Supplementary Information

In vivo Hox binding specificity revealed by systematic changes to a single cis regulatory module

Carlos Sánchez-Higueras et al.



Supplementary Figure 1: Biochemical and *in silico* predicted site affinity to *main* and *overlapping* class 3 sites or to the *overlapping* site

(a) Sequence comparison of the oligo containing the *main* and *overlapping* wild type sites with that carrying the class 1 mutated site (green) that secondarily generates a higher affinity class 3 site. (b) oligo probes used to test in EMSA the affinity to probes containing variants for both sites or only containing the *overlapping* site after mutation of the main site.

(c) Ubx binding to both the *main* class 2 site and *overlapping* class 3 containing oligo (blue lanes); to a mutant for both sites (grey); to a *main* class 1 site that transformed the *overlapping* site into a High class 3 (green). (d) Ubx binding to both the *main* class 2 site and *overlapping* class 3 containing oligo (blue); to a mutant *main* with a class 3 *overlapping* site (light green); or to a mutant *main* with a high affinity class 3 *overlapping* site (light purple). (e) *In silico* predicted affinity of Ubx to sites shown in (c-d).

Comparison of blue lanes with light green labelled lanes in d shows Ubx binding is maintained or even increased in the absence of the *main* site, indicating Ubx binding to the oligos containing two sites is mostly occurring through the *overlapping* class 3 site. This conclusion is reinforced by the strong binding of Ubx observed in the light purple labelled lanes where only the *overlapping* site remains. These EMSA results indicate that the trunk expression observed in *S1+55cl1* embryos is mediated through the *overlapping* site. Source data are provided as a Source Data file.

