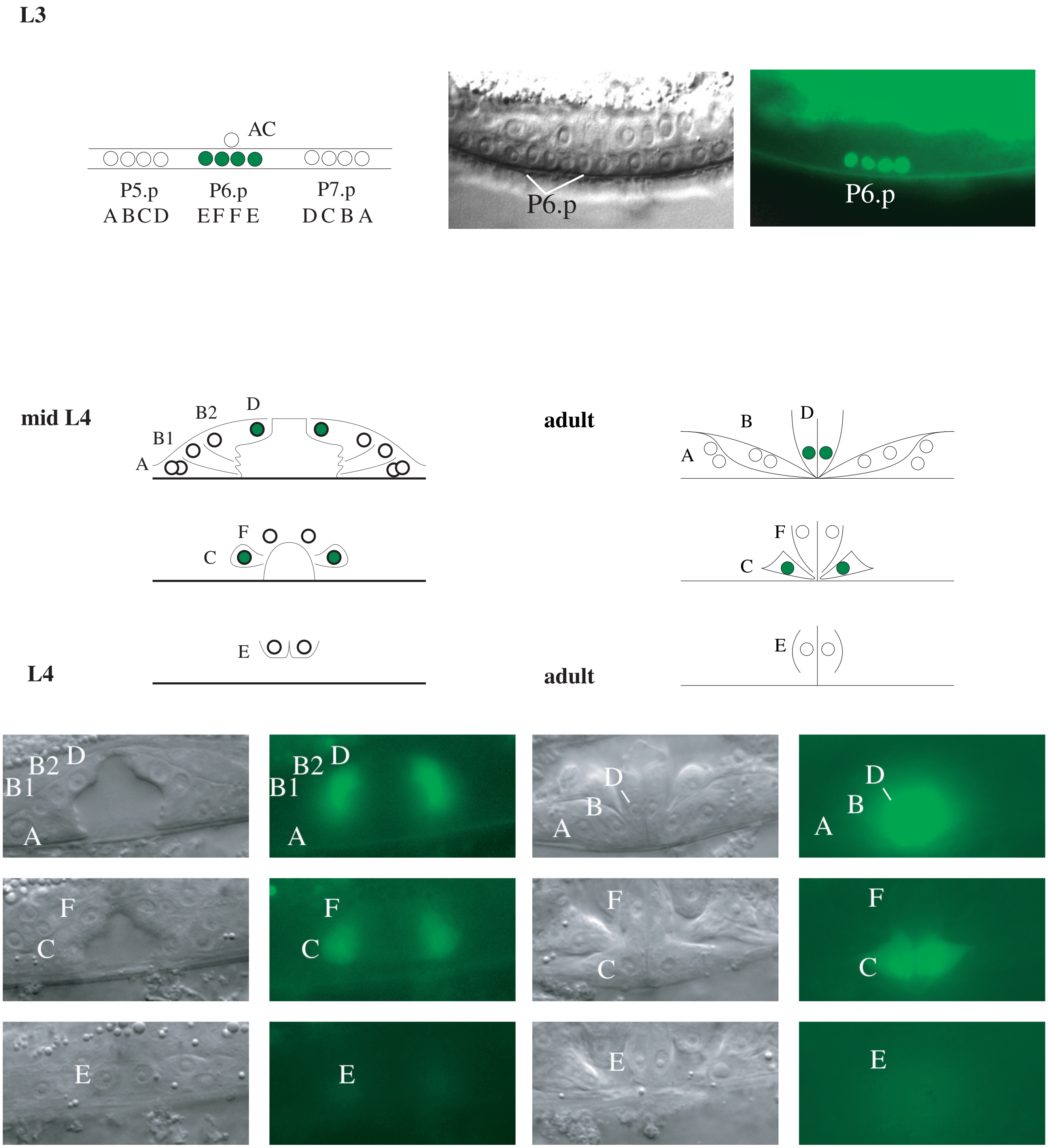
Figure 2: Available vulva marker genes expression pattern in *C. elegans*

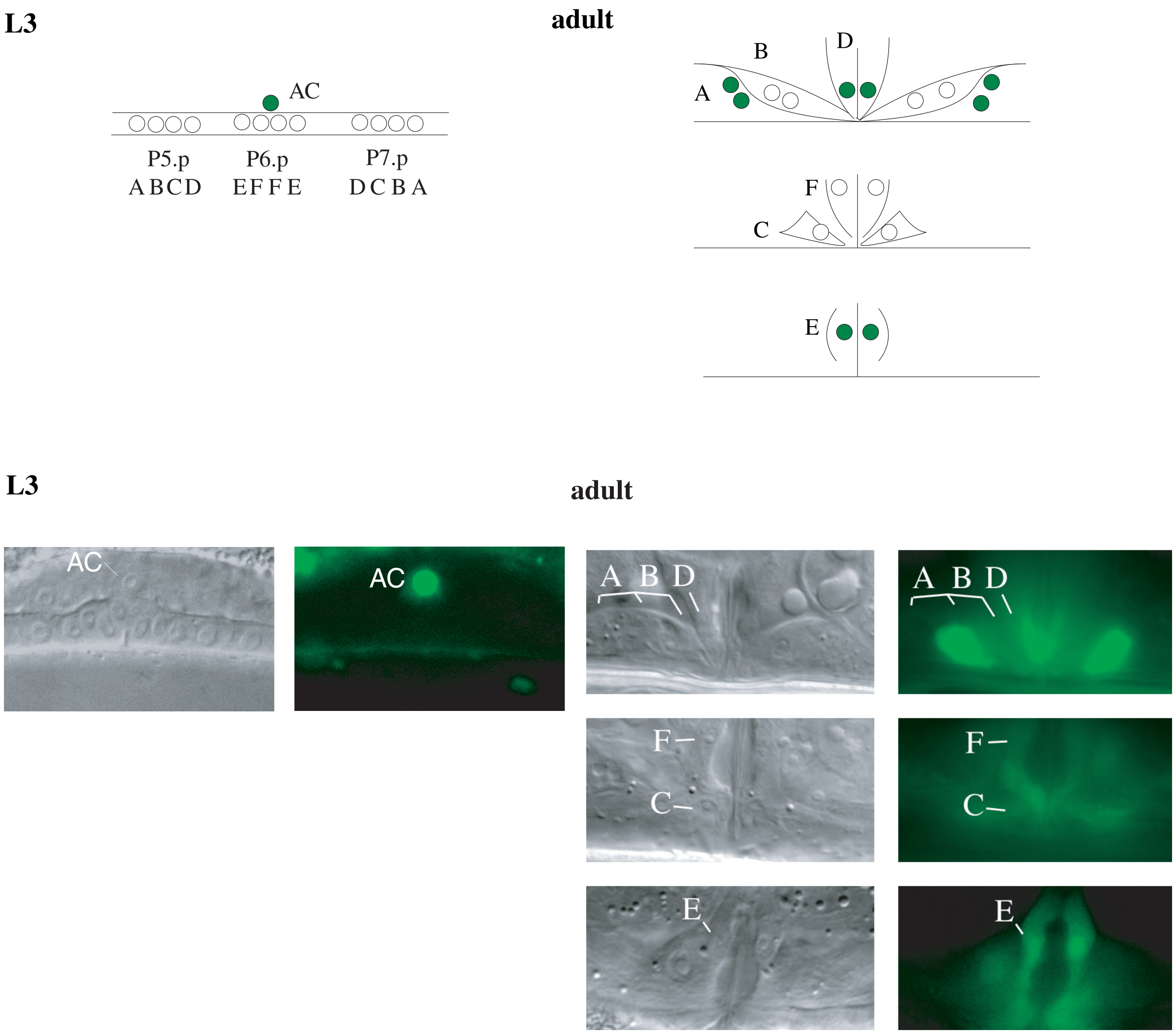
from Inoue *et al.*, submitted

Figure 3: *egl-17::gfp*



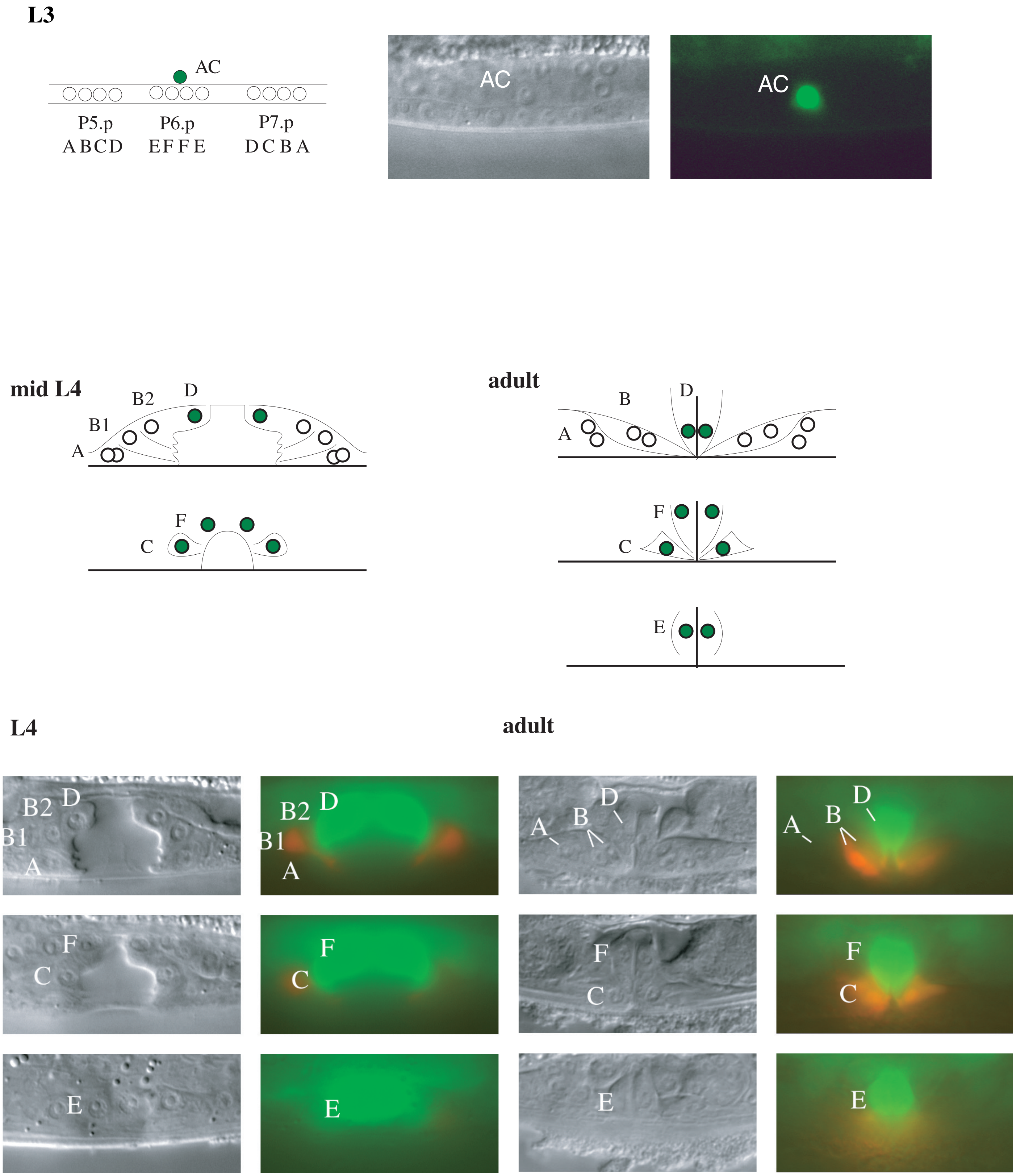
(Adapted from Inoue *et al.*,submitted)

Figure 4: *zmp-1::gfp*



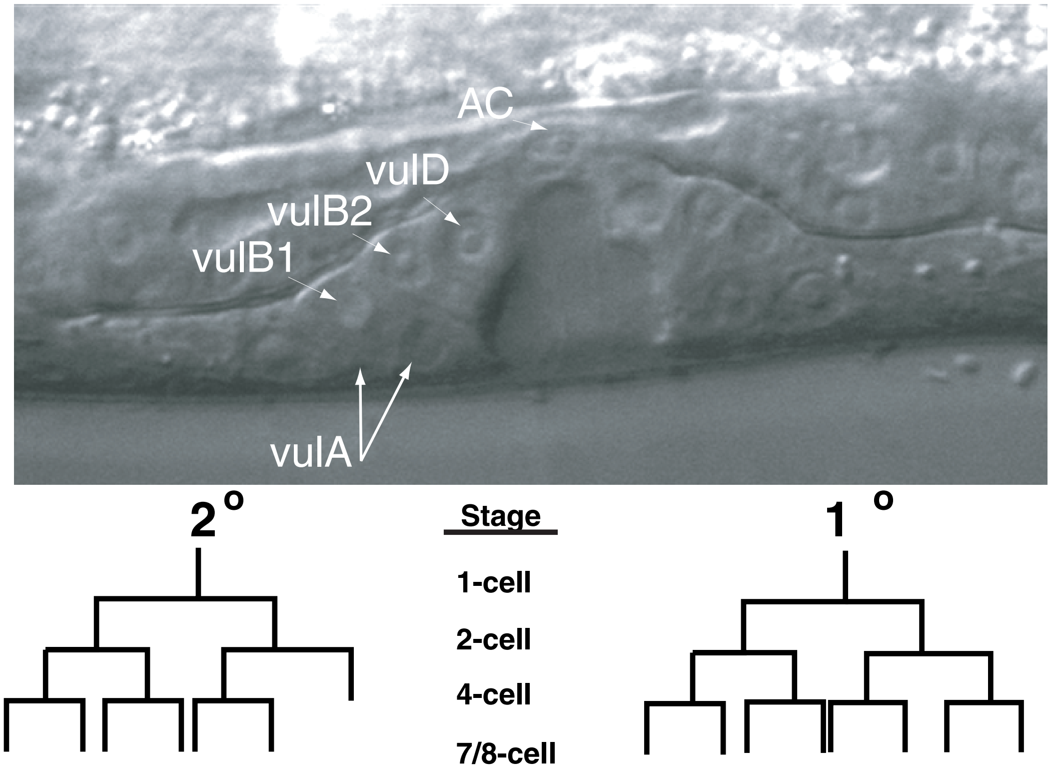
(Adapted from Inoue *et al.*,submitted)

