

Figure S1. Analysis of *achaete-scute* mutants and various *achaete-scute* cis-regulatory elements in *D. melanogaster*. Related to Figure 2. (A) Analysis of *achaete-scute* mutants in *D. melanogaster* reveals a 5-kb candidate region driving expression in hypandrial bristles.

Genomic organization of the *AS-C* locus in *D. melanogaster*. Arrows and arrowheads indicate the coding regions of *yellow* (*y*), *achaete* (*a*), *scute* (*sc*), *lethal of scute* (*l(1)sc*), *pepsinogen-like* (*pcl*), *asense* (*ase*) and *cytochrome P450-4g1* (*Cyp4g1*) genes, respectively. The *achaete-scute* mutants are represented in blue if they display a wild-type phenotype of 2 bristles and in red if they show a mutant phenotype with a reduced number of hypandrial bristles. Point mutations are represented as crosses, insertions as triangles, deletions as solid horizontal bars, inversions as thin horizontal bars terminated by an arrow, and translocations as thin horizontal bars terminated by a bracket. Comparison of all the tested *achaete-scute* mutants suggests that a 5-kb region (black arrows) located 45 kb away from *sc* 5' coding region drives expression in the developing hypandrial bristles. Fig. S3. Three cis-regulatory elements of the *achaete-scute* complex (*AS-C*) govern expression in the hypandrium in *D. melanogaster*. (B-Q) Three cis-regulatory elements of the *achaete-scute* complex (*AS-C*) govern expression in the hypandrium in *D. melanogaster*. Expression of *UAS-singed.RNAi* results in singed hypandrium bristles compared to wild-type (B) with *15E09-* (C), *18C05-* (D) and *054839-GAL4* (E). Expression of *UAS-sc* (F-Q) in wild-type (N-Q), *sc*²⁹ (F-I) or *sc*^{M6} (J-M) mutant background produces extra bristles with *18C05-GAL4* (H,L,P) but not in absence of a *GAL4* reporter line (F,J,N) nor with *15E09-* (G,K,O) and *05439-GAL4* (I,M,Q). White and black circles indicate zero or two hypandrial bristles, respectively. Average bristle numbers are shown in squares (H,L,P). n: number of scored individuals. (B-Q) Scale bars indicate 25 μ m.

