Correlated Evolution of two Sensory Organs via a Single Cis-Regulatory Nucleotide Change

Olga Nagy[1], Isabelle Nuez[1], Rosina Savisaar[1†], Alexandre E. Peluffo[1], Amir Yassin[2], Michael Lang[1], David L. Stern[3], Daniel R. Matute[4], Jean R. David[2,5] and Virginie Courtier-Orgogozo[1]

[1]Institut Jacques Monod, CNRS UMR7592, Univ. Paris-Diderot, 75013 Paris, France.

[2]Institut Systématique Évolution Biodiversité (ISYEB), Centre National de Recherche Scientifique, MNHN, Sorbonne Université, EPHE, 57 rue Cuvier, CP 50, 75005 Paris, France

[3]Janelia Research Campus, 19700 Helix Drive, Ashburn, Virginia 20147.

[4]Biology Department, University of North Carolina, Chapel Hill, USA.

[5]Laboratoire Evolution, Génomes, Comportement, Biodiversité (EGCE), CNRS, IRD, Univ. Paris-sud, Univ. Paris-Saclay, 91198 Gif-sur-Yvette, France.

*Correspondence: virginie.courtier@normalesup.org, @Biol4Ever, +33 1 57 27 80 43

†current address: The Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK.

One Sentence Summary: We identify one nucleotide substitution in a gene regulatory region contributing to evolutionary change of two distinct organs.

SUMMARY

Diverse traits often covary between species. The possibility that a single mutation could contribute to the evolution of several characters between species is rarely investigated as relatively few cases are dissected at the nucleotide level. *Drosophila santomea* has evolved additional sex comb sensory teeth on its legs and has lost two sensory bristles on its genitalia. We found that a single nucleotide substitution in an enhancer of the *scute* gene contributes to both changes. The mutation alters a binding site for the Hox protein Abdominal-B in the developing genitalia, leading to bristle loss, and for another factor in the developing leg, leading to bristle gain. Our study shows that morphological evolution between species can occur through a single nucleotide change affecting several sexually dimorphic traits.

RESULTS AND DISCUSSION

"*Variability is governed by many unknown laws, of which correlated growth is probably the most important*."[1]

Correlated evolution of traits is widespread among taxa [1,2] and can be due to pleiotropy, where a single locus causally affects several traits [3]. Pleiotropy imposes large constrains on the paths of evolution [4,5], making it crucial to assess the extent of pleiotropy to understand the evolutionary process. Empirical studies suggest that many loci influence multiple traits [3,6,7] and current data cannot reject the idea that all genetic elements have pleiotropic roles [3,8,9]. Several pleiotropic substitutions have been associated with natural variation [10–12]: all are coding changes and all underlie intraspecific changes (www.gephebase.org). Nevertheless it remains unclear whether pleiotropic mutations contribute also to interspecific evolution, as experimental evidence suggests that the mutations responsible for interspecies evolution may be less pleiotropic than the mutations underlying intraspecific variation [13].

Here we focused on male sexual bristle evolution between *Drosophila yakuba* and *Drosophila santomea*, which diverged approximately 0.5-1 million years ago [14] and can produce fertile F1 females in the laboratory [15], facilitating genetic mapping. We found that hypandrial bristles – two prominent mecanosensory bristles located on the ventral part of male genitalia in all *D. melanogaster* subgroup species – are missing in *D. santomea* males (Fig. 1). Examination of many inbred stocks and 10 closely related species revealed that the absence of hypandrial bristles is a derived *D. santomea*-specific trait (Fig. 1, Tables S1-S2). No other genital bristle type was noticeably variable in number between *D. yakuba* and *D. santomea* (Fig. S1).

We performed whole-genome QTL mapping between *D. santomea* and *D. yakuba* and found that the left tip of chromosome X has a large effect on hypandrial bristle number (confidence interval = 7 Mb for the *D. santomea* backcross and 2.6 Mb for the *D. yakuba* backcross, Fig. 2A). Duplication mapping in rare *D. santomea-D. melanogaster* hybrid males narrowed down the causal region to a 84.6 kb region of the *achaete-scute* complex (*AS-C*) (Fig. 2B-C, Table S2).

The *AS-C* locus contains four genes, but only two, *achaete* (*ac*) and *scute* (*sc*), are required for bristle formation [16]. Both genes are co-expressed, share *cis*-regulatory elements and act redundantly to specify bristles [17,18]. The elaborate expression pattern of *ac* and *sc* genes prefigures the adult bristle pattern and is controlled by numerous *cis*-regulatory elements [17]. We tested which of the two genes, *ac* or *sc*, contributes to loss of bristles using null mutants in *D. melanogaste*r. All ac^{CAMI} null mutant males had 2 hypandrial bristles (n=15) and sc^{M6} null mutants had none (n=15), indicating that *sc* is required for hypandrial bristle development.