•	P07548 DFD_DROME B8RCB0 B8RCB0_BRALA	1	MSSFLMGYPHAPHHVQSPMSMGNGLDPKFPPLADDYHHYNGHYSMTASTGHMSGAVGG HIMSSFLMN	58 20
a	P07548 DFD_DROME B8RCB0 B8RCB0_BRALA	59 21	GAGVGSVGGGGAGGMTGHPHSMHPADMVSDYMANHHNPHSHSHSHTHSLPHHHSNSAISG	118 20
Dm-Dfd Al-Hox4	P07548 DFD_DROME B8RCB0 B8RCB0_BRALA	119 21	HHQASAGGYSSNYANATPPSHPHSHPHAHPHQSLGYYVHHAPEFISAGAVHSDPTNGYGP EEYSQNSYIPSQIPDYYGSPSPQAAQNYGP	178 50
	P07548 DFD_DROME B8RCB0 B8RCB0_BRALA	179 51	AANVPNTSNGGGGGGGGAVGGGANGGYGGGGGGGGGGGGGGGGGGGGG	238 65
	P07548 DFD_DROME B8RCB0 B8RCB0_BRALA	239 66	MMDLFLQCSSTEPPTNTALGLQ-ELGLKLEKRIEEAVPAGQOLQELGHRLRCDDMGSEND VSQAARNVPPPQPSHGQAPGPGPGPETPPPHSRPTSTLSTVQHNH	297 110
	P07548 DFD_DROME B8RCB0_B8RCB0_BRALA	298 111	DMSEEDRLML-DRSPDELGSNDNDDDLGSDSDEDLMAETTDGERIIYPMNKKHVAGVA NVSPPAPSQPTPPPPPPATSTTAASTVSDGSPPSNGPPTVPMVYPMNKVHSNTG-	356 165
	P07548 DFD_DROME B8RCB0 B8RCB0_BRALA	357 166	NGSYOPOME PKROKTANING LISI EKEMINYARIA TARAR ISIAHI UUSEROIKI WOO -SISYNOQOTANSATANING UUSEREEHINARIA TARAR ISIAHI UUSEROIKI WOO	416 224
	P07548 DFD_DROME B8RCB0 B8RCB0_BRALA	417 225	RAMKWKKON KLINTKNVRKKTVDANGNPTPVAKKPTKRAASKKQQAQQQQQQQQQQQQQ AMKWKKQBRLPNTKTRSSSASGSSCPAVNSTVVRHNHHGVTQC	476 268
	P07548 DFD_DROME B8RCB0 B8RCB0_BRALA	477 269	TPVMNECIRSDSLESIGDVSSSLGNPPYIPAAPETTSSYPGSQQHLSNNNNNGSGNNNNN NVSDHGMIEL	536 278
	P07548 DFD_DROME B8RCB0_B8RCB0_BRALA	537 279	II II NNNNNSNLNNNNNNMGHTNLHGHLQQQQSDLMTNLQLHIKQDYDLTAL	586 278
Deformed YPWM A-Hox4 YPWM	PKRQRTAYTRHQ TKRSRTAYTRQQ	ILELE	KEFHYNRYLTRRRRIEIAHTLVLSERQIKIWFQNI KEFHFNRYLTRRRRIEIAHSLGLTERQIKIWFQNI	RRMKWKKDN RRMKWKKDN
Consensus	FKR ^S RTAYTRSQ	UELE	KEFH _e nryltrrrrieiah _e lelerqikiwfqn	RRMKWKKDN
W motif		Heli	x 1 Helix 2 Helix 3	3 -
Willott			HOMEODOMAIN	
b	B8RCA9 B8RCA9 BRALA P10105 LAB_DROME	1	MMDVSSNYGNIPHHHHPHANAYDGYSTTTASAANASSYFAPQQHQPHLQLQQQQQQQQUUQ	0 60
Om Labial	B8RCA9 B8RCA9_BRALA P10105 LAB_DROME	1 61	OPOQHLTYNGYESSBFGNYYPQQOAQLTPPPTSSHQVVQQHQQQQQAQQQQLYPHSHLFS	0 120
Dm-Labial	B8RCA9 B8RCA9_BRALA P10105 LAB_DROME	121	PSAAEYGITTSTTTCNPGTPLHPSSHSPADSYYESDSVHSYYATAAVATVAPPSNSSPIT	0
	B8RCA9 B8RCA9_BRALA P10105 LAB_DROME	181	AANASATSNTQQQQQAAIISSENGMYYNLDCHYPTAQAQAPVHGYAGQIEEKYAAVLH	0 240
	B8RCA9 B8RCA9 BRALA P10105 LAB_DROME	1 241	MEQNDTARMNSYVDYSLCNGDONT ASYAPGMVLEDQDPHMQQATQSQMWHHQQHLAGSYALDANOSLGWHAHHHRGLPHGHLGN	24 300
	B8RCA9 B8RCA9 BRALA P10105 LAB_DROME	25 301	CSPRSYGQDYGVPAYQSCAMNNVDRHVTMGPSGQLPPSAGPGPGPVPGPSPYDPPVIMSN LAN	84 328
	B8RCA9 B8RCA9 BRALA P10105 LAB_DROME	85 329	GDPONFTTYSYNHYSHPGGHHMGNGYGTNNHAAMYSGSFGAELAGSYSSYNSGMNGTVAP QSPAANQQHHQNSVSPNGGNN	144 358
	B8RCA9 B8RCA9 BRALA P10105 LAB_DROME	145 359	PPLDSQYGYMHHHTGQDPMISTSCNPPAPSPPVATYDWHKLKRNPPRTGKP	195 416
	B8RCA9 B8RCA9_BRALA P10105 LAB_DROME	196 417	ASGIASMHDYQMNGQLDMCRGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	195 476
	B8RCA9 B8RCA9_BRALA P10105 LAB_DROME	196 477	-GEYGFTTSG AMAGSAHPNGMGVGLGSGSGLSSGSLSSNTNNSKRTNFYNKGL/TELEKEFHFNKYL/FAA	234 536
	B8RCA9 B8RCA9_BRALA P10105 LAB_DROME	235 537	KVETAAALNLNETQVKIWFQNRRHKQKKRKKENENGPSTPGSGGSPAGEDSPSKST RIELANTQUAFTQVKIWFQNRRHKQKKRKKKKKKSLINDALLTQHSTSVISEKPPQQQOPOP	288 596
	B8RCA9 B8RCA9_BRALA P10105 LAB_DROME	289 597	PELQLKSQGSDLGGNELATGAPSTPTTAMTLTAPTSKQS	288 635
Labial YKWM A-Hox1 YDWM Consensus	NNSGRINFIKO PNNGRINFIKO BURGRINFIKO		KEFHFNRYLTRARRIELANTLOLNETOVKIWFON KEFHYNKYLTRARRVELAAALNINETOVKIWFON KEFHLNRYLTRARRVELAAALNINETOVKIWFON	
		Heli	x 1 Helix 2 Helix 3	3 -
W motif			HOMEODOMAIN	
C	P83949 UBX_DROME B8RCB5 B8RCB5_BRALA	1	MNSYFPQASGFYGHPHQATGMAMGSGGHHQTASAAAAAYRGFPLSLGMSPYANHHLQRT —MSSYFPQASGFYGHPHQATGMAMGSGGHHQTASAAAAAYRGFPLSLGMSPYANHLQRT	60 12
Dm-Ubx	P83949 UBX_DROME B8RCB5 B8RCB5_BRALA	61 13	TODS-PYDASITAACNKIYGDGAGAYKODCLNIKADAVNGYKDINNTGGSNGGG YQPGESLSSSNGPAGVPSCSYGELASTARGGYHAGAYGPYPATNSGPROPTN	113 64
AI-HOX/	P83949 UBX_DROME B8RCB5 B8RCB5_BRALA	114 65	GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	173 93
	P83949 UBX_DROME B8RCB5 B8RCB5_BRALA	174 94	SPVSHRGGSAGGNVSVSGGNGNAGGVQSGVQAGAGTAMANCTISGAAAQTAAASSLEQ SGAACSYGAACSYGAACSY	233 111
	P83949 UBX_DROME B8RCB5 B8RCB5_BRALA	234 112	ASNHTFYPWMAIAGECPEDPTKSKIRSDLTQYGGISTDMGKRY-SESLAGSLLPDW 	288 151
	P83949 UBX DROME B8RCB5 B8RCB5_BRALA	289 152	LOTINGLERRGROYTRYOT DELEKTERVINTARREITENANALCHTEROTRUKTERRE RSTAPE KRGROYTRYOT DELEKTERVINTARREITENANACCHTEROTRUKTERRE	348 211
	P83949 UBX DROME B8RCB5 B8RCB5_BRALA	349 212	KLEKETQAIKELNEQEKQAQAQKAAAAAAAAAAAAAVQGGELDQ MAKKANKLES-LKQQPAESETSSTTS	389 236
Ubx YPWM		TLEL	EKEFHTNHYLTRRRRI EMAHALCLTERQIKIWFQN	RRMKLKKEI
Consensus	RERCROTYTRYC	TLE	EXEMPTION AND THE REAL AND TH	RRWK_KKE.
	K	Lin's		2 N
W motif		Heli		3
			nomeodomain	