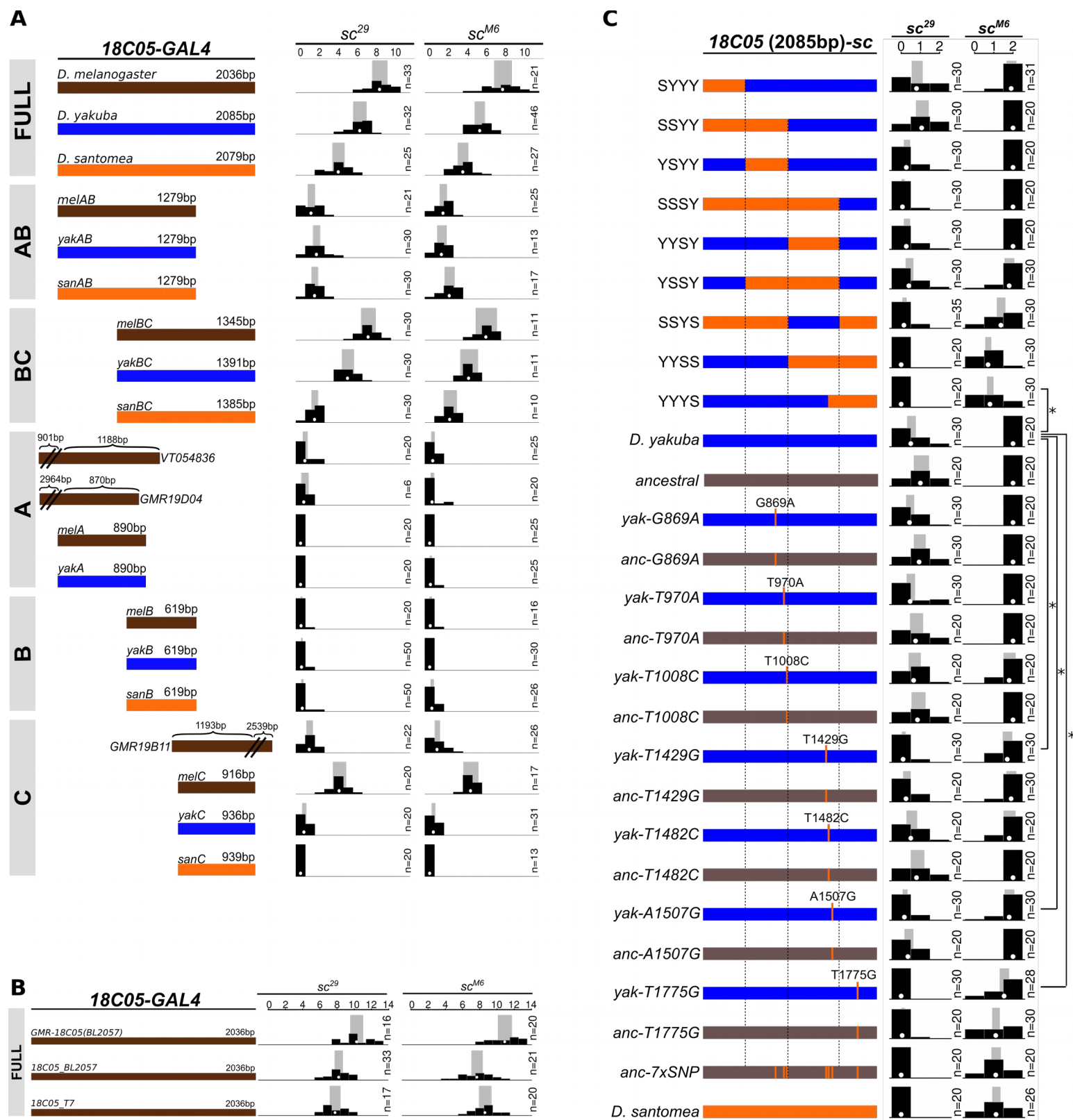


Figure S2. Various 18C05 constructs and their effect on hypandrial bristle number. Related to Figure 3.

(A) Full-length 18C05 is required for complete expression in the hypandrial bristles. Rescue of the hypandrial bristle loss of *sc²⁹* (left column) and *sc^{M6}* (right column) *D. melanogaster* mutants by expression of GAL4 with UAS-sc driven by various genomic regions from *D. melanogaster* (brown), *D. yakuba* (blue) and *D. santomea* (orange). Distribution of hypandrial bristle number (black histogram), together with mean (white dot) and 95% confidence interval (grey rectangle) from a fitted GLM Quasi-Poisson model are shown for each genotype. n: number of scored

individuals. For all three species we found that smaller segments induced significantly fewer bristles than the corresponding full region (GLM-Quasi-Poisson, $F(19, 509) = 161.7$, all $p < 0.02$ for sc^{29} ; $F(19, 415) = 125.9$, $p < 0.03$ for sc^{M6}). Furthermore, the ~1400 bp 3' region (region BC) from *D. santomea* rescues fewer bristles than the corresponding region from *D. melanogaster* and *D. yakuba*. Moreover, the ~900 bp 3' region (region C) from the three species show striking differences: the region from *D. melanogaster* rescues about 4 bristles on average, the region from *D. yakuba* rescues about 0.4 bristles on average, and the region from *D. santomea* rescued no bristles. Addition of DNA 5' of the *18C05* region increases hypandrial bristle numbers (*VT054836* different from 0 (*melA*), Exact-Poisson, $p < 10^{-16}$), whereas addition of DNA 3' of *18C05* reduces the number of bristles (*GM19B11* vs. *melC*; GLM-Quasi-Poisson, $F(19, 509) = 161.7$, $p < 10^{-16}$ for sc^{29} ; $F(19, 415) = 125.9$, $p < 10^{-16}$ for sc^{M6}).