We detected 64 nucleotide differences in the *sc* coding region between *D. yakuba* and *D. santomea*, and all were synonymous substitutions, indicating that coding changes in *sc* are not responsible for the evolved function of *sc*. Using molecularly mapped chromosomal aberrations, we identified a 5-kb region located 46 kb downstream of the *sc* promoter that is required in *D. melanogaster* for hypandrial bristle development (Fig. S2, Tables S3-4). Independently we screened 55 *GAL4* reporter constructs tiling the entire *AS-C* locus and identified three *GAL4* lines (*15E09*, *054839* and *18C05*) that drive expression in hypandrial bristles (Fig. 2C, Tables S5-6). Only one of these lines, *18C05*, increased hypandrial bristle number with *UAS-scute* in a *sc* mutant background or in a sc^{+} background (Fig. 2C, Fig. S3E-P). The 2036-bp *18C05* region is located within the 5-kb candidate region identified with *ac-sc* structural mutations (Fig. 2C),

suggesting that *18C05* is a good candidate region for hypandrial bristle evolution.

To test whether loss of hypandrial bristles in *D. santomea* resulted from changes(s) in the *18C05 cis*-regulatory region, we assayed whether orthologous *18C05* regions from *D. melanogaster*, *D. yakuba* and *D. santomea* driving a *sc* coding region could rescue hypandrial bristles in a *D. melanogaster sc* mutant (Table S7-8). The *D. melanogaster 18C05* enhancer rescued two bristles in both sc^{29} and sc^{M6} mutant backgrounds, indicating that this construct mimics normal levels of *sc* expression. The *D. yakuba 18C05* enhancer rescued on average 2 hypandrial bristles in sc^{M6} and 0.5 bristles in sc^{29} whereas the *D. santomea 18C05* enhancer rescued significantly fewer bristles (1.1 in sc^{M6} and 0 bristles in sc^{29}, Fig. 3). For another measure of *18C05* enhancer activity, we compared the ability of enhancer-*GAL4* constructs containing the *18C05* region from *D. melanogaster*, *D. yakuba* or *D. santomea* to induce extra bristles in *sc* mutants using the *UAS-GAL4* system with *UAS-sc*. In this assay the *D. santomea 18C05* region also induced fewer bristles than the corresponding *D. yakuba* region (Fig. 3, GLM-Quasi-Poisson, $F(19, 509) = 161.7$, $p < 10^{-5}$ for sc^{29}; $F(19, 415) = 125.9$, $p < 10^{-5}$ for sc^{M6}). Together, these results suggest that changes(s) within *18C05* contributed to hypandrial bristle evolution in *D. santomea*.

To narrow down the region responsible for hypandrial bristle loss, we dissected the *18C05* element from *D. melanogaster*, *D. yakuba* and *D. santomea* into smaller overlapping pieces and quantified their ability to produce hypandrial bristles with the *GAL4* rescue experiment. For all three species we found that smaller segments rescued significantly fewer bristles than the corresponding full region (Fig. S4-5). Thus, transcription factor binding sites scattered throughout the entire ~2 kb of the *18C05* element are required to drive full expression in the hypandrial bristle region.

Sequence alignment of the *18C05* region from multiple species revealed 11 substitutions and one indel that are fixed and uniquely derived in *D. santomea*. Among them, seven substitutions altered sites that are otherwise conserved in the *D. melanogaster* subgroup (Fig. S6-7). We tested the effect of these seven *D. santomea*-specific nucleotide changes by introducing them one at a time or all together, into either a *D. yakuba 18C05* enhancer or into the inferred ancestral enhancer driving *sc* expression (Tables S7-11). The ancestral *18C05* sequence was resurrected by reverting the *D. santomea*-specific and *D. yakuba*-specific mutations to their ancestral states and it produced the same number of bristles as the *D. yakuba* construct (Fig. 3, Fig. S8). Four substitutions (*G869A*, *T970A*, *T1008C* and *T1482C*) had no effect, whether in the *D. yakuba* or in the ancestral background (GLM-Quasi-Poisson, $p>0.6$). Three substitutions (*T1429G*, *A1507G* and *T1775G*) decreased the number of rescued bristles in both the *D. yakuba* and the ancestral sequence, and these effects were highly significant, except for *A1507G* in the *D. yakuba* background, which was slightly above statistical threshold (using the most stringent correction method) (Fig. 3). These results are consistent with analysis of smaller pieces of *18C05* and of *18C05* chimeric constructs containing DNA fragments from *D. yakuba* and *D. santomea* (Tables S7-11, Fig. S8). When combined into the *D. yakuba* background, the seven *D. santomea*-specific substitutions rescued the same number of bristles as the *D. santomea 18C05* construct (Fig. 3, Fig. S8, GLM-Quasi-Poisson, $p>0.9$ in sc^{M6}). We conclude that at least three fixed substitutions within a 350-bp enhancer of *sc* contribute to the reduction in hypandrial bristle number in *D. santomea*.

