATTAGATTGTACACGACTTGGACCTAGTTCATCT 1083
CELd        ACAAATCCTTCTCCTCAGTATGGAGGATTAGATTGTACACGACTTGGACCTAGTTCATCT 1260
CELa        ------------------------------------------------------------
CELb        ACAAATCCTTCTCCTCAATATGGAGGATTAGATTGTACACGACATGGACCTAATTCATCT 906

CELc        ACACGTGCTTGCAATCAACATGCCTGTCCAATTTTGCTAAAGAGAGTATGCCAAGATAAA 1143
CELd        ACACGTGCTTGCAATCAACATGCCTGTCCAATTTTGCTAAAGAGAGTATGCCAAGATAAA 1320
CELa        --------------------------------TTTGCTAAAGAGAGTATGCCAAGATAAA 789
CELb        ACACGCTCTTGCAATCAACATGCCTGTCCAATTTTGCTAAAGAGAGTATGCCAAGATAAA 966
                                            ****************************

CELc        ACTCTACAAATCAAATGTGCACCCGGTTACATAGTGAGAATTCATGCCGCACTGTACGGC 1203
CELd        ACTCTACAAATCAAATGTGCACCCGGTTACATAGTGAGAATTCATGCCGCACTGTACGGC 1380
CELa        ACTCTACAAATCAAATGTGCACCCGGTTACATAGTAAGAATTCATGATGCACTGTACGGC 849
CELb        ACTCTACAAATCAAATGTGCACCCGGTTACATAGTAAGAATTCATGATGCACTGTACGGC 1026
            *********************************** **********  ************

CELc        CGCATTTCAAAAGGATATTGCGGATGGAGATTCTACTACAGCAAGAGGTGCAGATCTCAT 1263
CELd        CGCATTTCAAAAGGATATTGCGGATGGAGATTCTACTACAGCAAGAGGTGCAGATCTCAT 1440
CELa        CGCATGTCAAAAGCATATTGCGGATGGAGATTCTACTACAGCAAGGGGTGCAGATCTCAT 909
CELb        CGCATGTCAAAAGCATATTGCGGATGGAGATTCTACTACAGCAAGGGGTGCAGATCTCAT 1086
            ***** ******* ******************************* **************

CELc        TCTTCAGTTTCTCATACAGCGAATATGTGCAATAACAAGAGAACATGTTCTGTTTCGGCA 1323
CELd        TCTTCAGTTTCTCATACAGCGAATATGTGCAATAACAAGAGAACATGTTCTGTTTCGGCA 1500
CELa        TCTTCAGTTTCTCAAACAGCGAATATGTGCAATAACAAGAGAACATGTTCTGTTTCGGCA 969
CELb        TCTTCAGTTTCTCAAACAGCGAATATGTGCAATAACAAGAGAACATGTTCTGTTTCGGCA 1146
            ************** *********************************************

CELc        ACAAATAAAGTGTATGGCAACCCGTGTTTTGGAACTCGTAAATATCTAGAAGTCAAATAC 1383
CELd        ACAAATAAAGTGTATGGCAACCCGTGTTTTGGAACTCGTAAATATCTAGAAGTCAAATAC 1560
CELa        ACAAATAAAGTGTATGGCAACCCGTGTGTTGGAACTCATAAATACCTAGAAGTCAAGTAC 1029
CELb        ACAAATAAAGTGTATGGCAACCCGTGTGTTGGAACTCATAAATACCTAGAAGTCAAGTAC 1206
            *************************** ********* ****** *********** ***
```

```
CELc          TCCTGCGAGCCCAGGACTACTGTTCTTCAACTCAAACGAGTATGCCAGGACAAAACCATG 1443
CELd          TCCTGCGAGCCCAGGACTACTGTTCTTCAACTCAAACGAGTATGCCAGGACAAAACCATG 1620
CELa          TCCTGCGAGCCCAGGACTACTGTTCTTCAACTCAAACGAGTATGCCAGGACAAAACCATG 1089
CELb          TCCTGCGAGCCCAGGACTACTGTTCTTCAACTCAAACGAGTATGCCAGGACAAAACCATG 1266
              ************************************************************

CELc          ACAATCAACTGTGGAAAAAAGAGAATTCATATCAAGAGTGCAAGTTATGGTCGCACTTCA 1503
CELd          ACAATCAACTGTGGAAAAAAGAGAATTCATATCAAGAGTGCAAGTTATGGTCGCACTTCA 1680
CELa          ACAATCAACTGTGGAAAAAAGAGAATTCATATCAAGAGTGCAAGTTATGGTCGCACTTCG 1149
CELb          ACAATCAACTGTGGAAAAAAGAGAATTCATATCAAGAGTGCAAGTTATGGTCGCACTTCG 1326
              ***********************************************************