Figure 5: *cdh-3::gfp*



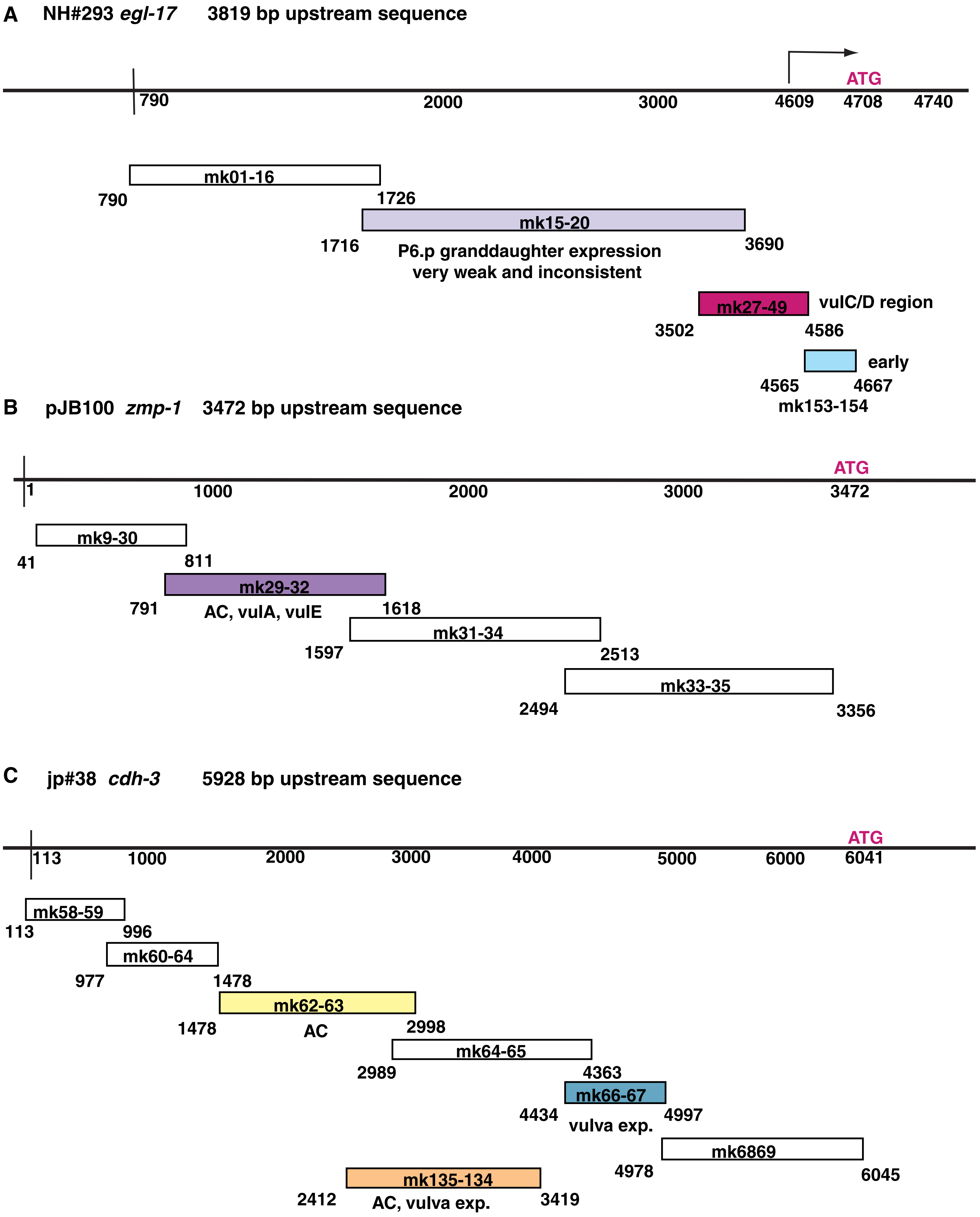
(Adapted from Inoue *et al.*,submitted)

Figure 1: Marker gene expression summary



	vulA	vulB	vulC	vulD	vulE	vulF	vulF	vulE	AC
cdh-3			L4	L4	L3/L4	L3/L4	L3/L4	L3/L4	L3
egl-17			L4/YA	L4/YA	L3	L3	L3	L3	
zmp-1			YA/A		L4/YA			L4/YA	L3

Figure 2: Initial dissection of *egl-17*, *zmp-1*, and *cdh-3* regulatory regions



A NH#293 *egl-17* 3819 bp upstream sequence

B pJB100 *zmp-1* 3472 bp upstream sequence

C jp#38 *cdh-3* 5928 bp upstream sequence

Figure 3: Upstream regions that direct *egl-17* expression

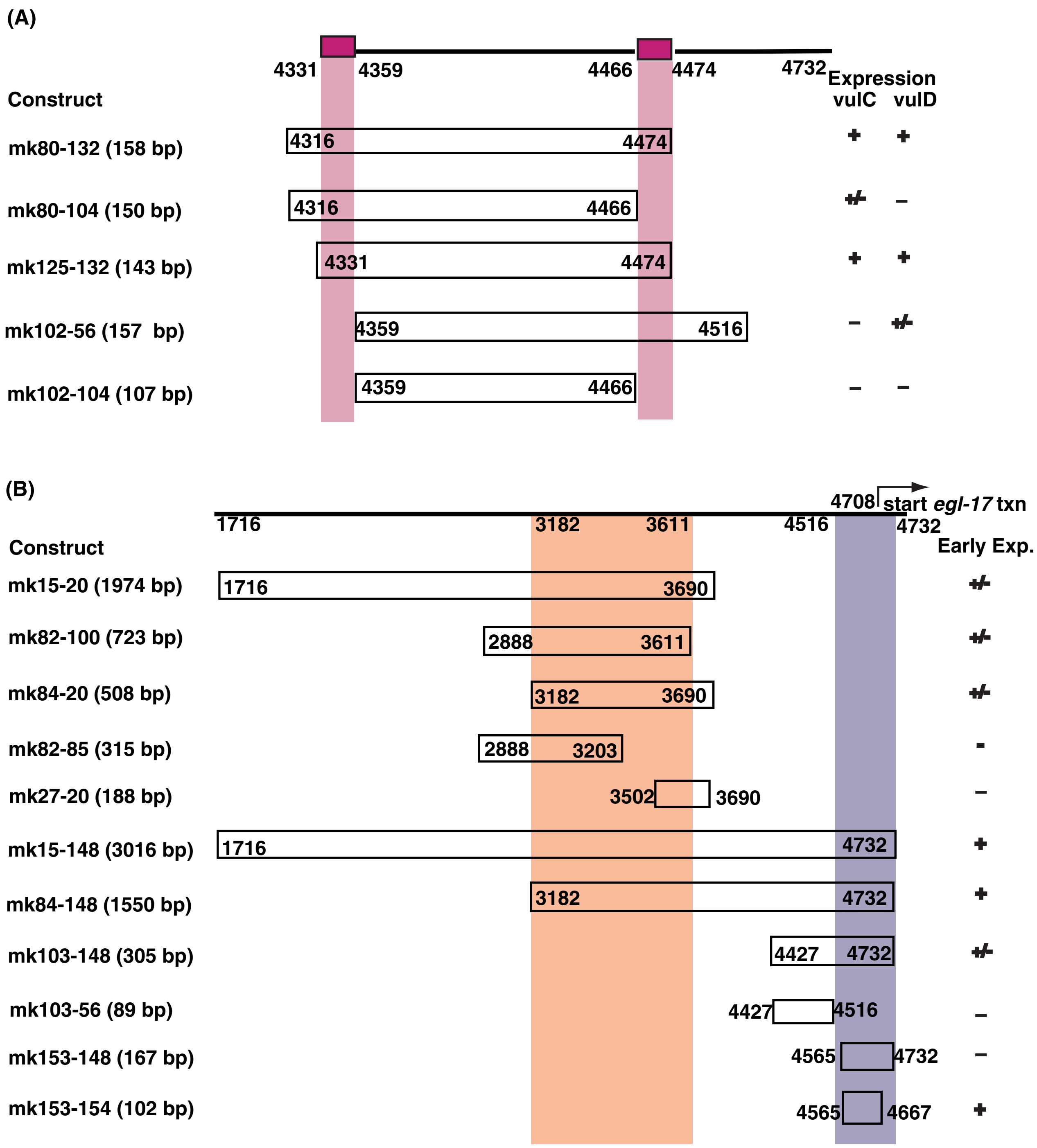
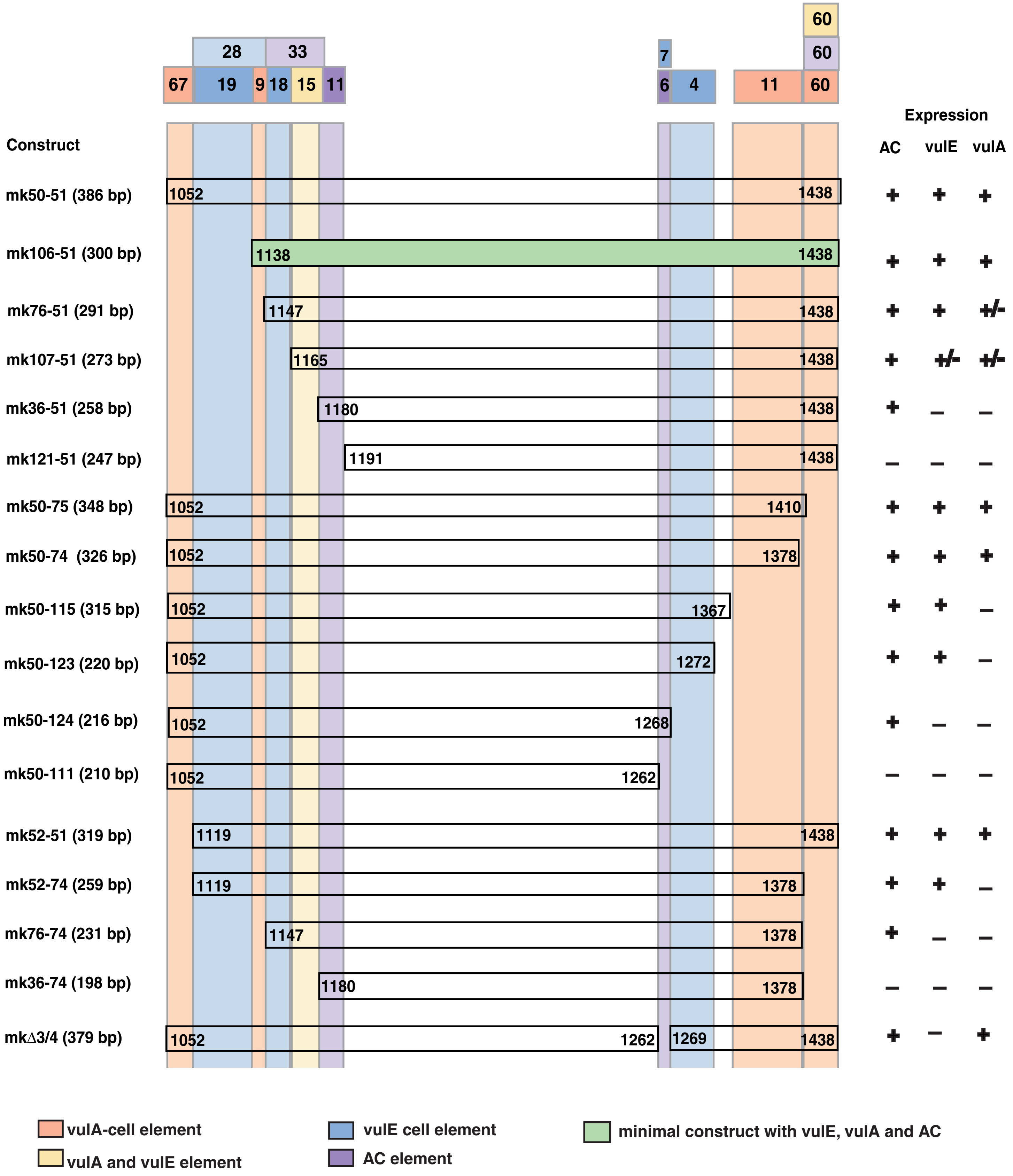


Figure 5: Multiple Regions Direct *zmp-1* expression



Construct			Expression		
			AC	vulE	vulA
mk50-51 (386 bp)	1052	1438	+	+	+
mk106-51 (300 bp)	1138	1438	+	+	+
mk76-51 (291 bp)	1147	1438	+	+	+/-
mk107-51 (273 bp)	1165	1438	+	+/-	+/-
mk36-51 (258 bp)	1180	1438	+	–	–
mk121-51 (247 bp)	1191	1438	–	–	–
mk50-75 (348 bp)	1052	1410	+	+	+
mk50-74 (326 bp)	1052	1378	+	+	+
mk50-115 (315 bp)	1052	1367	+	+	–
mk50-123 (220 bp)	1052	1272	+	+	–
mk50-124 (216 bp)	1052	1268	+	–	–
mk50-111 (210 bp)	1052	1262	–	–	–
mk52-51 (319 bp)	1119	1438	+	+	+
mk52-74 (259 bp)	1119	1378	+	+	–
mk76-74 (231 bp)	1147	1378	+	–	–
mk36-74 (198 bp)	1180	1378	–	–	–
mkΔ3/4 (379 bp)	1052 1262	1269 1438	+	–	+

🟧 vulA-cell element
🟨 vulA and vulE element
🟦 vulE cell element
🟪 AC element
🟩 minimal construct with vulE, vulA and AC