(B) Comparison of *18C05 melanogaster-GAL4* constructs. Three different *D. melanogaster 18C05* sequences were tested with *UAS-sc* in the hypandrium in sc^{29} (first column) and sc^{M6} (second column). Distribution of bristle number (black histogram), together with mean (white dot) and 95% confidence interval (grey rectangle) from a fitted GLM Quasi-Poisson model are shown. Hypandrial bristle number, for *GMR-18C05 (BL2057)* is significantly higher than *18C05_BL2057* and *18C05_T7* in both backgrounds (GLM-Quasi-Poisson, $F(2, 63) = 16.88$, both $p < 10^{-6}$ for sc^{29} ; $F(2, 58) = 20.9$, $p < 10^{-10}$ and $p < 10^{-5}$ for sc^{M6}). *GMR-18C05 (BL2057)* comes from the Janelia Research Campus collection [41]. *18C05_BL2057* and *18C05_T7* were cloned in this study. *GMR-18C05* and *18C05_BL2057* are the same sequence (from *D. melanogaster* Bloomington Stock Center Strain #2057), cloned in opposite direction. *18C05_T7* contains the *18C05* sequence of *D. melanogaster T.7* strain. The *GMR-18C05* fragment is inserted in the expression vector 3'-5' compared to the *D. melanogaster* genome sequence. In contrast, the *18C05_BL2057* and *18C05_T7* are cloned 5'-3'. Comparing bristle number between *GMR-18C05* and *18C05_BL2057* shows that the orientation of the cis-regulatory region has an effect on bristle number. All our *18C05* constructs were inserted in the same orientation, 5'-3'.

(C) Three *D. santomea*-specific substitutions in *18C05* cause the loss of hypandrial bristles.

Various full-length *18C05* regions were cloned in front of the *D. melanogaster sc* coding region and tested for their ability to rescue hypandrial bristles in sc^{29} (left column) and sc^{M6} (right column) mutant backgrounds. The *18C05* sequence from *D. yakuba* (blue, labeled with Y) and from *D. santomea* (orange, labeled with S) were divided into four subregions which were then fused together in different chimeric combinations. *D. santomea*-specific substitutions are represented as dark orange bars. Seven *D. santomea*-specific substitutions were introduced into either the *D. yakuba* region (blue) or the ancestrally reconstructed *18C05* region (dark grey). n: number of scored individuals. *: $p < 0.05$. Replacing the 5' half (1-1020 bp) of *D. yakuba 18C05* by *D. santomea* corresponding region has little or no effect (GLM-Quasi-Poisson, $F(26,658) = 12.09$, SSYY versus *D. yakuba*: $p = 0.031$ in sc^{29} , YSY Y not different from *D. yakuba*: $p = 0.432$ in sc^{29} , SYYY not different from *D. yakuba*: $p = 0.432$ in sc^{29}) whereas replacing the 3' end (1573-2088 bp) decreases bristle number (GLM-Quasi-Poisson, $F(26,618) = 16.04$, YYYS versus *D. yakuba*: $p < 10^{-16}$ in sc^{M6} , YYSS versus *D. yakuba*: $p < 10^{-16}$ in sc^{M6}). The 3' middle part (1021-1572 bp) has no effect (YYSY not different from *D. yakuba*: $p = 0.432$ in sc^{29} , YYSS versus YYYS: $p = 0.579$ in sc^{M6}) except in presence of the *D. santomea* 5' half (1-1020 bp), in which case it decreases bristle number (SSSY versus SSYY: $p < 10^{-4}$ in sc^{29}). Our analysis of chimeric constructs thus indicates that at least two changes, in regions 1021-1572 and 1573-2088, contribute to the reduced ability of *D. santomea 18C05* to produce hypandrial bristles. Three substitutions significantly decreased the number of rescued bristles in the *D. yakuba* and/or the ancestral sequence (T1429G: $p = 0.021$ and $p = 0.009$ in sc^{29} , respectively; A1507G: $p = 0.066$ and $p = 0.03$ in sc^{29} , respectively; T1775G: $p = 0.013$ and $p = 10^{-8}$ in sc^{M6} , respectively).