Supplementary Figure 2: Sequence alignment of *D. melanogaster* and *A. lanceolatus* orthologous proteins

Full sequence Uniprot alignments of (a) Dfd and Al-Hox4, (b) Lab and Al-Hox1 and (c) Ubx and Al-Hox7. In all cases most conservation maps to the homeodomain especially on the DNA-contacting third helix and on the cofactor interacting hexapeptide motif. Specific alignments of the hexapeptide motif (also known as W motif) and the homeodomain performed with Jalview are shown under the full sequence alignment

Note that the most similar cofactor interacting motives (labelled in orange) between UbxIVa and AI-Hox7 do not align when using either Uniprot or Clustal w software with default settings. UbxIVa contains 4 described W motifs ³⁵.



Supplementary Figure 3: *Amphioxus* Hox monomeric function in *Drosophila*

Expression of *S2-GFP* (a-b) or *vvl1+2-mCherry* (red) and *S2-GFP* (a-b, right panels) in *hth*^{P2} mutant embryos that express in the maxilla and labium (asterisks) the *UAS-AHox7* (a) or *UAS-AHox4* (b) with *sal-Gal4*. AHox-7 can rescue *vvl1+2* but not *S2* expression, while AHox-4 cannot rescue any of them. We did not test the rescue capacity of AHox1 in *hth*^{P2} embryos as AHox1 is unable to rescue any of these constructs in a wild type embryo (see Fig. 4I). Scalebar 50µm.