Figure 6: Regions that direct *cdh-3* Expression

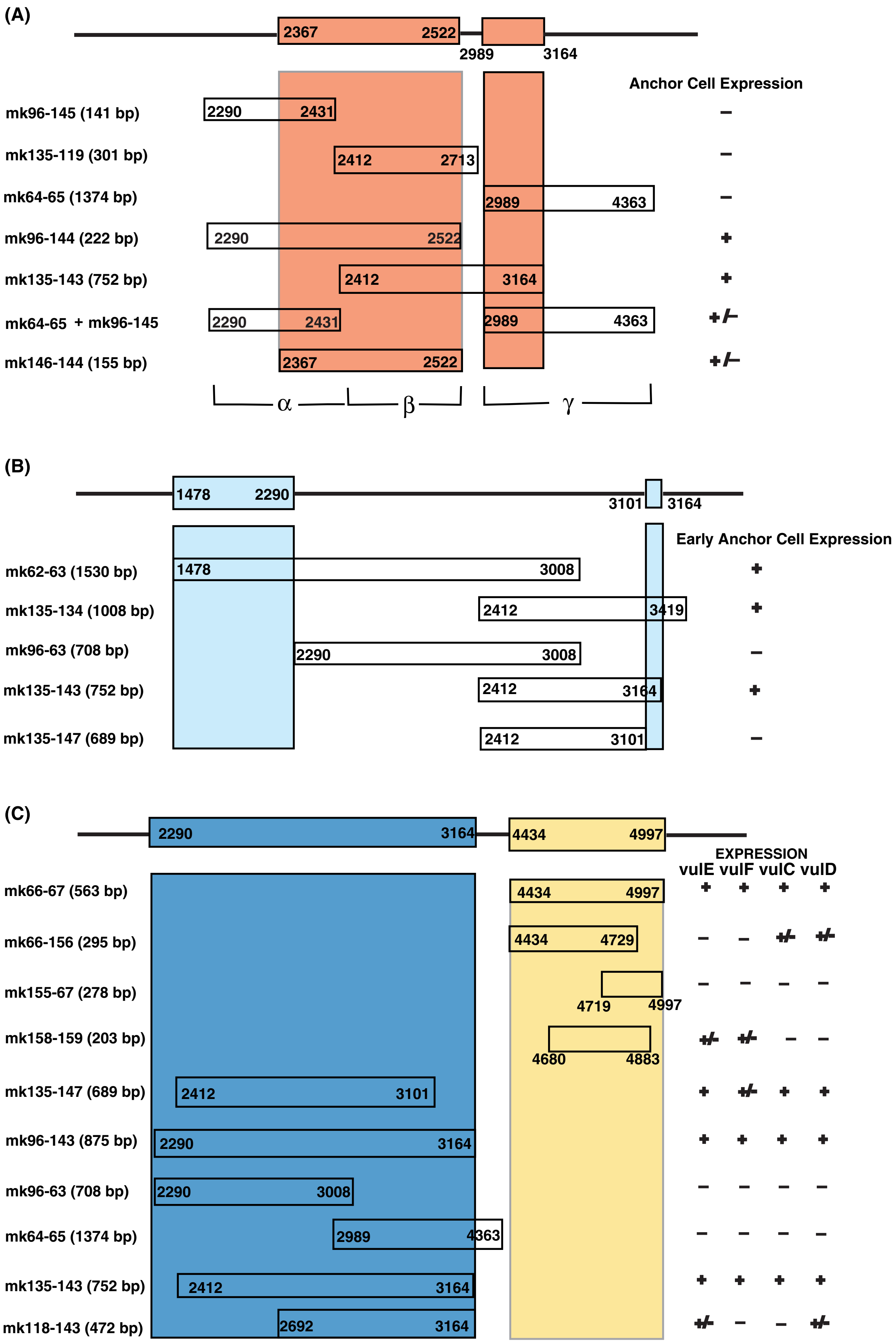
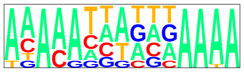
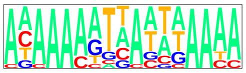


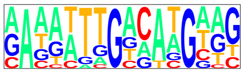
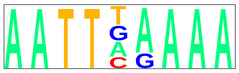
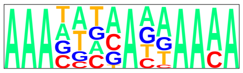
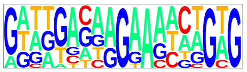
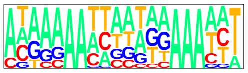
Table 2: AlignACE predictions of overrepresented sequences

A.

Expression	Regions examined	Gene	8 bp motif	10 bp motif
	mk84-148	*egl-17*	5 (14)	3 (6)
	mk50-51	*zmp-1*	1 (5)	0 (2)
	mk96-134	*cdh-3*	2 (12)	1 (8)
	mk66-67	*cdh-3*	0 (5)	0 (7)
Vulval cdh-3	mk96-134/mk66-67	*cdh-3*	0 (8)	0 (12)
Anchor cell	mk50-51/96-134	*zmp-1/cdh-3*	2 (10)	4 (11)
Vulval general	*all	*all	5 (12)	8 (15)

B.

Region	Motif	MAP	Consensus	Sites	Comments
mk84-148 *egl-17*	1.8	17.15		124, 269, 395, 419, 433, 455, 510, 581, 646, 667, 683, 708, 876, 917, 1111, 1158, 1232, 1248, 1327	Sites 1158 & 1232 are located between mk125 & mk102 and mk102 & mk103 regions respectively imp. for vulC/D expr.
	2.8	12.38		161, 217, 277, 295, 433, 453, 647, 667,874, 917, 974, 1110, 1157,1413, 1502	Site 1502 is located between mk154 & mk148 region important fidelity early expression.
	3.8	11.88		124, 394, 405, 471, 486, 510, 594, 618, 633, 686, 712, 818, 876, 1252, 1355	Overlaps with multiple motif 1.8 sites.
	4.8	11.08		74, 160, 214, 273, 296, 392, 417, 518, 581, 683, 854, 916, 992, 1230, 1336, 1361, 1379	Site 1230 is between mk102 & mk103 region imp. for vulC/D expression. Overlaps with multiple motif 2.8 sites.
	5.8	10.79		176, 190, 365, 374, 392, 453, 475, 631, 653, 733, 742, 1127, 1187, 1380, 1434	Sites 1380, 1434,1187 betw. mk153 & mk154, & several in distal region imp. for early expr. Site 1127 betw. mk102 & mk103 region imp. for vulC/D expr.
	6.10	19.87		212, 267, 296, 392, 417, 518, 579, 681, 706, 816, 854, 1109, 1230, 1325, 1361	Multiple sites overlap motif 1.8 sites.
	7.10	19.69		49, 158, 218, 296, 430, 455, 518, 540, 562, 580, 643, 668, 776, 875, 910, 991, 1064, 1110, 1367	Multplie sites overlap with motif 2.8 sites.
	8.10	13.91		124, 289, 392, 403, 471, 486, 510, 592, 616, 631, 684, 712, 818, 876, 1252	Multiple sites overlap with motif 3.8 sites.
mk50-51 *zmp-1*	1.8	11.30		16, 31, 50, 58, 101, 110, 123, 173, 182, 239, 316, 342	Site 316 between mk74 & mk115, region imp. for vulA expression. Site 110 between mk107 & mk36, region imp. for vulA and vulE expr.

Region	Motif	MAP	Consensus	Sites	Comments
mk96-134 *cdh-3*	1.8	17.17		104, 129, 148, 166, 555, 687, 710, 750, 770, 812, 842, 1112	Multiple sites in alpha, beta, and gamma. Site could be important for vuval expression (few sites between mk136 & mk164).
	2.8	10.62		64, 91, 128, 369, 380, 416, 438, 471, 606, 764, 811, 836, 866, 1035	Multiple sites in alpha, beta, and gamma.
	3.10	16.99		103, 129, 147,166, 687, 709, 749, 770, 812, 846	All sites overlap motif1.8 sites.
mk50-51 *zmp-1* mk96-134 *cdh-3* mk66-67 *cdh-3* mk84-148 *egl-17*	1.8	16.08		[116, 285] [319, 333, 353, 418, 610, 957] [35] [397, 669, 689, 816, 885, 1348, 1487]	
	2.8	16.04		[105, 169, 298] [38, 388, 623, 1056, 1098] [31, 482] [213, 277, 395, 582, 706, 864, 1110, 1232]	
	3.8	12.14		[91] [89, 463, 600, 720, 833, 1026, 1055] [205, 354, 390, 483] [160, 218, 296, 432, 454, 582, 646, 875, 912, 1120, 1158]	
	4.8	11.75		[25, 95] [94,113, 465, 770] [ ] [1, 114, 160, 230, 448, 647, 1115, 1220, 1531]	No sites present in mk66-67
	5.8	10.98		[278, 372] [81, 156, 324, 416, 610] [171, 261, 382] [112, 169, 496, 656, 918, 1113, 1255, 1267, 1516]	
	6.10	21.39		[51, 168, 239, 285] [42, 321, 341, 381, 610, 1081] [35] [49, 212, 255, 296, 390, 518, 576, 678, 1227, 1360]	
	7.10	21.09		[88, 236, 283, 316] [40, 254, 351, 404, 512, 718, 952, 1023] [29, 202] [167, 212, 293, 392, 414, 451, 527, 579, 643, 681, 705, 862, 923, 1108, 1227]	
	8.10	16.85		[31, 173, 360] [38, 120, 180, 263, 384, 472, 619, 710, 1036] [31, 336, 548] [258, 395, 570, 674, 708, 832, 1011, 1232, 1477]	

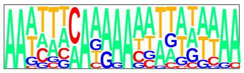
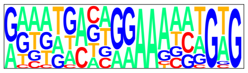
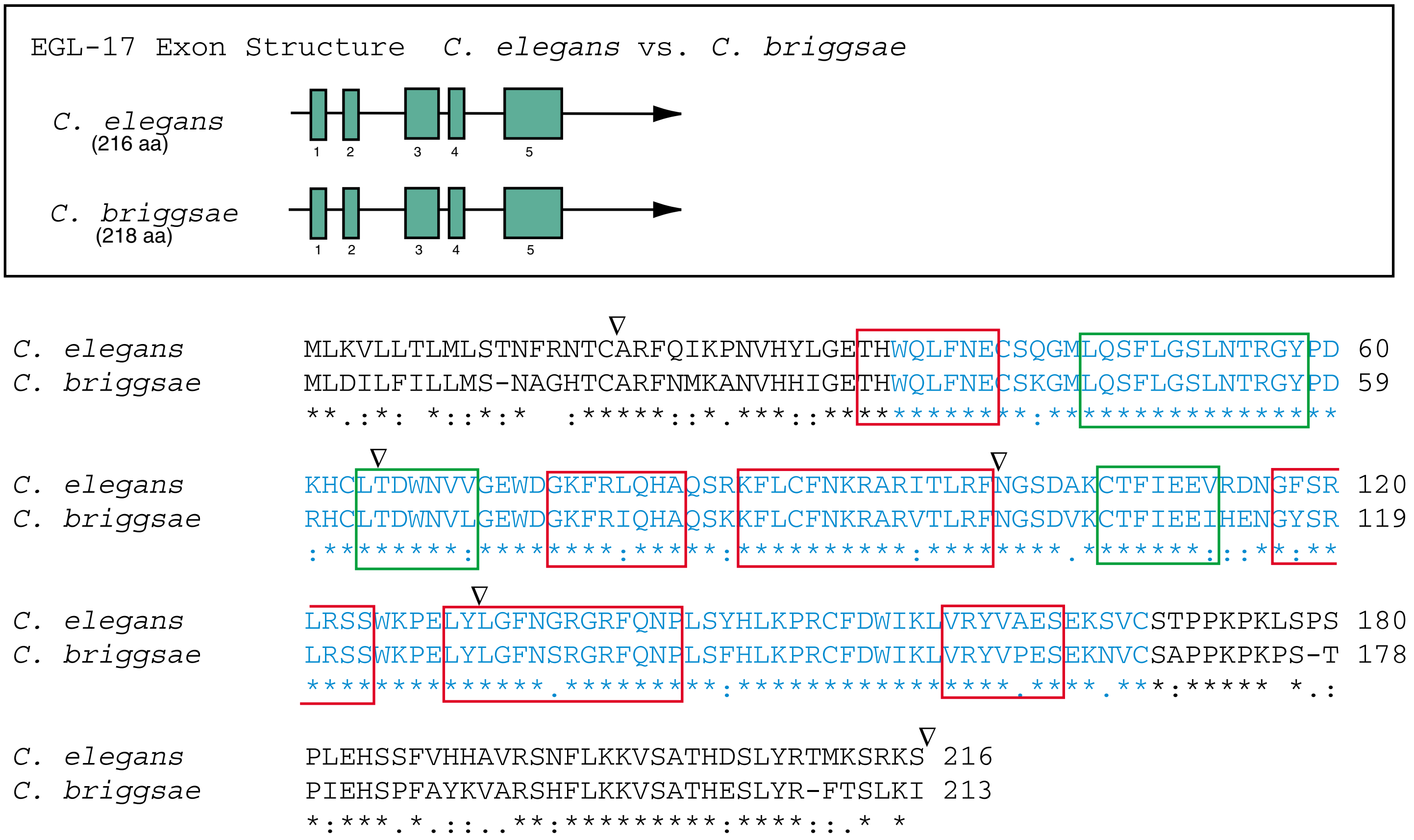
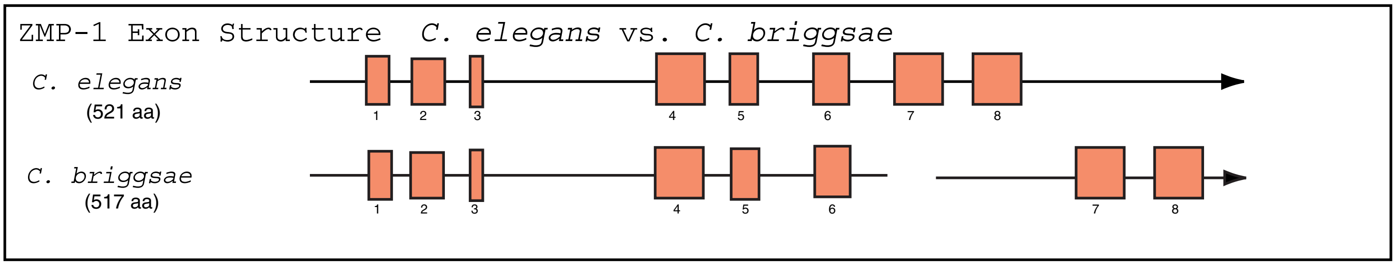
Region	Motif	MAP	Consensus	Sites	Comments
mk50-51 *zmp-1* mk96-134 *cdh-3* mk66-67 *cdh-3* mk84-148 *egl-17*	9.10	15.47		[91, 239, 352] [38, 374, 484, 945,1026] [60] [29, 269, 296, 395, 417, 510, 582, 674, 698, 856, 916, 1234]	
	10.10	15.42		[25, 95, 158] [71, 113, 166, 465, 604, 770] [356, 477] [21, 159, 448, 556, 646, 1115]	
	11.10	15.36		[6] [0, 743, 845] [454, 497, 525] [72, 367, 479, 622, 735, 1291, 1467, 1514]	
	12.10	15.29		[25, 280, 320] [81, 324, 416, 593, 612, 979] [157, 263] [114, 448, 656, 702, 920, 1115, 1516]	
	13.10	11.55		[275] [15, 63, 100, 685, 765, 840, 897] [195] [5, 452, 647, 1110, 1127, 1500]	
mk50-51 *zmp-1* mk96-134 *cdh-3*	1.8	11.60		[126] [65, 105, 129, 149, 166, 687, 751, 770, 812]	Site 126 between mk36 & mk121 imp. for ac expression in mk50-51. Mulitple sites in mk96-134 regions alpha, beta and gamma.
	2.8	10.76		[91, 239, 286] [51, 148, 238, 257, 302, 333, 354, 374, 948, 964]	Site 91 between mk76 & mk36 imp. for ac expression in mk50-51. Multiple sites in mk96-134 regions alpha, beta, and gamma.
	3.10	15.10		[53, 105, 173, 286, 359] [47, 133, 180, 268, 330, 354, 383, 618, 816, 920, 1040, 1097]	Site 105 between mk76 & mk36 imp. for ac expression in mk50-51. Multiple sites in mk96-134 regions alpha, beta, and gamma.
	4.10	10.63		[54, 112, 173] [120, 339, 370, 384, 417, 472, 619, 846, 867, 952, 1036]	Site 112 between mk76 & mk36 imp. for ac expression in mk50-51. Multiple sites in mk96-134 regions alpha, beta, and gamma.
	5.10	10.63		[126] [65, 100, 129, 166, 687, 812]	Multiple sites overlap with motif 1.8 sites. See 1.8 comments.
	6.10	9.92		[50, 101, 168, 184, 239, 286] [37, 72, 263, 333, 374, 948, 964, 1027, 1055]	Site 101 between mk76 & mk36 imp. for ac expression in mk50-51. Multiple sites in mk96-134 regions alpha, beta, and gamma.