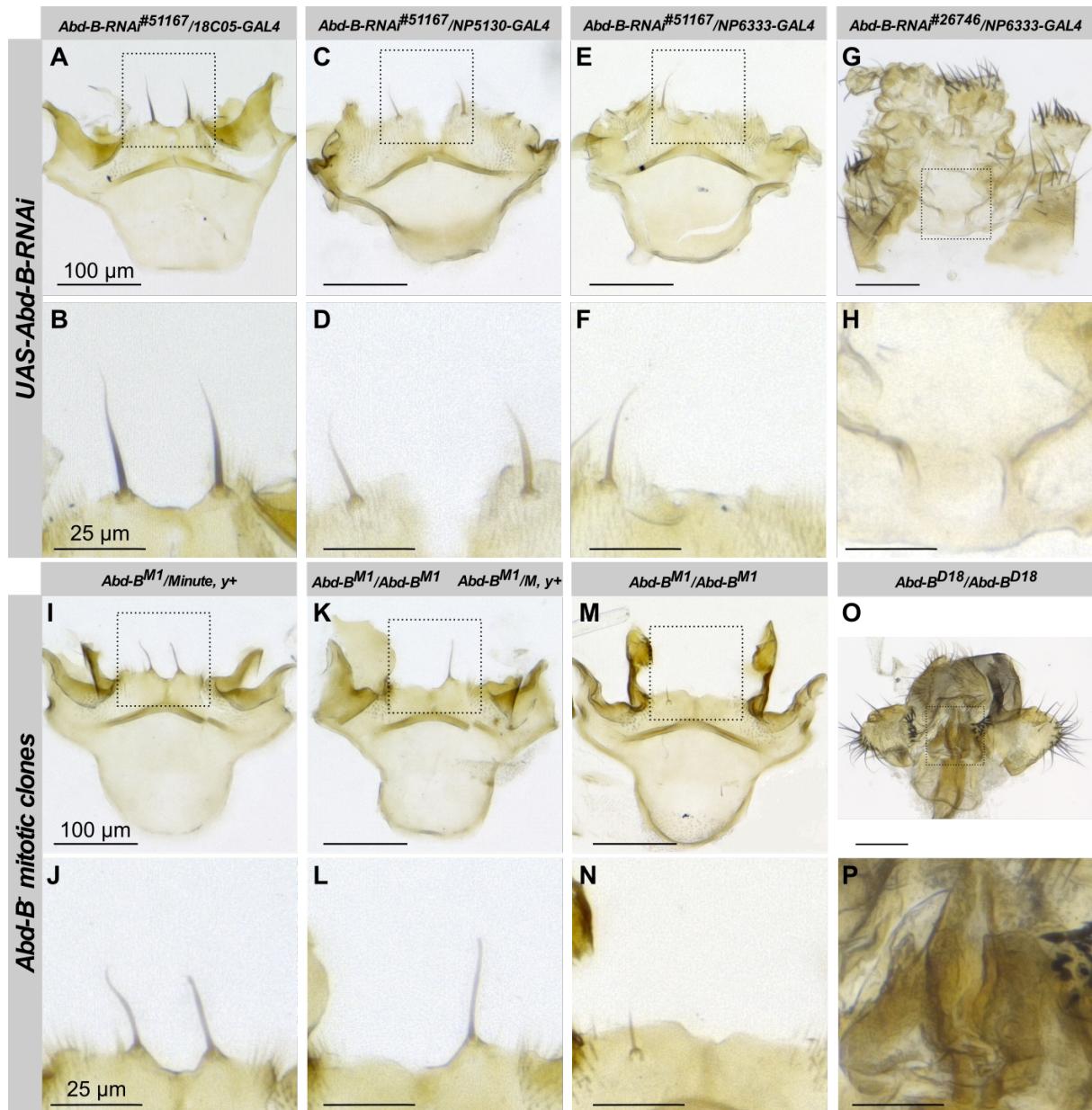


Figure S3. Hyandrial bristle development affected in *Abd-B-RNAi* and *Abd-B* clones. Related to Figure 4.

Wild-type hyandrial bristles. (B) In *Abd-B-RNAi*^{#51167}/*GMR18C05-GAL4* males the hyandrial bristles are thinner. (C) In *Abd-B-RNAi*^{#51167}/*NP5130-GAL4* hyandrial bristles are thinner and shorter. (D) In *Abd-B-RNAi*^{#51167}/*NP6333-GAL4* hyandrial bristles are thinner or lost. (I-N) In *Abd-B* clones hyandrial bristles are lost. Clones were selected in *Minute, y⁺* background. One hypandrium with two *Minute, y⁺* bristles (I-J), one hypandrium with only *Minute, y⁺* bristle (K-L) and one hypandrium with no hypandrial bristle (M-N) are shown. Extreme transformations of genitalia with abnormal hypandrium are shown in *Abd-B-RNAi*^{#26746}/*NP6333-GAL4* (G-H).