Supplementary Figure 4: Predicted site affinity of Lab to *vvl1+2* and *vvl1+2vvm*

NRLB predicted relative affinity of Lab-Exd (a) or of monomer Lab (b) to *vvl1+2* or to the remaining Hox binding sites in the *vvl1+2vvm*. Asterisk and discontinuous vertical line in (a) label the position of the *main* site.



Supplementary Figure 5: Predicted site affinity of Dfd to *vvl1+2* and *vvl1+2vvm*

NRLB predicted relative affinity of Dfd-Exd (a) or of monomer Dfd (b) to *vvl1+2* or to the remaining Hox binding sites in the *vvl1+2vvm*. Asterisk and discontinuous vertical line in (a) label the position of the *main* site.



Supplementary Figure 6: Predicted site affinity of Ubx IVa to *vvl1+2* and *vvl1+2vvm*

NRLB predicted relative affinity of UbxIVa-Exd (a) or of monomer UbxIVa (b) to *vvl1+2* or to the remaining Hox binding sites in the *vvl1+2vvm*. Diamond and discontinuous vertical line in (a) label the position of the *overlapping* site.



Supplementary Figure 7: Predicted site affinity of Abd-B to *vvl1+2* and *vvl1+2vvm*

NRLB predicted relative affinity of AbdB-Exd (a) or of monomer AbdB (b) to vvl1+2 or to the remaining Hox binding sites in the vvl1+2vvm. Diamond and discontinuous vertical line in (a) label the position of the *overlapping* site.

<i>vvl</i> 1+2	GGTTAATGATGGCCACACAGGCGACAACTGGCGACTGAGATCTTGTCCCCGCAATTTCCCG	60				
<i>vvl</i> 1+2 vvm	GGTTggTGATGGCCACACAGGCGACAACTGGCGACTGAGATCTTGTCCCGCAATTTCCCG	60				
<i>vvl</i> 1+2	ATCATTTGCTCAGATACGATACGGATTCTGCGAAGTCACCGGGCAAAGGCATCCCACTGA	120				
<i>vvl</i> 1+2 vvm	ATCATTTGCTCAGATACGATACGGATTCTGCGAAGTCACCGGGCAAAGGCATCCCACTGA	120				
<i>vvl</i> 1+2	GTTTGTTGGCTTCTTGTCAAACAAATCATTTACTGCGGATACTCCTCGATATTCCTCAAG	180				
<i>vvl</i> 1+2 vvm	GTTTGTTGGCTTCTTGTCAAACAAAgCAgggaCTGCGGATACTCCTCGATATTCCTCAAG	180				
<i>vvl</i> 1+2	ATACTATCCATTTCACTGTGTAGGTGTGAGCTGCAATTTTCCCTGGAAAAATTACGTCCA	240				
<i>vvl</i> 1+2 vvm	ATACTATCCATTTCACTGTGTGTGTGTGTGTGCAATTTTCCCTGGAAAAAggACGTCCA	240				
<i>vvl</i> 1+2	GCACGGGATTTGATTAATTTATTACCGCTTGTCGAAGGGAAAGTGATCTTCGGGTTCCTA	300				
<i>vvl</i> 1+2 vvm	GCACGGGATTTGATcAATTTcTTcCCGGCTTGTCGAAGGGAAAGTGATCTTCGGGTTCCTA	300				
<i>vvl</i> 1+2	ACGGTTGCGGATCGTAAAAAACATTCGGCAGACACAATTGTTGAATTATCTGCGGGCTGC	360				
<i>vvl</i> 1+2 vvm	ACGGTTGCGGAgCGgAAAAAACATTCGGCAGACACAATTGTTGAAggATCTGCGGGCTGC	360				
<i>vvl</i> 1+2	TGTTGTGGGCACTTTTTTATGAGTTATTTATGTGCGACTGTGGCGCCAACAGGATCATAA	420				
<i>vvl</i> 1+2 vvm	TGTTGTGGGCACTTTgTTATGgGTTATgTATGTGCGACTGTGGCGCCCAACAGGATCATgg	420				
<i>vvl</i> 1+2	AATATGTAGTTATGGGTAAATCTGTGAAAATAAATGTAAGCGTAATCTGGAAAATATTGA	480				
<i>vvl</i> 1+2 vvm	AATATGTAGTTATGGGTAAATCTGTGAAAATAAATGTAAGCGTggTCTGGAAAATATTGA	480				
<i>vvl</i> 1+2	GTGGCCTAGATTCACCACATTAGTCATTTCAAAGAAATTAAAGAAAACTAAAGTAGAAGT	540				
<i>vvl</i> 1+2 vvm	GTGGCCTAGATTCACCACAggAGTCATTTCAAAGAAAggAAAGAAAACTAAAGTAGAAGT	540				
<i>vvl</i> 1+2		600				
<i>vvl</i> 1+2 vvm	ATAAAAATTTTAAATTGTggTGAgggAAATATATCAAATggTGCAAATTTTAAAAAAggA	600				
<i>vvl</i> 1+2	ATTTGATTACTGCTTTATTAAAATTTTTTTTTTTCATATTAATATGATTTATCGTGCATACG	660				
<i>vvl</i> 1+2 vvm	ATTTGAggACTGCTTTAggAAAATTTTTATTTCATAggAATATGAgggATCGTGCATACG	660				
<i>vvl</i> 1+2	GAAATTAAACTGGATTATGG 680					
<i>vvl</i> 1+2 vvm	GAAAggAAACTGGAggATGG 680					
04 14 2 0001						
S1: [1 → 209] S2: [210 → 418] S3: [419 → 680]						