Figure 1: EGL-17 clustalW alignment *C. elegans* and *C. briggsae*



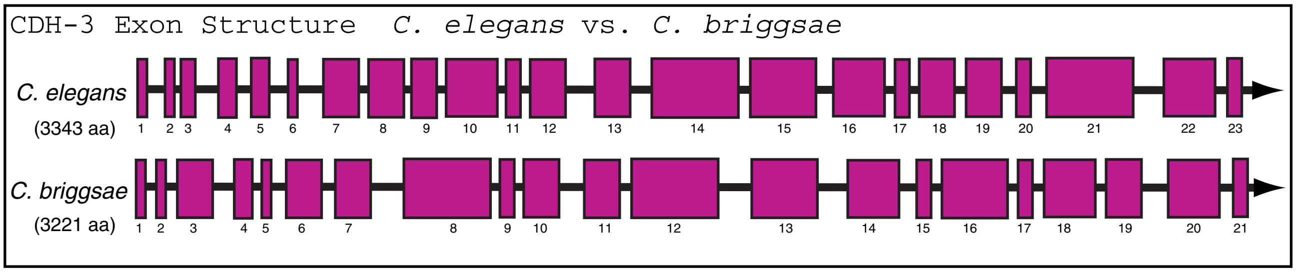
```
                                 ∇
C. elegans      MLKVLLTLMLSTNFRNTCARFQIKPNVHYLGETHWQLFNECSQGMLQSFLGSLNTRGYPD 60
C. briggsae     MLDILFILLMS-NAGHTCARFNMKANVHHIGETHWQLFNECSKGMLQSFLGSLNTRGYPD 59
                **.:*: *::*:*   :*****::*.***::*********:*****************

                   ∇                                    ∇
C. elegans      KHCLTDWNVVGEWDGKFRLQHAQSRKFLCFNKRARITLRFNGSDAKCTFIEEVRDNGFSR 120
C. briggsae     RHCLTDWNVLGEWDGKFRIQHAQSKKFLCFNKRARVTLRFNGSDVKCTFIEEIHENGYSR 119
                :********:********:*****:**********:********.*******:::**:**

                         ∇
C. elegans      LRSSWKPELYLGFNGRGRFQNPLSYHLKPRCFDWIKLVRYVAESEKSVCSTPPKPKLSPS 180
C. briggsae     LRSSWKPELYLGFNSRGRFQNPLSFHLKPRCFDWIKLVRYVPESEKNVCSAPPKPKPS-T 178
                **************.*********:****************.*****.**:***** *.:

                                                    ∇
C. elegans      PLEHSSFVHHAVRSNFLKKVSATHDSLYRTMKSRKS 216
C. briggsae     PIEHSPFAYKVARSHFLKKVSATHESLYR-FTSLKI 213
                *:***.*.::..**:*********:****::.*  *
```

Figure 2: ZMP-1 clustalW alignment in *C. elegans* and *C. briggsae*



```
                                         ∇
C. elegans      MFTGLHDILIILFLLVTLKIAQNVDHTKFLQKYGYLTSGDNQLSSESLSDALKNMQRMAG 60
C. briggsae     MFH--YDLLVVFLTIIPFLDTRNVDHTEFLQKYGYLPRGSNQLSSTSLSEALKNMQRMAG 58
                **   :*:*:::: ::.:  ::*****:********. *.***** ***:**********

                                            ∇                      ∇
C. elegans      LEETGELDERTIQMMERPRCGHPDVEDHQKSRGKRYAPPQFKWKEKIITYGCKAVGTSTR 120
C. briggsae     LEETGELDERTKRMMERPRCGHPDVQEDRHSRGKRYAQPQFKWTDKVITYGCKSAGTSTR 118
                *********** :*************::.::******* *****.:*:******:.*****

                                                       ∇
C. elegans      ISLDDLRRTMHQAASQWSELADVEIVESSVKNPMMVISAGRENHYPCTVRFDTKTLAHAF 180
C. briggsae     ISLDDLRRTMHQAASQWSELADVEILESSVEKPNIQISAGRSSHYPCNVRFDSQTLAHAF 178
                *************************:****::* : *****..****.****::******

                                         ∇
C. elegans      FPTNGQIHINDRVQFAMTNYTERMGANSLYSVVAHEMGHALGFSHSPDIDSVMFAYDTPR 240
C. briggsae     FPPHGQIHINDNVHFVMTNFTERIGGNSLYSVVAHEMGHALGFSHSPDSDSVMFAYDTPR 238
                **.:*******.*:*.***:***:*.********************** ***********

C. elegans      KWKFTSMDKYNMRSYYGAKASKKENEEEERKTENEDKRRKTEKDRGRTREHESDDIRPNE 300
C. briggsae     EWKFTPMDKYRMEMYYGTKKNPTKPKEDERKKE---ERKKEERGRKTEKEHEKDDIRPNE 295
                :****.****.*. ***:* . .: :*:***.*   :*:* *:.*   :***.*******

                                   ∇
C. elegans      CRVENPIVVQYRGEYLIFKSQWVWRVSSDWKRLIIKAVPINQLFPGLPNPIDAAVTVGHN 360
C. briggsae     CRVDNPIVVQYRGEYLVFKSQWVWRVSSDWKRLIIKPVPIDQLFPGLPNPIDAAVSVGHH 355
                ***:************:*******************.***:**************:***:

                                                           ∇
C. elegans      LWVFVGEMIYVIYGNHMVHAPLRLSDIGINEKYVDLAYEWHYFNPPAVYIWKGSRYWKLD 420
C. briggsae     LWVFVGEKIYVIYGNRMAHAPLTLSEIGIDEKYIDLAYEWHYFNPPAIYIWKGSRYWKLD 415
                ******* *******:*.**** **:***:***:*************:************

C. elegans      EKMYHRRVDERYPKDTDLNWARVPKGVHSAFTYEKEIHLLRGNQVFRMNSSRSVFDIADG 480
C. briggsae     EKMDHRRVDERYPKAIDLNWARVPKGVHSAFTYEKEIHFTRGDQVFRMNSSRSVFDVADG 475
                *** **********  **********************: **:*************:***

                                                           ∇
C. elegans      YPQPLQSFFGFCPRNEKLVLNSSSSHFSLIYATITILILIF-- 521
C. briggsae     YPQPLQSFFGFCPRNEKLVLN-ISPTFHPITALLTVLILNFYL 517
                *********************  *. *  * * :*:*** *
```

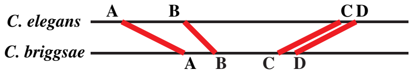
Figure 3: CDH-3 clustalW alignment in *C. elegans* and *C. briggsae*



```
                                                     ∇
C. elegans      MTIRIFFSIFLLNHLIFFHLFNFTHQFSEETIKFSVSEDAKLNTIIGHLEAEIGYTYRLS 60
C. briggsae     ----MLIRHHLFLVFLLTLIFKFSRQFSEETVKFSIAEDAPIDTIIGHLTPENGYSYRLS 56
                 :  :::  .*:  :::  :*:*::******:***::*** ::****** .* **:****

                             ∇
C. elegans      RGNSKIKFDEQTLELSVSSPLDRESENAIDMLIITSPPSIIHILIDVLDVNDNSPIFPID 120
C. briggsae     RGNSKIKFDEETLEFSVSSALDRESENAIDMLIVSSPPSIIHVLIDILDINDNPPKFPLE 116
                **********:***:****.*************::*******:***:**:***.* **::

                 ∇'
C. elegans      VQRVEIPETAPIGWRVQISGATDPDEGKNGTIGKYELVDSLATVDTMSP--FGIVQSDGF 178
C. briggsae     IQNVEIPETAPIGWRVPISGATDPDQGKNGSIGKYELDEITVDGDTPTPPLFRLLQSDGF 176
                :*.************* ********:****:****** :  .  ** :*  * ::*****

                                               ∇
C. elegans      LFLEVTGKLDRETRDLYSMRLTAIDQGVPELSSSCHLNILILDINDNPPNFGIRSLTLNW 238
C. briggsae     VYLEVVGTLDREMRDFYSMRLTASDEGVPELSASCLLNIRILDINDNPPDFGIRQIHLKW 236
                ::***.*.**** **:******* *:******:** *** *********:****.: *:*

                                                                ∇
C. elegans      NGLPNTKLFSLNATDLDSNENSLLTYRILPSGPTSEMFSISDENILVTQNNTECLQRCEF 298
C. briggsae     NGHKNAKLFLLNATDADSGDNGILKYRIQGGEEIFGILEEKDGRFLVTKNSTKCSPICEF 296
                **  *:*** ***** **.:*.:*.***  .     ::. .* .:***:*.*:*   ***

                                                  ∇
C. elegans      VVEARDSGVPPLSTTLNIVVNMEYGNEHEPNINIRFYPSDYPFIIVQPEDVNGKTLAILS 358
C. briggsae     VIEAKDSGIPPLATTLNVVVQMEYGNEHEPNINIRFYPSDFPFIVVQPEDVNGKTLAILS 356
                *:**:***:***:****:**:*******************:***:***************

C. elegans      ITDSDGPLGANSTIWIENGNEQSIFSLISRQSINILTVKHVENANQEQYILEFRANDGQS 418
C. briggsae     LTDPDGPLGPSSKIWIDSGNDQSIFSLISRQSINILTLKNVELAEKEEYTLEFAANDGQG 416
                :**.*****..*.***:.**:****************:*:** *::*:* *** *****.

C. elegans      PADRITRKELKIFFKKYVKSTQIHVERESHVTVEKDTVPGSFVAHVETNCTDMCSFELAN 478
C. briggsae     PEEKICRRSLKIFFKKFVKSTVIQVAREIHVELERDTVAGSFVAHVETNCSEMCRFELQQ 476
                * ::* *:.*******:**** *:* ** ** :*:***.***********::** *** :

                           ∇
C. elegans      -SDVFKIDPFNGIIVTSSILPEGVTSYHLPIRIHLPPPSTQLVEADVFVKVIQES-VPKN 536
C. briggsae     SSDVFRIDGMNGIIVTSSELPSDVSSYHLPIRIHPPPPSTQILETDVFVKILQASSVPKN 536
                 ****:** :******** **..*:********* ******::*:*****::* * ****

C. elegans      LIRSSESPIHLKRAYTFTTWQDVSLGTVIGRLPKAQIYSTIDTVSELGVFPDGSVFVGKT 596
C. briggsae     LIRSTDPPVHLKRAYTFDTWQNVTVGTVVGKLPKAQIYSTRDLESELGVFPDGSIFVGKS 596
                ****::.*:******** ***:*::***:*:********* *  **********:****:

                              ∇
C. elegans      ITSDFVTLPVTLVNRNTTQTSIITLIVKPLNQHSPICQITEIHVLENAPIGTIFGRIQAR 656
C. briggsae     ISGDLVRLEVILENRNTTQTAVVTVLLKPLNRHSPVCADTKVRVLESEKPGNFIGKLHAT 656
                *:.*:* * * * *******:::*:::****:***:*  *:::***.   *.::*:::*
```

Figure 4: Seqcomp and Family Relations predictions for *egl-17, zmp-1* and *cdh-3*

(A) *egl-17*



egl-17 element A

```
C. elegans   TGCCTCGCCTCATCGAATTATTGGATGAGGTCAGTTATGGC
                     |||||| || |||||||||||||||||||
C. briggsae  CCAGCAAACTCATCAAACTATTGGATGAGGTCAGTTAGTAT
```

egl-17 element B

```
C. elegans   CTCCTTTTCTTTCACGTCCTTGGTAATTTTCATATGTATGTT
                 | ||| ||||||||||||||||||||||||||||
C. briggsae  TTCTTCTTCCTTCACGTCCTTGGTAATTTTCATATGTAGAGG
```

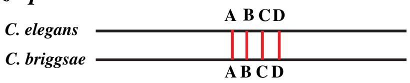
egl-17 element C

```
C. elegans   CATCTTAATATGATGTCCAGTCAATAGTTTTCCTCAG
                  ||  ||||||  |||||||||||| | ||
C. briggsae  ATGAGTATAATGATGCGCAGTCAATAGTTATTCTTTC
```

egl-17 element D

```
C. elegans   ACAATTGTCCGACAACTTCAAGTTGTAATTACAATGTGTTTTGAAAGAAAAAAGT
             ||||||| |||||  |||||||||||||||||||||||||||||  ||||| |
C. briggsae  ACAATTGCCCGACGTCTTCAAGTTGTAATTACAATGTGTTTTGACGGAAAATAAA

C. elegans   AAAAAAGTGACAAAAAAGTTGATTAAATTCTTGTTTCTGATCTGATTTC
             ||||||||||  |||||||||||||||||||||  ||||||    ||||
C. briggsae  AAAAAAGTGATGAAAAAGTTGATTAAATTCTTGCGTCTGATTCTTTTTC
```

(B) *zmp-1*



zmp-1 element A

```
C. elegans   ATGCGCCCCTCGAGAGAAAGATGTATTTCGTAACCCATTTCAAAAGAAGGACGGCTCGTTGAACAG
               |   ||||||||||| |  \ \\\\\\\\\\\\\\\\\\  \\\     \\\\\\\\\\\\\
C. briggsae  TGGTGTCCCTCGAGAGAGACTCCTTCTATTTCGTAACCCATTTCGTAAGTATCGGCTCGTTGAACG
```

zmp-1 element B

```
C. elegans   AAGTATTCGAGTACGTTTACACTGGTTCTG
                   |||||||||||||||||
C. briggsae  CGAATCTCGAGTACGTTTACACTTTGGTTT
```

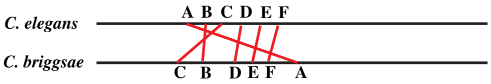
zmp-1 element C

```
C. elegans   AGATGCAAACTGATTCATGTGTACGTATTGCTTGAAAAAAAGA
                    |||||||||||||||  || |\\\\\
C. briggsae  TCTCCAGAACTGATTCATGTGTGTGTGTTTGCTTATTGAAC
```

zmp-1 element D

```
C. elegans   GTAGAAGGGTATTAGTCGTAGTAGTAGTATTCAGT

C. briggsae  GCCAGTTTTACTACTACCAATACTAATACTACCTC
```