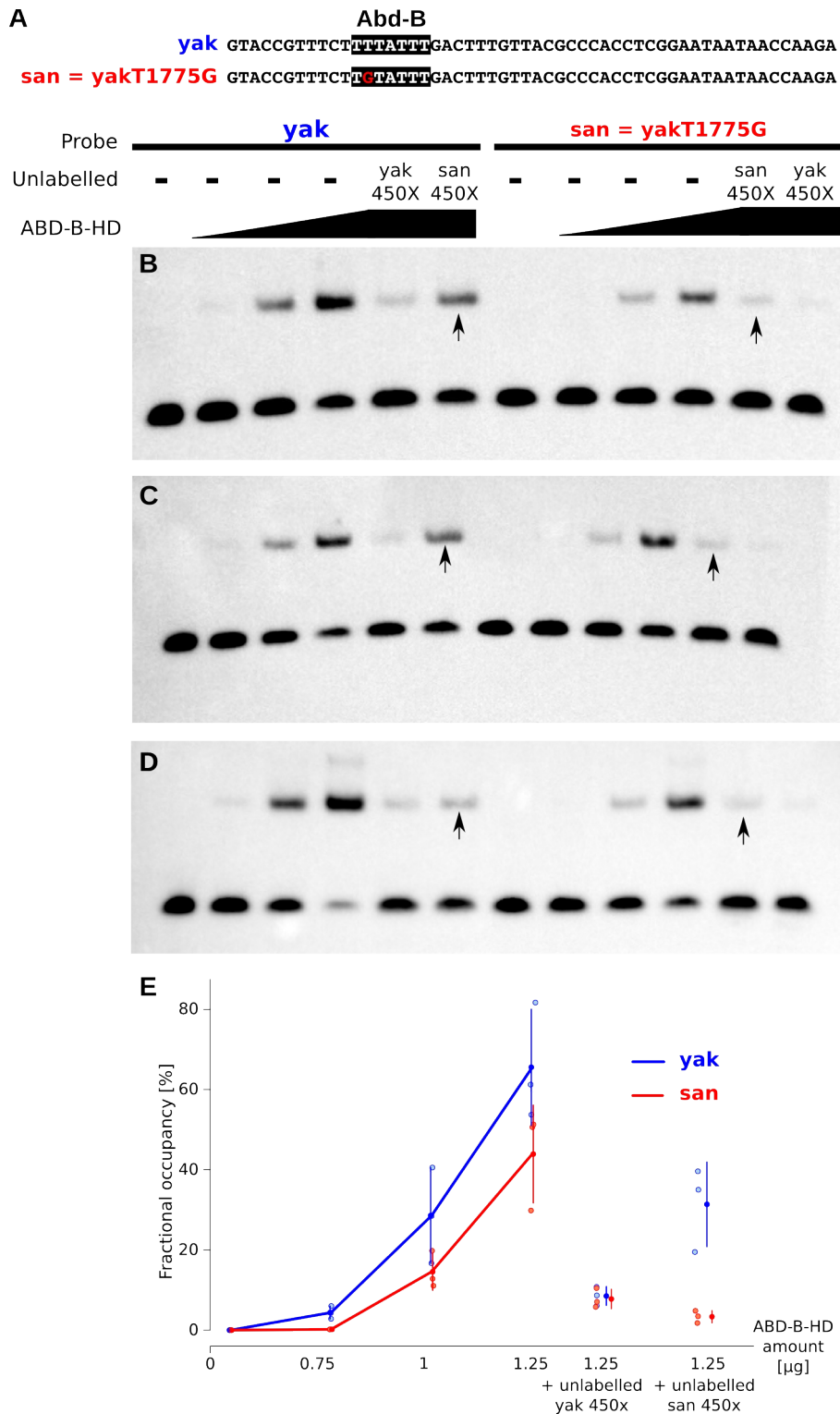


Figure S4. *D. santomea*-specific substitution T1775G affects Abd-B binding. Related to Figure 4.