vvl 1+2 genomic location: (Chr3L: 6,764,714-6,765,393)

Supplementary Figure 8: Aligned *vvl1+2* and *vvl1+2vvm* sequences

The vvl1+2 sequence aligned to the vvl1+2vvm sequence indicating the mutated sites. The *S1* sequence is shown in grey, *S2* in green and *S3* in brown. Genomic location coordinates for the vvl1+2 enhancer are included.

REPORTER GENE	Intercalary Segment (ic)	Maxillary Segment	Labial Segment (Ib)	Trunk Segments (T2-Δ7)	Eighth Abdominal Segment (A8)	MAIN&OL Hox/Exd sites	Global Hox Monomer Input per fragment
<i>vvl</i> 1+2	×	অত্যত	অত্যত		অব্যব্য	Main cl2 OL cl3	high
<i>vvl</i> S1+S2	×	<u>N</u> MN	<u>NNN</u>	NANA	NNN	Main cl2 OL cl3	high
vvl S2	X	NAN	×	NNN	<u>N</u> NN	Main cl2 OL cl3	high
<i>vvl</i> S1+55	×	$\mathbf{\nabla}\mathbf{\nabla}$	<u>N</u> N	×	×	Main cl2 OL cl3	weak
<i>vvl</i> S1+55 cl1	N	NNN	<u>N</u> MN		×	Main cl1 Higher OL cl3	weak
<i>vvl</i> S1+55 cl3	×	M	NAN		V	Main cl3	weak
vvl S1+55 cl10	F MANA	NN	NAN	×	×	Main cl1OF	weak
<i>vvl</i> S1+55 cl2O	F 🗵	<u>N</u> NN	<u>N</u> NN	×	×	Main cl2OF	weak
vvl S1+55 cl3O	F 🗵	NN	NN	V		Main cl3OF	weak
<i>vvl</i> S1+55 mut	×	×	×	×	×	none	weak
vvl 1+2 Main&OL n	nut 🔀			MMM	NNN	none	high
vvl S2 Main&OL mu	it 🗵	×	×	×	<u>NNN</u>	none	high
<i>vvl</i> 1+2 vvm	×	×	×	×	×	none	weak
<i>vvl</i> 1+2 vvm-R	×	VVV	NN	VVV	NNN	Main cl2 OL cl3	weak
vvl S1	×	×	×	×	×	none	weak
vvl S3	×	×	×	×	×	none	high
	Highest optimal expression	Medium expres	sion Low Express	ion Occasiona	I not significant express	sion	

Supplementary Table 1: Summary of the segmental expression in different reporter genes

At a glance summary of the expression driven by the different reporters. The levels of expression in each segment are indicated using a subjective five degree scale: occasional not significant expression; low expression; medium expression; high expression; highest optimal expression.

The presence or absence of the *main* and *overlapping* sites and the binding class to which they belong, plus the abundance of monomer sites in each reporter are indicated in the last two columns.