CELc          AGAAGATACTGCGGTTACAAAATCTTTCATCAGTCAAACTGTCACGCCAAATCATCCTCG 1563
CELd          AGAAGATACTGCGGTTACAAAATCTTTCATCAGTCAAACTGTCACGCCAAATCATCCTCG 1740
CELa          AGAAGATACTGCGGTTACAAAATCTTTCATCAGTCAAACTGTCACGCCAAATCATCCTCG 1209
CELb          AGAAGATACTGCGGTTACAAAATCTTTCATCAGTCAAACTGTCACGCCAAATCATCCTCG 1386
              ************************************************************

CELc          CGTGTTGTGACGTCACGATGTGAAAATAAGAAACGATGCATTGTGAAAGCATCAAACAAA 1623
CELd          CGTGTTGTGACGTCACGATGTGAAAATAAGAAACGATGCATTGTGAAAGCATCAAACAAA 1800
CELa          CGTGTTGTGACGTCACGATGTGAAAATAAGAAACGATGCATTGTGAAAGCATCAAACAAA 1269
CELb          CGTGTTGTGACGTCACGATGTGAAAATAAGAAACGATGCATTGTGAAAGCATCAAACAAA 1446
              ************************************************************

CELc          GTTTTTGGTAATCCATGCTTTCTCACGTCAAAGTACTTGGAAGTGTCGTATACTTGCAAA 1683
CELd          GTTTTTGGTAATCCATGCTTTCTCACGTCAAAGTACTTGGAAGTGTCGTATACTTGCAAA 1860
CELa          GTTTTTGGTAATCCATGCTTTCTCACGTCGAAGTACTTGGAAGTGTCGTATACTTGCAAA 1329
CELb          GTTTTTGGTAATCCATGCTTTCTCACGTCGAAGTACTTGGAAGTGTCGTATACTTGCAAA 1506
              ***************************** ******************************

CELc          TAACATGTGGAT 1695
CELd          TAACATGTGGAT 1872
CELa          TAACATGTGGAT 1341
CELb          TAACATGTGGAT 1518
              ************
```

B

```
CELd   MVNLRLSFVVLFIVFMYTNSGFAKSSKRVCENKTMRISCSSGRVIQIRRAMYGRYTTSYCGWNIFQGRCNSHSSRSLHEVKLRCQNKRSCSVNAVNGVFGDPCFGTRKYLEVDYYCRNLPV
CELc   MVNLRLSFVVLFIVFMYTNSGFAKSSKRVCENKTMRISCSSGRVIQIRRAMYGRYTTSYCGWNIFQGRCNSHSSRSLHEVKLRCQNKRSCSVNAVNGVFGDPCFGTRKYLEVDYYCRNLPV
CELb   MVNLRLSFVVLFIVFMYTNSGFAKSSKRVCENKTMRISCSSGRVIQIRRAMYGRYTTSYCGWNIFQGRCNSHSSRSLHEVKLRCQNKRSCSVNAVNGVFGDPCFGTRKYLEVDYYCRYLPV
CELa   MVNLRLSFVVLFIVFMYTNSGFAKSSKRVCENKTMRISCSSGRVIQIRRAMYGRYTTSYCGWNIFQGRCNSHSSRSLHEVKLRCQNKRSCSVNAVNGVFGDPCFGTRKYLEVDYYCRYLPV

CELd   DGGYTKWNIQRCSRTCGGGTKTFTRTCTNPSPKFGGRNCEMLGPATFTEACNSEPCPI
CELc   DGGYTKWNIQRCSRTCGGGTKTFTRTCTNPSPKFGGRNCEMLGPATFTEACNSEPCPI
CELb   DGGYTKWNIHRCSRSCGGGTKTFTRACTNPSPKFGGRNCEMLGPATFTEACNSEPCPI
CELa   DGGYTKWNIHRCSRSCGGGTKTFTRACTNPSPKFGGRNCEMLGPATFTEACNSEPCPI

CELd   NGGFSKWGKYGVCSVKCGGGMQSRRRTCTNPSPQYGGLDCTRLGPNSSTRACNQHACPI
CELc   NGGFSKWGKYGVCSVKCGGGMQSRRRTCTNPSPQYGGLDCTRLGPNSSTRACNEHACPI
CELb   NGGFSKWGKYGVCSVKCGGGMQSRRRTCTNPSPQHGGLYCTRLGPNSSTRACNQHACPI
CELa   NGGFSKWGKYGVCSVKCGGGMQSRRRTCTNPSPQHGGLYCTRLGPNSSTRACNQHACP


CELd   NGGFRKWGKYGVCSVKCGGGMQSRRRTCTNPSPQYGGLDCTRLGPNSSTRACNQHACPI
CELc   NGGFSKWGKYGVCSVKCGGGMQLRKRTCTNPSPQYGGLDCTRLGPSSSTRACNQHACPI
CELb   NGGFSKWRKYGVCSVKCGGGMQSRRRTCTNPSPQYGGLDCTRHGPNSSTRSCNQHACP
CELa   ----------------------------------------------------------