(A) Two double-stranded oligonucleotides were used in EMSA, one with *D. yakuba* sequence (yak, blue) and one with *D. santomea* sequence (san=yakT1775G, red). The putative Abd-B binding site is shown in black. (B-D) Three parallel EMSA were performed with increasing amount of Abd-B-HD (0.75 µg, 1.0 µg and 1.25 µg) and yak probe or san probe. In competitor lanes 450x excess of unlabelled yak probe or san probe were added to the binding reaction. In all experiments san probe binds less than yak probe to Abd-B-HD, and unlabelled san probe competes less the band shift of the labelled probe compared to the yak probe (indicated with arrows). (E) Quantification of EMSA shifts. Fractional occupancy (ratio of bound/(free+bound) probe) is shown for three parallel

experiments. Mean (dot) and standard deviation (bar) are shown for *yak* and *san* probes, in blue and red, respectively.

Name	Origin and full genotype	Description
sc^{M6}	Bloomington Stock Center #52668 $sc^{M6}/FM7i, P\{w^{+mC}=ActGFP\}JMR3$	Nonsense Glu114X mutation in <i>scute</i> .
ac^{CAMI}	Bloomington Stock Center #36540 $y^1 P\{w^{+mW.hs}=GawB\}CG32816^{NP6014} ac^{cami}$	Deletion from 263861 to 264099, which includes nucleotides 28 to 265 of <i>achaete</i> coding sequence, and insertion of 21 bp at the same location (remains of an excised P-element).
sc^6	Bloomington Stock Center #108 $sc^6 w^a$	Deletion from between ca. 315550 and 316484 to between 338537 and 341499.
ase^1	Bloomington Stock Center #104 $Df(1)ase-1, sc^{ase-1} pn^1/C(1)DX, y^1 f^1$	Deletion from between 341500 and 343845 to between 360269 and 360967.
sc^5	Bloomington Stock Center #178 $y^1 sc^5$	1.2-kb deletion between 346503 and 347999.
ac^1	Bloomington Stock Center #8715 $y^1 ac^1 w^{1118}; P\{w^{+mC}=GAL4-ac.13\}1$	Deletion from ca. 242000 to between 257781 and 259255.
sc^1	Bloomington Stock Center #176 $y^1 sc^1$	Gypsy insertion between 322067 and 322068.
ac^{sbm}	Given by P. Simpson. ac^{sbm}	P element insertion between 264099 and 264100, after nucleotide 27 counting 5' from the <i>achaete</i> ATG.
ac^{Hw-1}	Bloomington Stock Center #109 $Df(1)sc10-1, sc^{10-1}/y^1 ac^{Hw-1}$ (this stock produces only yellow males)	Gypsy insertion between 264219 and 264671. Transcription of <i>achaete</i> terminates within Gypsy sequence. The resulting truncated transcript is overabundant compared to wildtype.
sc^{29}	Bloomington Stock Center #1442 $In(1)sc^{29}, sc^{29} w^a eag^{sc^{29}}$	Inversion. Left breakpoint is between 336621 and 337006.
$ac^1 sc^1$	Bloomington Stock Center #4596 $y^1 ac^1 sc^1 pn^1$	See ac^1 and sc^1 .
sc^H	Bloomington Stock Center #4055 $C(1)DX, y^1 f^1; T(1;4)sc^H, sc^H$	Transposition. Left breakpoint is between 317578 and 318088.
sc^9	DGRC Kyoto Stock Center #102028 $In(1)sc^9, sc^9 w^a f^1 Bx^1$	Inversion. Left breakpoint is between 318656 and 320436.
sc^{S2}	Bloomington Stock Center #3333 $T(1;2)sc^{S2}, y^+ sc^{S2}: cn^1 M(2)53^1/+; CyO$	Translocation associated with a deletion which is between 3.3 kb and 4.4 kb. Left breakpoint of the deletion is between 329,139 and 333,308. Minimal extent of the deletion is 330,042-333,308.
sc^7	Bloomington Stock Center #723 $Df(1)B/In(1)sc^7, In(1)AM, sc^7 ptg^4$	Inversion associated with a deletion which is between 2.5 kb and 3.1 kb. Left breakpoint of the inversion is 37 kb 3' of the <i>sc</i> structural gene [S3]. Minimal extent of the deletion is 332225- 336481.