(C) *cdh-3* mk96-134 homologous region



cdh-3 element A

```
C. elegans   GTTCAGATTCCCAAACAGAAAAAAAAACAAATAAAAAGGCAC

C. briggsae  TGGCTCCATTTTCTTCTATGTTTTTTTCTCTCTTTTGTTTTC
```

cdh-3 element B

```
C. elegans   TAGCCAATGTTTATGTGTCATGAATAATGAATGGTTTGGA
                 //////// | |||  |||||||||||||| |
C. briggsae  TCTAATGTTTACGTATGTTGTGAATAATGAATGGGTGGTT
```

cdh-3 element C

```
C. elegans   TTTCCTGTTAATTTTATTTGACGACAACTTAATGAAAT

C. briggsae  AGTTATCAAAAAGTTGTCGACTACTAAAATGTAGTTCG
```

cdh-3 element D

```
C. elegans   GAAATTTGAAATGTAGCTACCAAAAAATTGCCTTGTTCTGAA

C. briggsae  TACACATTTTGTTAGTGATCAATTTTTTGTTAGCTAATTTGC
```

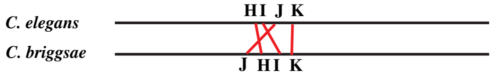
cdh-3 element E

```
C. elegans   TCCCTATACAAAACGGACCGACCGTCCCAAAA
                   ||||||||| ||||  ||||
C. briggsae  CTAACTTACAAAACGTACCGCGCGTCGTGTAT
```

cdh-3 element F

```
C. elegans    CACTACCACCTGCCTTTTGTGTGTTCGTTCGCGCGTGTCCCGCCTGT
                  ///////   |||||||||||||||||||||| ||| /
C. briggsae  CCTAGCCACCTGTCTTTTGTGTGTTCGTTCGCGCGTATCCACCATTC
```

(D) *cdh-3* mk66-67 homologous region



cdh-3 element H

```
C. elegans   GATTCTCTTGTATATCCAATTTCGAGATTGTCCTTCACACCACACAGTGCCAATTGTCTTTCC
               | ||||  ||||||||||||\\\\\\\\\\\\ |   | | |||||||||||||||||||
C. briggsae  GTTCCTCTCATATATCCAATTTTCGAGATTGTCCTCCTCCAAAACAGTGCCAATTGTCTTTCG
```

cdh-3 element I

```
C. elegans   CGAGATTGTCCTTCACACCACACAGTGCCA
              |    ||| ||||||||| |||||
C. briggsae  GGCCTGTGTACTTCACACCTCACAGCCAAT
```

cdh-3 element J

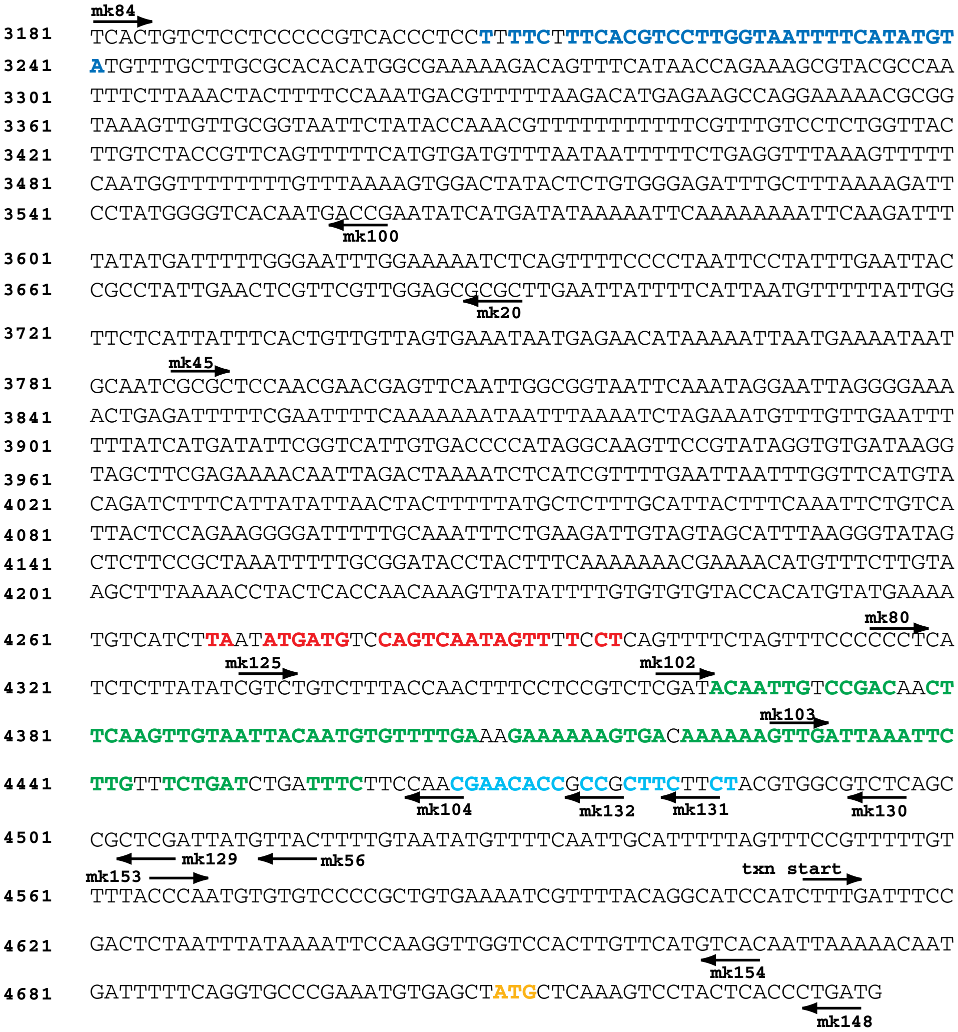
```
C. elegans   GGGCGGTCTCTTTCTGTTCCTCTCATAGTTCACACCTTTT
                    ||||| ||||||||||||||  || //  ///
C. briggsae  CCTCTTCCTCTTCCTGTTCCTCTCATATATCCAATTTTCG
```

cdh-3 element K

```
C. elegans   CCAATCCAATATGTCCTTTTGATATGCTAATTTGCATTCTCTGTCCGCG
                  |||||||||||||||||||||||||||| | || ||
C. briggsae  TTCCGCCAATATGTCCTTTTGATATGCTAATTTTCTTTTTCCTCTTCTT
```

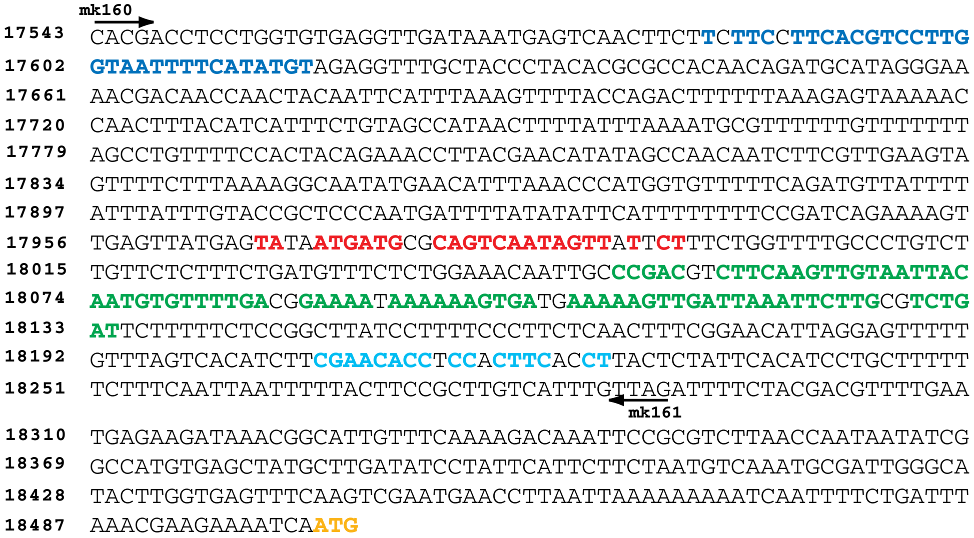
Figure 5: *egl-17* nucleotide sequences of important regions

(A) *C. elegans egl-17* mk84-148



3181 TCACTGTCTCCTCCCCCGTCACCCTCC**TTTTCTTTCACGTCCTTGGTAATTTTCATATGT**
3241 **A**TGTTTGCTTGCGCACACATGGCGAAAAAGACAGTTTCATAACCAGAAAGCGTACGCCAA
3301 TTTCTTAAACTACTTTTCCAAATGACGTTTTTAAGACATGAGAAGCCAGGAAAAACGCGG
3361 TAAAGTTGTTGCGGTAATTCTATACCAAACGTTTTTTTTTTTCGTTTGTCCTCTGGTTAC
3421 TTGTCTACCGTTCAGTTTTTCATGTGATGTTTAATAATTTTTCTGAGGTTTAAAGTTTTT
3481 CAATGGTTTTTTTTGTTTAAAAGTGGACTATACTCTGTGGGAGATTTGCTTTAAAAGATT
3541 CCTATGGGGTCACAATGACCGAATATCATGATATAAAAATTCAAAAAAAATTCAAGATTT
3601 TATATGATTTTTGGGAATTTGGAAAAATCTCAGTTTTCCCCTAATTCCTATTTGAATTAC
3661 CGCCTATTGAACTCGTTCGTTGGAGCGCGCTTGAATTATTTTCATTAATGTTTTTATTGG
3721 TTCTCATTATTTCACTGTTGTTAGTGAAATAATGAGAACATAAAAATTAATGAAAATAAT
3781 GCAATCGCGCTCCAACGAACGAGTTCAATTGGCGGTAATTCAAATAGGAATTAGGGGAAA
3841 ACTGAGATTTTTCGAATTTTCAAAAAAATAATTTAAAATCTAGAAATGTTTGTTGAATTT
3901 TTTATCATGATATTCGGTCATTGTGACCCCATAGGCAAGTTCCGTATAGGTGTGATAAGG
3961 TAGCTTCGAGAAAACAATTAGACTAAAATCTCATCGTTTTGAATTAATTTGGTTCATGTA
4021 CAGATCTTTCATTATATTAACTACTTTTTATGCTCTTTGCATTACTTTCAAATTCTGTCA
4081 TTACTCCAGAAGGGGATTTTTGCAAATTTCTGAAGATTGTAGTAGCATTTAAGGGTATAG
4141 CTCTTCCGCTAAATTTTTGCGGATACCTACTTTCAAAAAAACGAAAACATGTTTCTTGTA
4201 AGCTTTAAAACCTACTCACCAACAAAGTTATATTTTGTGTGTGTACCACATGTATGAAAA
4261 TGTCATCT**TA**AT**ATGATG**TC**CAGTCAATAGTT**T**T**C**CT**CAGTTTTCTAGTTTCCCCCCTCA
4321 TCTCTTATATCGTCTGTCTTTACCAACTTTCCTCCGTCTCGAT**ACAATTG**T**CCGAC**AA**CT**
4381 **TCAAGTTGTAATTACAATGTGTTTTGA**AA**GAAAAAAGTGA**C**AAAAAAGTTGATTAAATTC**
4441 **TTG**TT**TCTGAT**CTGA**TTTC**TTCCAA**CGAACACC**G**CC**G**CTTC**TT**CT**ACGTGGCGTCTCAGC
4501 CGCTCGATTATGTTACTTTTGTAATATGTTTTCAATTGCATTTTTAGTTTCCGTTTTTGT
4561 TTTACCCAATGTGTGTCCCCGCTGTGAAAATCGTTTTACAGGCATCCATCTTTGATTTCC
4621 GACTCTAATTTATAAAATTCCAAGGTTGGTCCACTTGTTCATGTCACAATTAAAAACAAT
4681 GATTTTTCAGGTGCCCGAAATGTGAGCT**ATG**CTCAAAGTCCTACTCACCCTGATG

(B) *C. briggsae egl-17* mk160-translational start site



17543 CACGACCTCCTGGTGTGAGGTTGATAAATGAGTCAACTTCT**TCTTC**C**TTCACGTCCTTG**
17602 **GTAATTTTCATATGT**AGAGGTTTGCTACCCTACACGCGCCACAACAGATGCATAGGGAA
17661 AACGACAACCAACTACAATTCATTTAAAGTTTTACCAGACTTTTTTAAAGAGTAAAAAC
17720 CAACTTTACATCATTTCTGTAGCCATAACTTTTATTTAAAATGCGTTTTTTGTTTTTTT
17779 AGCCTGTTTTCCACTACAGAAACCTTACGAACATATAGCCAACAATCTTCGTTGAAGTA
17834 GTTTTCTTTAAAAGGCAATATGAACATTTAAACCCATGGTGTTTTTCAGATGTTATTTT
17897 ATTTATTTGTACCGCTCCCAATGATTTTATATATTCATTTTTTTTCCGATCAGAAAAGT
17956 TGAGTTATGAG**TA**TA**ATGATG**CG**CAGTCAATAGTT**A**T**T**CT**TTCTGGTTTTGCCCTGTCT
18015 TGTTCTCTTTCTGATGTTTCTCTGGAAACAATTGC**CCGAC**GT**CTTCAAGTTGTAATTAC**
18074 **AATGTGTTTTGA**CG**GAAAA**T**AAAAAAGTGA**TG**AAAAAGTTGATTAAATTCTTG**CG**TCTG**
18133 **AT**TCTTTTTCTCCGGCTTATCCTTTTCCCTTCTCAACTTTCGGAACATTAGGAGTTTTT
18192 GTTTAGTCACATCTT**CGAACACC**T**CC**A**CTTC**AC**CT**TACTCTATTCACATCCTGCTTTTT
18251 TCTTTCAATTAATTTTTACTTCCGCTTGTCATTTGTTAGATTTTCTACGACGTTTTGAA
18310 TGAGAAGATAAACGGCATTGTTTCAAAAGACAAATTCCGCGTCTTAACCAATAATATCG
18369 GCCATGTGAGCTATGCTTGATATCCTATTCATTCTTCTAATGTCAAATGCGATTGGGCA
18428 TACTTGGTGAGTTTCAAGTCGAATGAACCTTAATTAAAAAAAAATCAATTTTCTGATTT
18487 AAACGAAGAAAATCA**ATG**