Unprocessed scans of EMSA experiments

Primer List

vvI S1+55

Forward vv/ S1 EcoRI: 5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

Rvs vv/ S1+55 EcoRI: 5'-AAAGAATTCTAATAAATTAATCAAATCCCGTGCTGGACGTAA-3'

vvl S1+55 Cl1

Forward vv/ S1 EcoRI: 5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

Reverse vv/ S1+55 Cl1 EcoRI: 5'-AAAGAATTCTAATAAATCAATCAAATCCCGTGCTGGACGTAA-3'

vvl S1+55 Cl3

Forward vv/ S1 EcoRI: 5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

Reverse vv/ S1+55 CI3 EcoRI: 5'-AAAGAATTCTAATAAATAAATCAAATCCCGTGCTGGACGTAA-3'

vvl S1+55 Cl1 OF

Forward vv/ S1 EcoRI: 5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

Reverse vv/ S1+55 CI1 OF EcoRI: 5'-AAAGAATTCTAAGTAATCAATCATATCCCGTGCTGGACGTAA-3'

vvl S1+55 Cl2 OF

Forward vv/ S1 EcoRI: 5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

Reverse S1+55 Cl2 OF: 5'- AAAGAATTCTAAGTCATTAATCATATCCCGTGCTGGACGTAA-3'

vvl S1+55 Cl3 OF

Forward vvl S1 EcoRI:

5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

Reverse vv/ S1+55 CI3 OF EcoRI:

5'-AAAGAATTCTAAGTCATAAATCATATCCCGTGCTGGACGTAA-3'

vvl S1+55 mutant

Forward vv/ S1 EcoRI:

5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

Reverse vvl S1+55 mutant EcoRI:

5'-AAAGAATTCTAATAAATTGATCAAATCCCGTGCTGGACGTAA-3'

vvl S2

Forward vv/ S2 EcoRI : 5':AAAGAATTCGCTGCAATTTTCCCTGGAAAAATTACGTCCAGCACGGGATTTGATTAATTTATA-3'

Rvs S2 EcoRI: 5'-AAAGAATTCATGATCCTGTTGGCGCCACAGTCGCACAT-3'

vvl S2 MAIN&OL mutant

Forward vv/ S2 Main&OL mut EcoRI: 5'-AAAGAATTCGCTGCAATTTTCCCTGGAAAAATTACGTCCAGCACGGGATTTGATCAATTTATTA-3'

Rvs S2 EcoRI: 5'-AAAGAATTCATGATCCTGTTGGCGCCACAGTCGCACAT-3'

vvl S1+S2

Forward vv/ S1 EcoRI: 5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

Rvs S2 EcoRI: 5'-AAAGAATTCATGATCCTGTTGGCGCCACAGTCGCACAT-3'

vvl 1+2 MAIN&OL mutant*

External Forward vv/ S1 EcoRI: 5'-AAAGAATTCGGTTAATGATGGCCACACAGGCGAC-3'

External Reverse vv/ S3 EcoRI: 5'-AAAGAATTCCCATAATCCAGTTTAATTTCCGTATGCACGATAAA-3' **Forward** *vvl* **S2 Main&OL mutant**: 5'- GATCAATTTCTTCCCGCTTGTCGAAGGGAAAGTG-3'

Reverse vv/ S2 Main&OL mutant: 5'-CCTTCGACAAGCGGGAAGAAATTGATCAAATCCCGTGCTGG-3'

* PCR mutagenesis in two steps (First Round PCRs and Second Round of Overlapping PCRs).

vvl 1+2 *vvm*-R*

External Forward *vvl***S1***vvm***EcoRI:** 5'-AAAGAATTCGGTTGGTGATGGCCACACAGGCGACAA-3'

External Reverse vvl S3 vvm EcoRI: 5'-AAAGAATTCCCATCCTCCAGTTTCCTTTCCGTATGCAC-3'

Fwd vv/ S2 Main&OL restored: 5'-AGCACGGGATTTGATTAATTTATTACCGCTTGTC-3'

Rvs vv/ S2 Main&OL restored: 5'-GACAAGCGGTAATAAATTAATCAAATCCCGTGCT-3'

* PCR mutagenesis in two steps (First Round PCRs and Second Round of Overlapping PCRs)