ac^3 sc^{10-1}	Bloomington Stock Center #36541 $In(1)ac^3, sc^{10-1} ac^3 w^1 sable^1/FM7i,$ $P\{w^{+mC}=ActGFP\};JMR3$	Inversion whose left breakpoint is between 263169 and 264010 (ac^3) and nonsense Glu163X mutation in <i>scute</i> (sc^{10-1}).
sc^4	Bloomington Stock Center #789 $In(1)sc^4, y^1 sc^4 ABO-X^1$	Inversion. Left breakpoint is 7 kb 3' of the <i>scute</i> transcribed region.
sc^8	Bloomington Stock Center #842 $T(1;3)sc^{260-15}, sc^{260-15}/FM6 B^1 dm^1 sc^8 y^{31d}$ (this stock produces only Bar males)	Inversion. Left breakpoint is between 275383 and 276564.
ac^1 sc^{19}	DGRC Kyoto Stock Center #107246 (ordered from Bloomington Stock Center with ancient number #3822) $Df(1)sc^{19}/y^1 ac^1; Dp(1;2)sc^{19}/In(2L)Cy, S^2 Cy^1$	See also ac^1 . Translocation. Left breakpoint is ca. 343845.

Table S1. *achaete-scute* mutant lines used. Related to Figure 2.

Coordinates are for *D. melanogaster* reference genome iso-1 (FB2013_03, version 3). Based on restriction sites we estimated that point 0 of [S1,S2] corresponds to position 330342 in iso-1.

Compared to other *D. melanogaster* strains, iso-1 contains a 6127-bp transposable element named 3S18{ }4/TE19523 at position 322,507-328,633. Coordinates were adjusted to account for the shift due to the transposable element.

	Number of Flies with 0 bristles	Number of Flies with 1 bristles	Number of Flies with 2 bristles	Number of Flies with 3 bristles	Total Number of Flies
Null Mutations (Coding)					
<i>sc</i> ^{M6}	15	0	0	0	15
<i>ac</i> ^{CAMI}	0	0	15	0	15
Deletions (Cis-Regulatory)					
<i>sc</i> ⁶	16	0	0	0	16
<i>ase</i> ¹	0	3	12	1	16
<i>sc</i> ⁵	0	0	17	0	17
<i>ac</i> ¹	0	0	10	0	10
Insertions					
<i>sc</i> ¹	17	0	0	0	17
<i>ac</i> ^{sbm}	1	0	14	0	15
<i>ac</i> ^{Hw-1}	0	0	15	0	15
Complex changes (inversions, translocations, etc.)					
<i>sc</i> ²⁹	16	0	0	0	16
<i>ac</i> ¹ <i>sc</i> ¹	15	0	0	0	15
<i>sc</i> ^H	6	0	0	0	6
<i>sc</i> ⁹	15	0	0	0	15
<i>sc</i> ^{S2}	5	0	0	0	5
<i>sc</i> ⁷	14	0	0	0	14
<i>ac</i> ³ <i>sc</i> ¹⁰⁻¹	14	0	0	0	14
<i>sc</i> ⁴	13	0	0	0	13
<i>sc</i> ⁸	0	0	23	0	23
<i>ac</i> ¹ <i>sc</i> ¹⁹	0	0	8	0	8

Table S2. Hypandrial bristle number in *achaete-scute* mutants. Related to Figure 2.

TF	1429 (yak)	T1429G (san)	1507 (yak)	A1507G (san)	1775 (yak)	T1775G (san)
Abd_B	-	-	-	-	0.92	0.78
ara	-	-	-	-	0.73	0.95
Awh	-	-	0.94	0.85	-	-
B-H1	0.64	0.93	-	-	-	-
B-H2	0.67	0.91	-	-	-	-
Br-Z2	-	-	-	-	0.9	0.62
CG1569G-RA	-	-	0.94	0.94	-	-
Dbx	-	-	0.94	0.87	-	-
lbe	-	-	0.95	0.74	-	-
lbl	-	-	0.93	0.74	-	-
Lim3	-	-	0.92	0.85	-	-
mirr	-	-	-	-	0.71	0.93
onecut	-	-	0.99	0.98	-	-
vvl	-	-	0.92	0.83	-	-

Table S3. JASPAR binding scores for 6 loci of interest within the 18C05 region. Related to Figure 4.