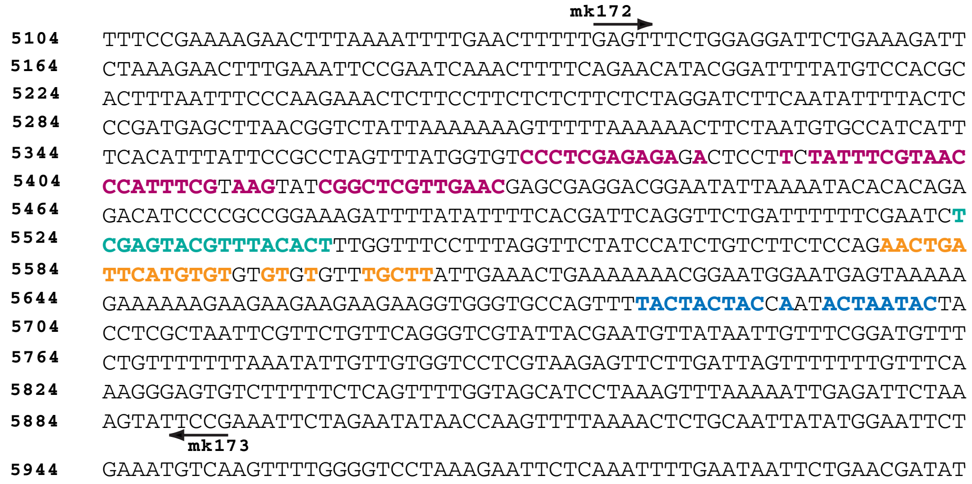
🟦 element B
🟥 element c
🟩 element D
🟦 element E
🟧 translational start site of EGL-17
⟶ indicate boundaries of primers

Figure 6: *zmp-1* nucleotide sequences

(A) *C. elegans zmp-1* mk50-51

992 TTTTTATGTAAGTTTAT**G**CGC**CCCTCGAGAGA**A**A**GA**T**G**TATTTCGTAACCCATTTC**AA**AA**
1052 **G**AAGGA**CGGCTCGTTGAAC**AGAATACACAGATTTCTGTTCCAATTGGAGATTTTTCCTTT
1112 TCTGTATTGATCATCAAAGTAT**TCGAGTACGTTTACACT**GGTTTCTGTTCTTTCCGTTTT
1172 TAATTTCTCCTGCCAGATGCA**AACTGATTCATGTGT**AC**GT**A**TTGCTT**GAAAAAAAGAGTA
1232 ACAAGAAAAAGTAGAAGG**GTATTAGT**CG**TAGTAGTAGTA**TTCAGTTGTAGTAATATATAT
1292 TTCTACTAATTTGTTTAGTTTCGCCACTTAAGATGGTCATCGCAATTTTCAATTAATTTT
1352 TTGGTGGACTTTTCAGAAGAGAAAACGTCGAAATATTTTATGAATGGAAAATGTGACAGT
1412 TTTTTTTGCATATTGGCCATTTTCTAG

(B) *C. briggsae zmp-1* mk172-173

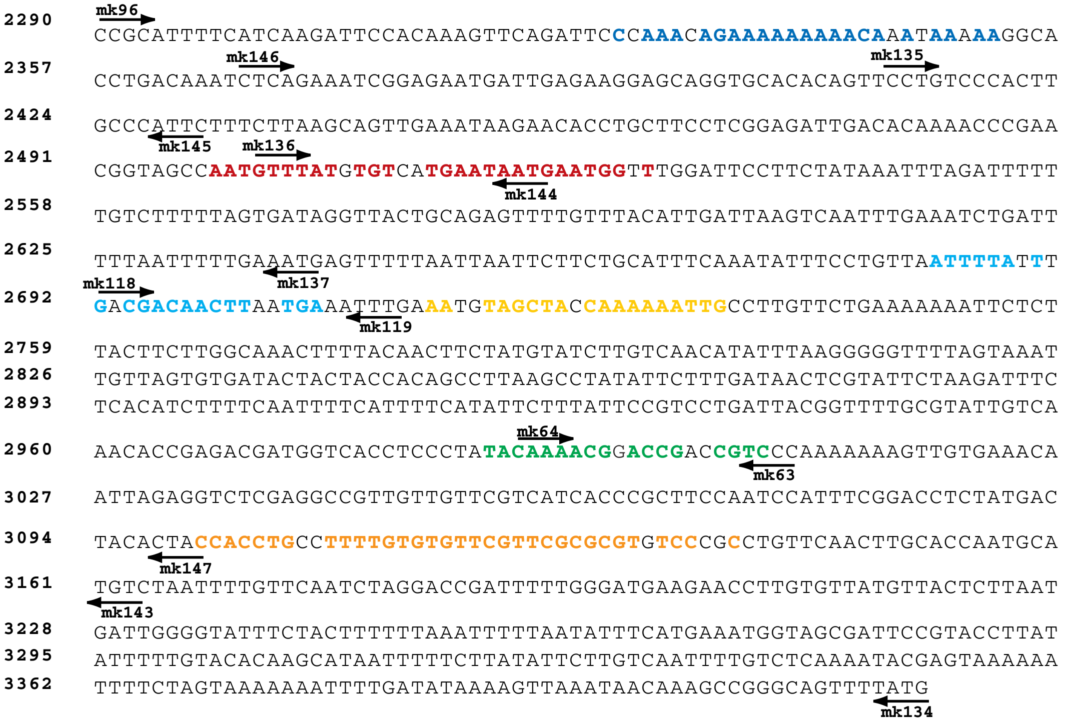


5104 TTTCCGAAAAGAACTTTAAAATTTTGAACTTTTTGAGTTTCTGGAGGATTCTGAAAGATT
5164 CTAAAGAACTTTGAAATTCCGAATCAAACTTTTCAGAACATACGGATTTTATGTCCACGC
5224 ACTTTAATTTCCCAAGAAACTCTTCCTTCTCTCTTCTCTAGGATCTTCAATATTTTACTC
5284 CCGATGAGCTTAACGGTCTATTAAAAAAAGTTTTTAAAAAACTTCTAATGTGCCATCATT
5344 TCACATTTATTCCGCCTAGTTTATGGTGT**CCCTCGAGAGA**G**A**CTCCT**T**C**TATTTCGTAAC**
5404 **CCATTTCG**T**AAG**TAT**CGGCTCGTTGAAC**GAGCGAGGACGGAATATTAAAATACACACAGA
5464 GACATCCCCGCCGGAAAGATTTTATATTTTCACGATTCAGGTTCTGATTTTTTCGAATC**T**
5524 **CGAGTACGTTTACACT**TTGGTTTCCTTTAGGTTCTATCCATCTGTCTTCTCCAG**AACTGA**
5584 **TTCATGTGT**GT**GT**G**T**GTT**TGCTT**ATTGAAACTGAAAAAAACGGAATGGAATGAGTAAAAA
5644 GAAAAAAGAAGAAGAAGAAGGTGGGTGCCAGTTT**TACTACTAC**C**A**AT**ACTAATAC**TA
5704 CCTCGCTAATTCGTTCTGTTCAGGGTCGTATTACGAATGTTATAATTGTTTCGGATGTTT
5764 CTGTTTTTTTAAATATTGTTGTGGTCCTCGTAAGAGTTCTTGATTAGTTTTTTTGTTTCA
5824 AAGGGAGTGTCTTTTTCTCAGTTTTGGTAGCATCCTAAAGTTTAAAAATTGAGATTCTAA
5884 AGTATTCCGAAATTCTAGAATATAACCAAGTTTTAAAACTCTGCAATTATATGGAATTCT
5944 GAAATGTCAAGTTTTGGGGTCCTAAAGAATTCTCAAATTTTGAATAATTCTGAACGATAT

🟦 **element A**
🟧 **element B**
🟩 **element C**
🟪 **element D**
→ **indicate primers location**

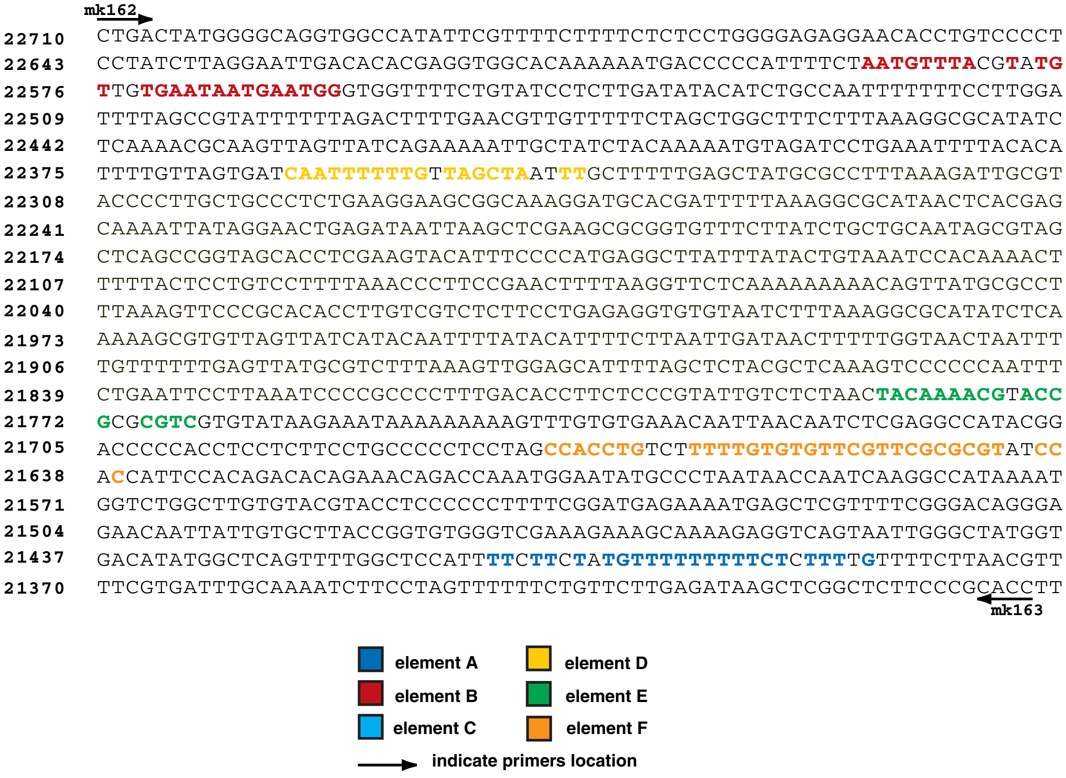
Figure 7: ***cdh-3*** **nucleotide sequences of mk96-134 and mk162-163**

(A) ***C. elegans cdh-3*** **mk96-134**



2290 CCGCATTTTCATCAAGATTCCACAAAGTTCAGATTC**CC****AAA**C**AGAAAAAAAAAACA**A**A**T**AAAAA**GGCA
2357 CCTGACAAATCTCAGAAATCGGAGAATGATTGAGAAGGAGCAGGTGCACACAGTTCCTGTCCCACTT
2424 GCCCATTCTTTCTTAAGCAGTTGAAATAAGAACACCTGCTTCCTCGGAGATTGACACAAAACCCGAA
2491 CGGTAGCC**AATGTTTAT**G**TGT**CA**TGAATAATGAATGG**T**T**TGGATTCCTTCTATAAATTTAGATTTTT
2558 TGTCTTTTTAGTGATAGGTTACTGCAGAGTTTTGTTTACATTGATTAAGTCAATTTGAAATCTGATT
2625 TTTAATTTTTGAAATGAGTTTTTAATTAATTCTTCTGCATTTCAAATATTTCCTGTTA**ATTTTA**T**T**T
2692 **G**A**CGACAACTT**AA**TGA**AATTTGA**AA**TG**TAGCTA**C**CAAAAAATTG**CCTTGTTCTGAAAAAAATTCTCT
2759 TACTTCTTGGCAAACTTTTACAACTTCTATGTATCTTGTCAACATATTTAAGGGGGTTTTAGTAAAT
2826 TGTTAGTGTGATACTACTACCACAGCCTTAAGCCTATATTCTTTGATAACTCGTATTCTAAGATTTC
2893 TCACATCTTTTCAATTTTCATTTTCATATTCTTTATTCCGTCCTGATTACGGTTTTGCGTATTGTCA
2960 AACACCGAGACGATGGTCACCTCCCTA**TACAAAACG**G**ACCG**AC**CGTC**CCAAAAAAAGTTGTGAAACA
3027 ATTAGAGGTCTCGAGGCCGTTGTTGTTCGTCATCACCCGCTTCCAATCCATTTCGGACCTCTATGAC
3094 TACACTA**CCACCTG**CC**TTTTGTGTGTTCGTTCGCGCGT**G**TCC**CG**C**CTGTTCAACTTGCACCAATGCA
3161 TGTCTAATTTTGTTCAATCTAGGACCGATTTTTGGGATGAAGAACCTTGTGTTATGTTACTCTTAAT
3228 GATTGGGGTATTTCTACTTTTTTAAATTTTTAATATTTCATGAAATGGTAGCGATTCCGTACCTTAT
3295 ATTTTTGTACACAAGCATAATTTTTCTTATATTCTTGTCAATTTTGTCTCAAAATACGAGTAAAAAA
3362 TTTTCTAGTAAAAAAATTTTGATATAAAAGTTAAATAACAAAGCCGGGCAGTTTTATG

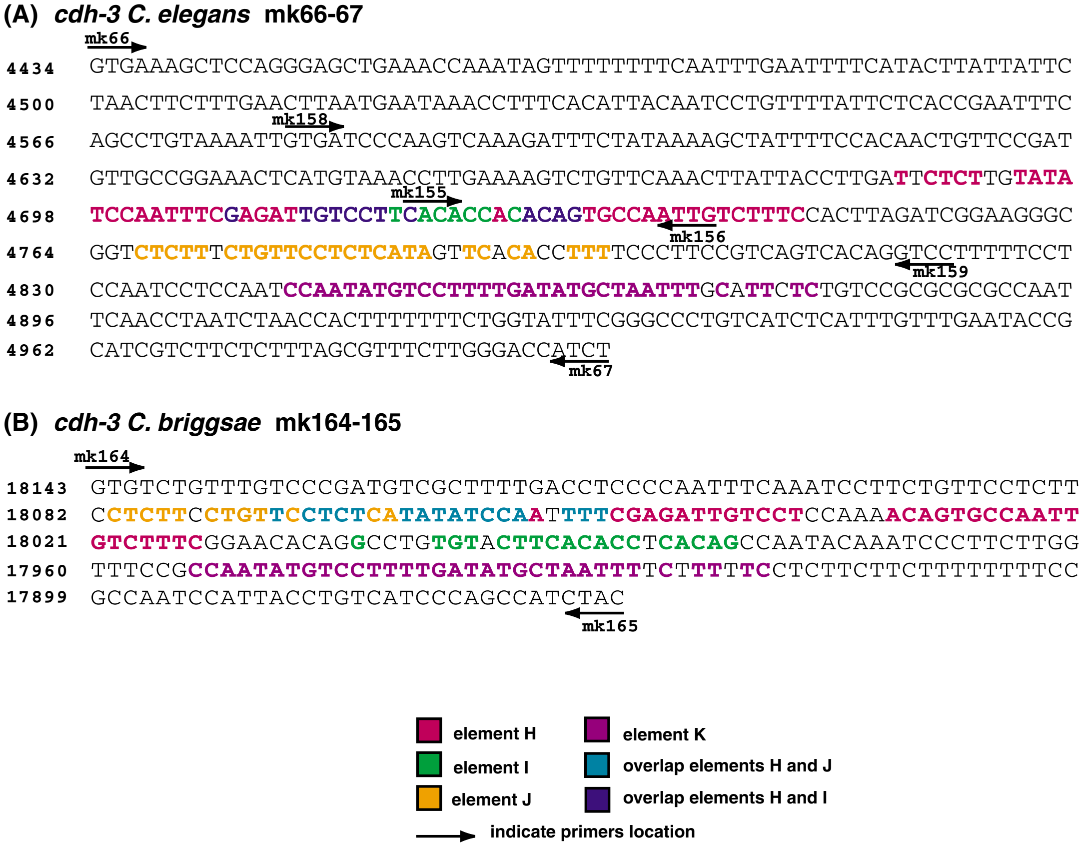
(B) ***C. briggsae cdh-3*** **mk162-163**