CELd   NGGFSKWGKYGVYSDKCGGGMQLRRRTCTNPSPQYGGLDCTRLGPNSSTRACNEHACPI
CELc   NGGFSKWGKYGVCSVKCGSGMQLRRRTCTNPSPQYGGLDCTRLGPNSSTRACNEHACP
CELb   -----------------------------------------------------------
CELa   -----------------------------------------------------------


CELd   NGGFSKWGKYGVCSVKCGGGMQLRKRTCTNPSPQYGGLDCTRLGPSSSTRACNQHACPILLKR
CELc   ----------------------------------------------------------ILLKR
CELb   ----------------------------------------------------------ILLKR
CELa   ----------------------------------------------------------ILLKR

CELd   VCQDKTLQIKCAPGYIVRIHAALYGRISKGYCGWRFYYSKRCRSHSSVSHTANMCNNKRTCSVSATNKVYGNPCFGTRKYLEVKYSCEPRTTVLQLKR
CELc   VCQDKTLQIKCAPGYIVRIHAALYGRISKGYCGWRFYYSKRCRSHSSVSHTANMCNNKRTCSVSATNKVYGNPCFGTRKYLEVKYSCEPRTTVLQLKR
CELb   VCQDKTLQIKCAPGYIVRIHDALYGRMSKAYCGWRFYYSKGCRSHSSVSQTANMCNNKRTCSVSATNKVYGNPCVGTHKYLEVKYSCEPRTTVLQLKR
CELa   VCQDKTLQIKCAPGYIVRIHDALYGRMSKAYCGWRFYYSKGCRSHSSVSQTANMCNNKRTCSVSATNKVYGNPCVGTHKYLEVKYSCEPRTTVLQLKR

CELd   VCQDKTMTINCGKKRIHIKSASYGRTSRRYCGYKIFHQSNCHAKSSSRVVTSRCENKKRCIVKASNKVFGNPCFLTSKYLEVSYTCK
CELc   VCQDKTMTINCGKKRIHIKSASYGRTSRRYCGYKIFHQSNCHAKSSSRVVTSRCENKKRCIVKASNKVFGNPCFLTSKYLEVSYTCK
CELb   VCQDKTMTINCGKKRIHIKSASYGRTSRRYCGYKIFHQSNCHAKSSSRVVTSRCENKKRCIVKASNKVFGNPCFLTSKYLEVSYTCK
CELa   VCQDKTMTINCGKKRIHIKSASYGRTSRRYCGYKIFHQSNCHAKSSSRVVTSRCENKKRCIVKASNKVFGNPCFLTSKYLEVSYTCK
```

Supplementary Fig. 1. Sequence alignment of the 4 cnidarian egg lectin (CEL) isoforms from Sylt. (A) *Nucleotides.* **(B)** *Protein. SNPs are shaded in the nucleotide sequence and the amino acid exchanges are highlighted in yellow in the protein sequence. Protein domains are marked in blue (signal peptide), red (lectin) and green (TSP).*

	CELd TSP1	CELd TSP2	CELd TSP3	CELd TSP4	CELd TSP5	CELc TSP1	CELc TSP2	CELc TSP3	CELc TSP4	CELb TSP1	CELb TSP2	CELb TSP3	CELa TSP1	CELa TSP2
CELd TSP1		43	43	60	54	93	48	59	54	91	43	41	92	43
CELd TSP2	43		97	95	94	54	98	96	95	54	97	96	54	97
CELd TSP3	43	97		95	94	53	97	96	95	55	97	95	55	97
CELd TSP4	60	95	95		96	59	96	98	96	59	94	94	59	94
CELd TSP5	54	94	94	96		54	94	96	99	54	94	94	54	94
CELc TSP1	93	54	53	59	54		59	59	54	95	53	53	95	53
CELc TSP2	48	98	97	96	94	59		97	95	59	97	96	59	97
CELc TSP3	59	96	96	98	96	59	97		97	59	95	94	59	95
CELc TSP4	54	95	95	96	99	54	95	97		54	94	94	54	94
CELb TSP1	91	54	55	59	54	95	59	59	54		55	53	100	55
CELb TSP2	43	97	97	94	94	53	97	95	94	55		96	55	100
CELb TSP3	41	96	95	94	94	53	96	94	94	53	96		53	96
CELa TSP1	92	54	55	59	54	95	59	59	54	100	55	53		55
CELa TSP2	43	97	97	94	94	53	97	95	94	55	100	96	55	

Supplementary Table 1. Similarity (in %) between the thrombospondin (TSP) repeats of different CEL isoforms..