Predicted transcription factors (TF) binding sites and their binding scores are shown for 3 loci of interest within the 18C05 region which underwent *D. santomea*-specific substitutions associated with a decrease in hypandrial bristle number. For each position the left column indicates binding scores for the *D. yakuba* sequence and the right column for the *D. santomea* sequence. All the TF predicted by JASPAR with binding scores higher than 0.9 for at least one locus are shown. “-” indicates cases where binding scores of the *D. yakuba* sequence is less than 0.9. In red are binding scores that were changed by the mutation. We found for T1429G a gain of two repressors (B-H1 and B-H2), for A1507G a loss of several TF binding sites (AWH, DDX, LBE, LBL, LIM3 and VVL), for T1775G a gain of ARA and MIRR binding and a loss of BR-Z and ABD-B binding sites.

Genotype	Number of dissected flies	Number of recognizable hypandrium	0 bristle	1 bristle	2 bristles	3 bristles	2 misplaced bristle
<i>Abd-B.RNAi</i> ²⁶⁷⁴⁶ / <i>NP6333-GAL4</i>	50	0					
<i>Abd-B.RNAi</i> ²⁶⁷⁴⁶ / <i>NP5130-GAL4</i>	50	0					
<i>Abd-B.RNAi</i> ⁵¹¹⁶⁷ / <i>NP6333-GAL4</i>	50	9		1(t,s)	6(t,s)	1(t,s)	1(L)
<i>Abd-B.RNAi</i> ⁵¹¹⁶⁷ / <i>NP5130-GAL4</i>	50	11			9(t,s)		1(L), 1(R)
<i>Abd-B.RNAi</i> ⁵¹¹⁶⁷ / <i>GMR18C05-GAL4</i>	10	10			7(n), 3(t)		
<i>yw hsflp122; FRT82B hs-CD2 y⁺ M(3) w¹²³ /FRT82B Abd-B^{M1} red¹ e¹¹ ro¹ ca¹</i>	30	10	2	2(M,y ⁺)	2(M,y ⁺), 1(M,y ⁻), 1(M,y ⁺ ; M,y ⁻)		2(R,M,y ⁻)
<i>yw hsflp122; FRT82B hs-CD2 y⁺ M(3) w¹²³ /FRT82B Abd-B^{D18}</i>	52	2	1	1(y ⁺)			

Table S4. Hypandrial bristle phenotypes in *ABD-B RNAi* lines and mitotic clones. Related to Figure 4.

Abd-B RNAi was induced with different *GAL4* drivers and *Abd-B^{M1}* and *Abd-B^{D18}* mutations were used in mitotic clones. The number of dissected flies and recognizable hypandrium with a shape similar to wild-type are shown. The number of hypandrium with 0, 1, 2, 3 or with 2 misplaced bristles are shown in separate columns. Different bristle phenotypes are indicated with the following labels: n: normal, t: thin, s: shorter than wild-type, M: Minute (thinner and slightly shorter than wild-type), y+: non yellow, y-: yellow, R: 2 bristles on the right, L: 2 bristles on the left.

Supplemental References:

- S1. Campuzano, S., Carramolino, L., Cabrera, C.V., Ruíz-Gómez, M., Villares, R., Boronat, A., and Modolell, J. (1985). Molecular genetics of the achaete-scute gene complex of *D. melanogaster*. *Cell* 40, 327–338.
- S2. Carramolino, L., Ruiz-Gomez, M., del Carmen Guerrero, M., Campuzano, S., and Modolell, J. (1982). DNA map of mutations at the scute locus of *Drosophila melanogaster*. *The EMBO journal* 1, 1185.
- S3. Cubas, P., De Celis, J.F., Campuzano, S., and Modolell, J. (1991). Proneural clusters of achaete-scute expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. *Genes & Development* 5, 996–1008.