22710 CTGACTATGGGGCAGGTGGCCATATTCGTTTTCTTTTCTCTCCTGGGGAGAGGAACACCTGTCCCCT
22643 CCTATCTTAGGAATTGACACACGAGGTGGCACAAAAAATGACCCCCATTTTCT**AATGTTTA**CG**TATG**
22576 **T**TG**TGAATAATGAATGG**GTGGTTTTCTGTATCCTCTTGATATACATCTGCCAATTTTTTTCCTTGGA
22509 TTTTAGCCGTATTTTTTTAGACTTTTGAACGTTGTTTTTCTAGCTGGCTTTCTTTAAAGGCGCATATC
22442 TCAAAACGCAAGTTAGTTATCAGAAAAATTGCTATCTACAAAAATGTAGATCCTGAAATTTTACACA
22375 TTTTGTTAGTGAT**CAATTTTTTG**T**TAGCTA**AT**TT**GCTTTTTGAGCTATGCGCCTTTAAAGATTGCGT
22308 ACCCCTTGCTGCCCTCTGAAGGAAGCGGCAAAGGATGCACGATTTTTAAAGGCGCATAACTCACGAG
22241 CAAAATTATAGGAACTGAGATAATTAAGCTCGAAGCGCGGTGTTTCTTATCTGCTGCAATAGCGTAG
22174 CTCAGCCGGTAGCACCTCGAAGTACATTTCCCCATGAGGCTTATTTATACTGTAAATCCACAAAACT
22107 TTTTACTCCTGTCCTTTTAAACCCTTCCGAACTTTTAAGGTTCTCAAAAAAAAACAGTTATGCGCCT
22040 TTAAAGTTCCCGCACACCTTGTCGTCTCTTCCTGAGAGGTGTGTAATCTTTAAAGGCGCATATCTCA
21973 AAAAGCGTGTTAGTTATCATACAATTTTATACATTTTCTTAATTGATAACTTTTTGGTAACTAATTT
21906 TGTTTTTTGAGTTATGCGTCTTTAAAGTTGGAGCATTTTAGCTCTACGCTCAAAGTCCCCCCAATTT
21839 CTGAATTCCTTAAATCCCCGCCCCTTTGACACCTTCTCCCGTATTGTCTCTAAC**TACAAAACG**T**ACC**
21772 **G**CG**CGTC**GTGTATAAGAAATAAAAAAAAAGTTTGTGTGAAACAATTAACAATCTCGAGGCCATACGG
21705 ACCCCCACCTCCTCTTCCTGCCCCCTCCTAG**CCACCTG**TCT**TTTTGTGTGTTCGTTCGCGCGT**AT**CC**
21638 A**C**CATTCCACAGACACAGAAACAGACCAAATGGAATATGCCCTAATAACCAATCAAGGCCATAAAAT
21571 GGTCTGGCTTGTGTACGTACCTCCCCCCTTTTCGGATGAGAAAATGAGCTCGTTTTCGGGACAGGGA
21504 GAACAATTATTGTGCTTACCGGTGTGGGTCGAAAGAAAGCAAAAGAGGTCAGTAATTGGGCTATGGT
21437 GACATATGGCTCAGTTTTGGCTCCATT**TT**C**TT**C**TATGTTTTTTTTTCT**C**TTTT**T**G**TTTTCTTAACGTT
21370 TTCGTGATTTGCAAAATCTTCCTAGTTTTTTCTGTTCTTGAGATAAGCTCGGCTCTTCCCGCACCTT

🟦 element A
🟥 element B
🟦 element C
🟨 element D
🟩 element E
🟧 element F
→ indicate primers location

Figure 8: ***cdh-3*** **nucleotide sequence mk66-67 and mk164-165**



(A) ***cdh-3 C. elegans*** **mk66-67**

```
4434  GTGAAAGCTCCAGGGAGCTGAAACCAAATAGTTTTTTTTCAATTTGAATTTTCATACTTATTATTC
4500  TAACTTCTTTGAACTTAATGAATAAACCTTTCACATTACAATCCTGTTTTATTCTCACCGAATTTC
4566  AGCCTGTAAAATTGTGATCCCAAGTCAAAGATTTCTATAAAAGCTATTTTCCACAACTGTTCCGAT
4632  GTTGCCGGAAACTCATGTAAACCTTGAAAAGTCTGTTCAAACTTATTACCTTGATTCTCTTGTATA
4698  TCCAATTTCGAGATTGTCCTTCACACCACACAGTGCCAATTGTCTTTCCACTTAGATCGGAAGGGC
4764  GGTCTCTTTCTGTTCCTCTCATAGTTCACACCTTTTCCCTTCCGTCAGTCACAGGTCCTTTTTCCT
4830  CCAATCCTCCAATCCAATATGTCCTTTTGATATGCTAATTTGCATTCTCTGTCCGCGCGCGCCAAT
4896  TCAACCTAATCTAACCACTTTTTTTCTGGTATTTCGGGCCCTGTCATCTCATTTGTTTGAATACCG
4962  CATCGTCTTCTCTTTAGCGTTTCTTGGGACCATCT
```

(B) ***cdh-3 C. briggsae*** **mk164-165**

```
18143  GTGTCTGTTTGTCCCGATGTCGCTTTTGACCTCCCCAATTTCAAATCCTTCTGTTCCTCTT
18082  CCTCTTCCTGTTCCTCTCATATATCCAATTTTCGAGATTGTCCTCCAAAACAGTGCCAATT
18021  GTCTTTCGGAACACAGGCCTGTGTACTTCACACCTCACAGCCAATACAAATCCCTTCTTGG
17960  TTTCCGCCAATATGTCCTTTTGATATGCTAATTTTCTTTTTCCTCTTCTTCTTTTTTTTTCC
17899  GCCAATCCATTACCTGTCATCCCAGCCATCTAC
```

🟥 element H

🟩 element I

🟧 element J

🟪 element K

🟦 overlap elements H and J

🟪 overlap elements H and I

→ indicate primers location

Figure 9: *C. briggsae* upstream regions

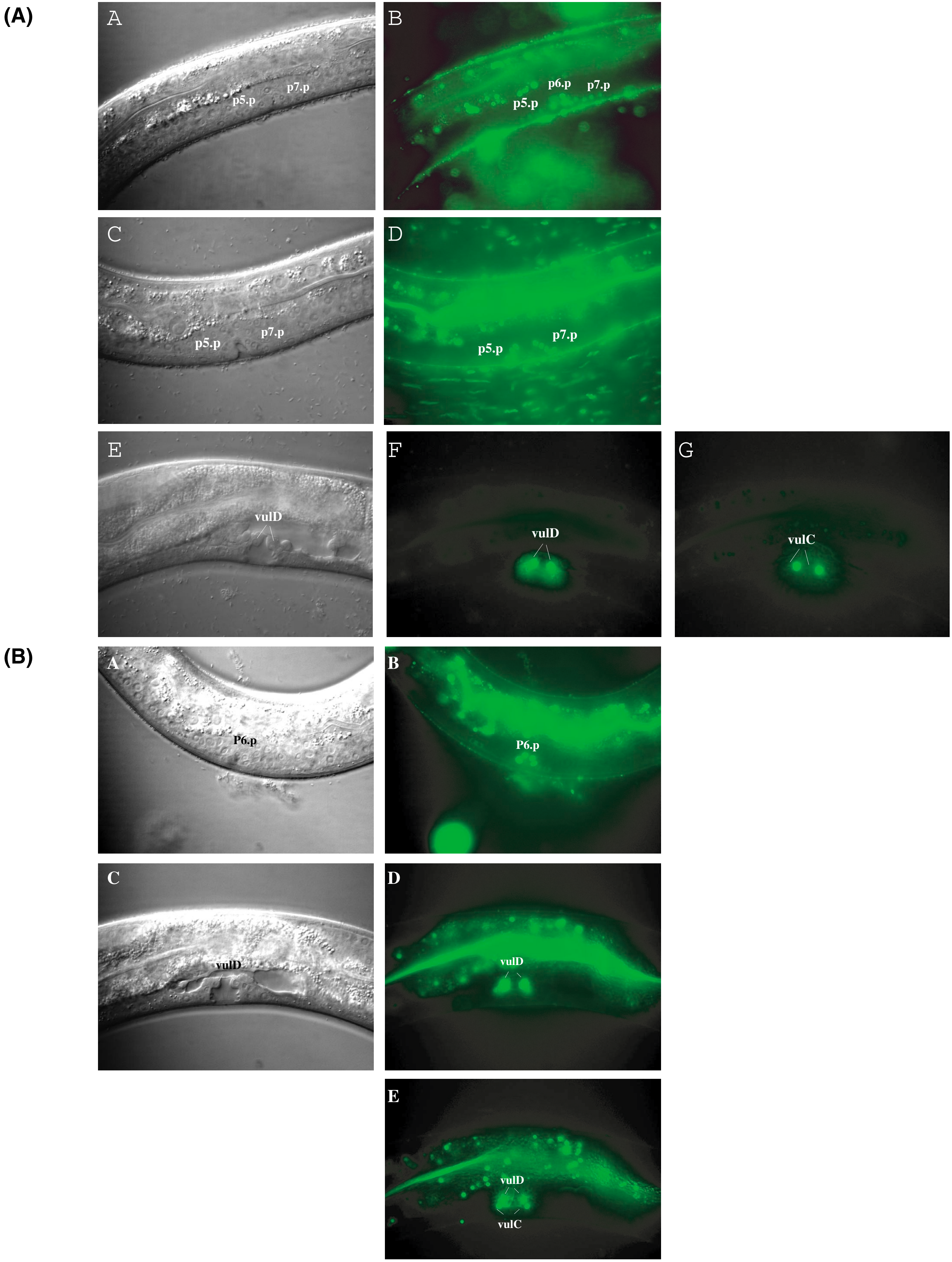
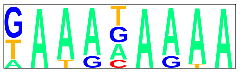
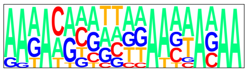
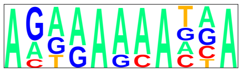
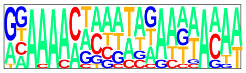
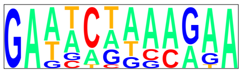


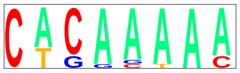
Table 3: AlignACE predictions of overrepresented sequences

A.

Expression	Regions examined	Gene	8 bp motif	10 bp motif
	mk160-161	*Cb-egl-17*	3 (7)	2 (7)
	mk172-173	*Cb-zmp-1*	3 (6)	2 (6)
	mk162-163	*Cb-cdh-3*	9 (12)	5 (7)
	mk164-165	*Cb-cdh-3*	1 (4)	2 (3)
Anchor cell	mk96-134/mk172-173	*Ce-cdh-3/Cb-zmp-1*	1 (12)	2 (13)

B.

Region	Motif	MAP	Consensus	Sites	Comments
mk160-161 *Cb-egl-17*	1.8	13.77		37, 206, 224, 296, 350, 378, 391, 447, 474, 546, 559, 591, 645, 705, 721	Site 37 is located in conserved element B imp. for early expr. Site 559 is located in conserved element D imp. for vulC/D expr.
	2.8	13.33		37,116,206, 226, 378, 453, 474, 521, 548, 591, 644, 704	Site 37 is located in conserved element B imp. for early expr. Sites 521 and 548 located in conserved element D imp. for vulC/D expr.
	3.8	10.84		37, 112, 168, 224, 293, 388, 446, 469, 542, 588, 609, 645, 700, 717	Site 37 located in conserved elem. B imp. for early expr. Sites 542 and part of 588 located in conserved element D imp. for vulC/D expr.
	4.10	11.96		62, 114, 206, 226, 280, 320, 344, 378, 443, 471, 521, 549, 579, 705	Sites 549 and 579 are located in conserved element D imp. for vulC/D expr. Multiple sites overlap motif 2.8 sites.
	5.10	10.19		38, 112, 168, 293, 307, 331, 388, 456, 469, 483, 592, 609, 700, 717	Site 38 is located in conserved element B imp. for early expr. Multiple sites overlap motif 3.8.
mk172-173 *Cb-zmp-1*	1.8	13.64		12, 22, 42, 122, 278, 361, 376, 411, 472, 584, 614, 716, 738	Site 278 is located in conserved element D. This is the only one of motifs in this element that is present in mk50-51
	2.8	13.46		30, 48, 62, 112, 131, 183, 367, 425, 480, 505, 575, 605, 624, 669, 700, 730	
	3.8	10.69		5, 22, 112, 249, 270, 352, 373, 408, 472, 497, 513, 575, 615, 673, 708, 738	Sites 249 and 270 are located in conserved element D.
	4.10	16.17		5, 22, 119, 270, 352, 373, 408, 472, 497, 513, 575, 615, 673, 708, 738	Site 270 is located in conserved element D. Multiple sites overlap 3.8 motif sites.
	5.10	13.21		1, 21, 94, 119, 174, 373, 408, 546, 705, 737	Site 546 is located in conserved element A. Multiple Sites overlap motif 1.8 sites.

Region	Motif	MAP	Consensus	Sites	Comments
mk162-163 *Cb-cdh-3*	1.8	23.92		232, 249, 265, 368, 384, 443, 495, 520, 549, 665, 715, 731, 804, 820	
	2.8	23.66		8, 45, 57, 610, 678, 887, 903, 1006, 1021, 1158, 1196, 1226, 1397	
	3.8	22.62		6, 56, 256, 381, 410, 450, 611, 662, 722, 888, 1022, 1108, 1159, 1194, 1249, 1278	
	4.8	20.83		259, 376, 396, 453, 489, 528, 657, 725, 812, 1062, 1103, 1151, 1217, 1379	Site 1062 is located in conserved element F imp. for vulval expression and in anchor cell gamma region.
	5.8	19.60		251, 377, 445, 502, 585, 658, 717, 813, 878, 997, 1026, 1256, 1378	Site 1028 is located in conserved element F imp. for vulval expression and in anchor cell gamma region.
	6.8	15.75		7, 44, 56, 103, 559, 673, 855, 882, 901, 1001, 1020, 1157	
	7.8	13.39		96, 154, 267, 288, 371, 594, 645, 733, 807, 968, 1045, 1085, 1212, 1338, 1366	Site 1045 is located in conserved element F imp. for vulval expression and in anchor cell gamma region.
	8.8	12.84		20, 180, 239, 270, 362 421, 567, 593, 651, 736, 962 1010, 1036, 1177, 1239, 1302, 1319, 1357	Site 362 in conserved elem. D. Site 1036 in conserved elem. F and Site 1302 and 1319 in conserved elem. A imp region for vulA cell expr.
	9.8	10.63		98, 150, 251, 383, 438, 631, 664, 717, 878, 1087, 1131, 1162, 1186, 1255, 1281	
	10.10	25.99		258, 376, 452, 657, 724, 812, 1061, 1379	Site 1061 is located in conserved element F imp. for vulval expression and in anchor cell gamma region.
	11.10	21.47		6, 43, 55, 397, 407, 608, 885, 1006, 1019, 1156, 1194	

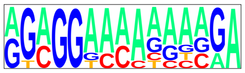
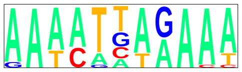
Region	Motif	MAP	Consensus	Sites	Comments
mk162-163 *Cb-cdh-3*	12.10	19.16		20, 180, 239, 270, 362, 736, 798, 958, 1036, 1177, 1239, 1302, 1319, 1357	All sites, except 798 and 958 are the same as motif 8.8. See 8.8 comments.
	13.10	16.61		26, 94, 232, 263, 301, 370, 641, 686, 729, 768, 786, 806, 952, 1044, 1189, 1239, 1296, 1320, 1336, 1365	Site 351 in conserved elem. D, Site 1044 in conserved elem. F, and Sites 1296 and 1239 in conserved elem. A
	14.10	10.43		79, 251, 382, 445, 663, 717, 818, 1228	Overlaps multiple sites with motif 9.8.
mk164-165 *Cb-cdh-3*	1.8	15.29		1, 26, 52, 70, 99, 121, 143, 183, 195, 221, 239, 258	Sites 99 and 121 are in conserved element H. Site 143 is located in conserved element I and Site 194 is located in conserved element K.
	1.10	17.49		3, 19. 45, 63, 92, 176, 214, 232	Site 45 is located in the overlap elements J and H. Site 63 is located in element. H.
	2.10	11.02		19, 45, 63, 92, 141, 176, 214, 232	Site 63 is located overlap elements J and H. Site 92 located in element H. Site 141 located in element I. Site 214 located in element K.
mk96-134 *cdh-3* mk172173 *Cb-zmp-1*	1.8	16.88		[38, 50, 81, 248, 263, 333, 374, 463, 484, 591, 610, 723, 820, 948, 964, 1027] [345, 367, 473, 501, 513, 623, 698, 722]	Sites 38 and 50 in mk96-134 are located in conserved element A of *cdh-3* , which is imp. for anchor cell (alpha region). No sites in conserved elements of *zmp-1*.
	2.10	16.32		[25, 45, 143, 264, 353, 384, 425, 472, 574, 719, 956, 1036, 1095] [86, 168, 373, 507, 612, 630, 662]	Site 25 in mk96-134 is partially in conserved element A of *cdh-3,* which is imp. for anchor cell (alpha region). No sites in conserved elements of *zmp-1*.
	3.10	16.01		[38,74 256, 333, 374, 416, 590, 616, 871, 948, 964, 1027] [31, 306, 345, 366, 473, 501, 623, 673, 699]	Site 38 in mk96-134 is located in conserved element A of *cdh-3,* which is imp. for anchor cell (alpha region). No sites in conserved elements of *zmp-1*. Multiple sites overlap motif 1.8 sites.

Figure 1: Selection of nematode species for comparative genomic analysis

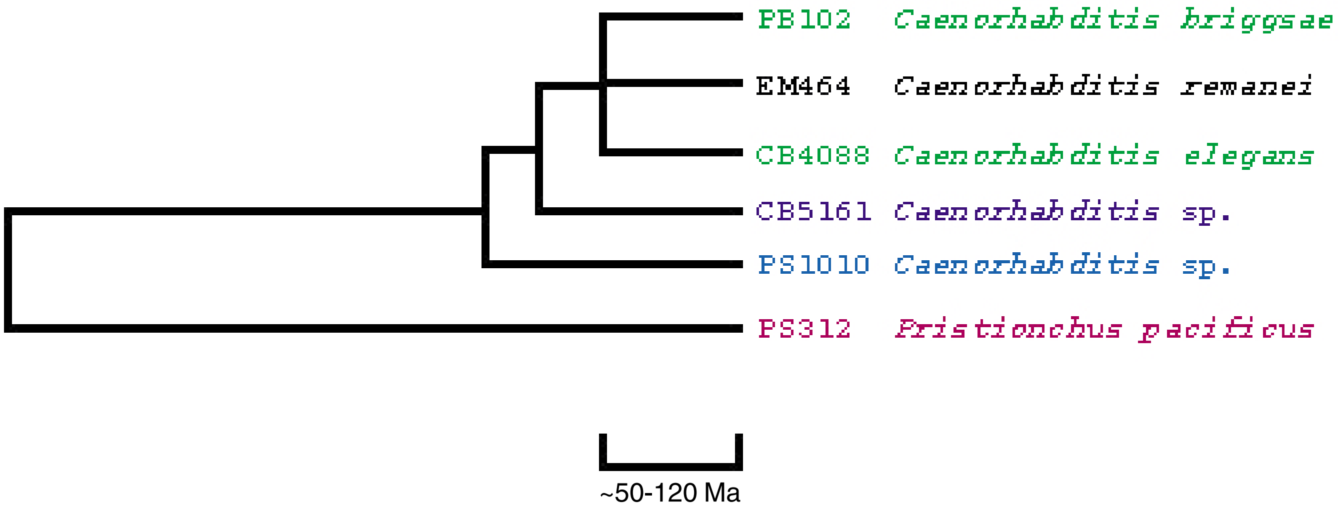
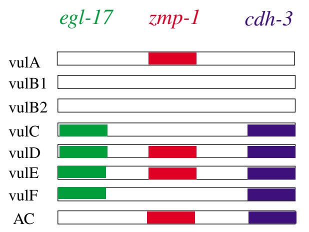
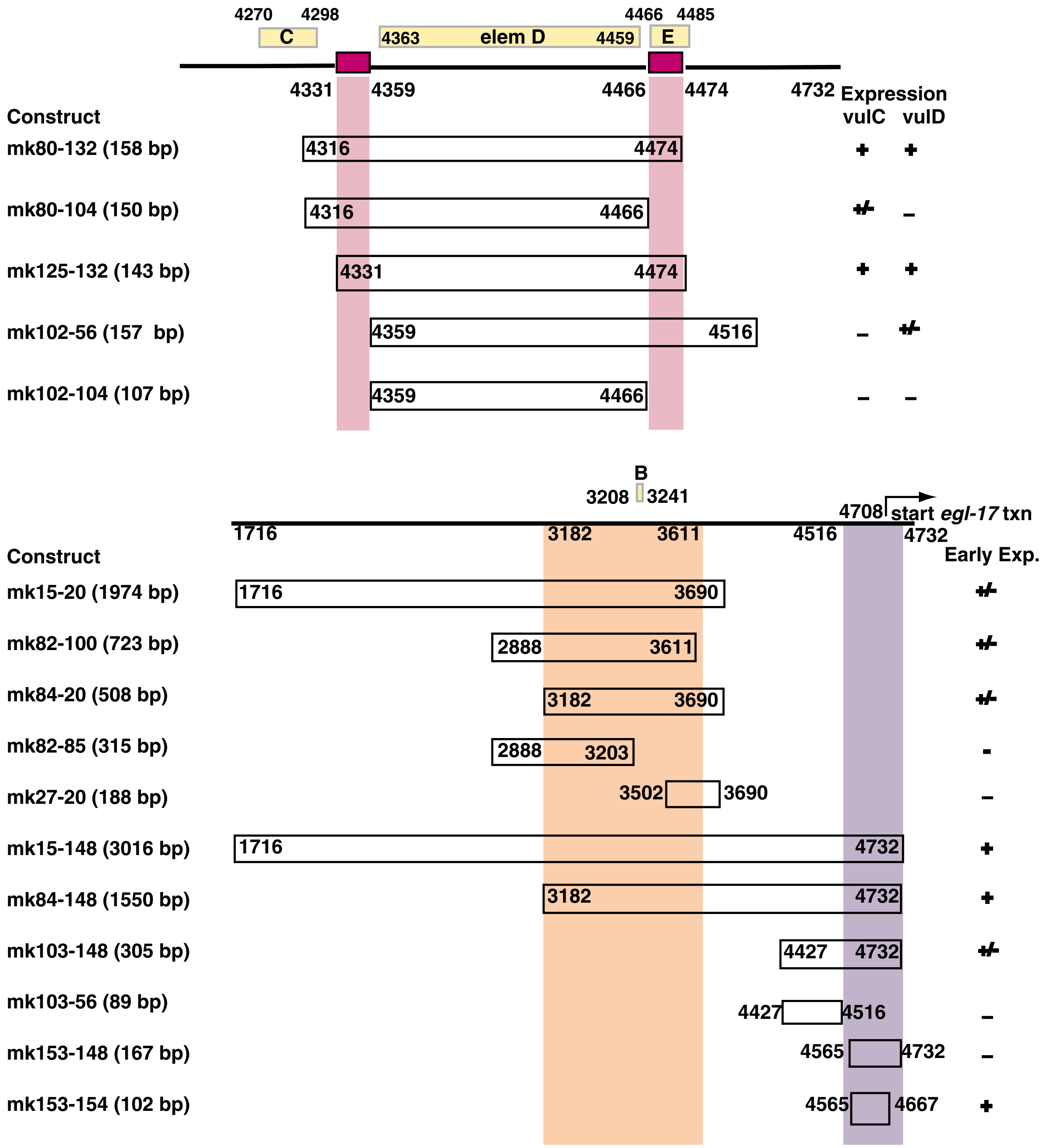


Figure 2: Upstream regions that direct *egl-17* expression





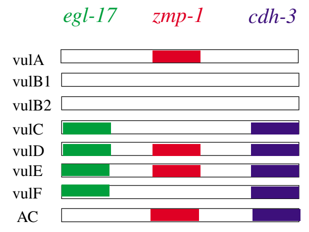


Figure 3: Multiple Regions Direct *zmp-1* expression

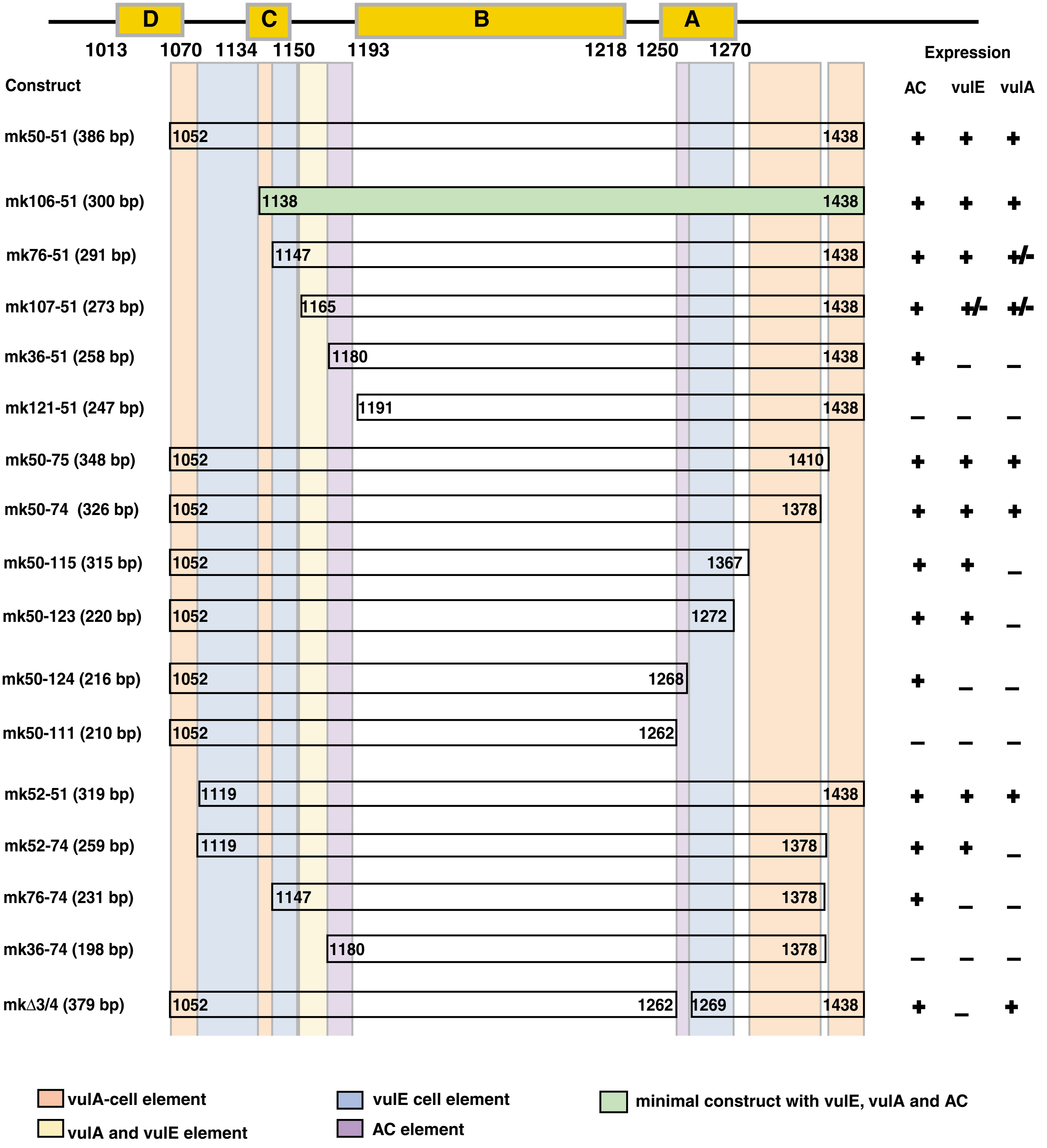
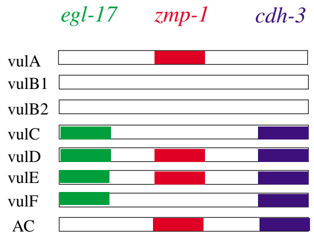


Figure 4:Regions that direct *cdh-3* Expression

2325	2352	A
2499	2529	B
2683	2707	C
2715	2735	D
2987	3006	E
3101	3138	F
4686	4878	HIJK



	egl-17	*zmp-1*	*cdh-3*
vulA		🟥	
vulB1			
vulB2			
vulC	🟩		🟪
vulD	🟩	🟥	🟪
vulE	🟩	🟥	🟪
vulF	🟩		🟪
AC		🟥	🟪