Analysis of *18C05-GAL4* and *18C05-GFP* reporter constructs revealed that the *18C05* region drives expression not only in male genital discs but also in male foreleg imaginal discs [19] (Fig. S9). Male forelegs harbor sex combs, sensory organs used for grasping the female

during copulation [20], which differ in bristle number between *D. santomea* and *D. yakuba*, and the difference maps to the X chromosome [21], where *sc* is located. These results prompted us to test whether the mutations contributing to hypandrial bristle evolution also affect sex combs. Sex comb tooth number was reduced in sc^{M6} mutants and significantly rescued with *yak18C05-sc* (Fig. 4, Fig. S10). Analysis of sc^{M6} mutants rescued with the *yak18C05-sc* constructs containing the *D. santomea*-specific substitutions showed that *T1429G* and *T1507G* have no effect and that *T1775G* increases the number of sex comb teeth (Fig. 4). We conclude that the *T1775G* substitution contributes to both the increase in sex comb tooth number and the loss of hypandrial bristles.

A bioinformatics search revealed that the *T1775G* substitution is predicted to alter a binding site for the Hox protein Abdominal-B (Abd-B) (Table S12). *Abd-B* is expressed only in the posterior part of the fly, where it directs the development of posterior-specific structures such as the genitalia [22]. We found that reducing *Abd-B* expression, using either genetic mutations or RNA interference, resulted in loss of hypandrial bristles (Fig. S11), indicating that normal levels of *Abd-B* expression are required for hypandrial bristle development. Electrophoretic mobility shift assays showed that Abd-B proteins bind more strongly to a 54-bp fragment of the *18C05* sequence containing the *D. yakuba*-specific T at position 1775 than the *D. santomea*-specific G at this position (Fig. S12). These results are consistent with the hypothesis that the *T1775G* substitution decreases ABD-B binding, contributing to reduction in *sc* expression levels, and ultimately reducing the number of hypandrial bristles. Since *Abd-B* is not expressed in developing legs, *T1775G* is expected to affect binding of other factors to increase sex comb tooth number. Overall, our study suggests that *T1775G* alters overlapping binding sites for distinct factors in the leg and the genitalia (Fig. 4E).

Intriguingly, the two organs affected by substitution *T1775G* – hypandrial bristles and sex combs – may both aid the male to position himself on top of the female during copulation [20,23]. Genitals are the most rapidly evolving organs in animals with internal fertilization [24]. To our knowledge, only two other mutations contributing to the evolution of genital anatomy are known. First, a 61-kb-deletion of a cis-regulatory region of the *androgen receptor (AR)* gene in humans is associated with loss of keratinized penile spines in humans compared to chimpanzees [25]. Second, an amino acid change in the *acetyltransferase* gene which probably appeared recently in laboratory strains of the nematode *C. elegans*, alters morphology in the presence of some mutations but not in a wild-type genetic background [10]. Both mutations appear to be pleiotropic: the *AR* deletion is associated with loss of facial vibrissae in humans and the *nath10* mutation affects egg and sperm production as well.The paucity of known mutations responsible for genital evolution makes it currently difficult to propose general rules for the causes of rapid genital evolution. Our results are reminiscent of Mayr's pleiotropy hypothesis [26], which posits that certain characters may evolve arbitrarily as a result of selection on other traits due to pleiotropic mutations.

We report here the first case of an evolved cis-regulatory substitution with pleiotropic effects. Given the large number of bristle types regulated by *sc* (>100 in adult flies), it is possible that no cis-regulatory mutation in *sc* can affect only one bristle type. Our results challenge the idea that cis-regulatory enhancers are tissue-specific modules underlying evolutionary changes in targeted traits [27]. Enhancer sequences evolve rapidly, with rapid turn over of individual binding sites while maintaining transcriptional output over millions of years by compensatory mutations [28]. Since pleiotropic mutations can have deleterious off-target effects, we propose that evolution of pleiotropic sites within enhancers should trigger the subsequent selection of

compensatory mutations in cis, thus contributing to rapid divergence of cis-regulatory sequences. Overall, our results suggest that pleiotropic cis-regulatory mutations may play a more important role in evolution than previously thought.

REFERENCES

1. Darwin, C. (1859). On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.

2. Saltz, J.B., Hessel, F.C., and Kelly, M.W. (2017). Trait Correlations in the Genomics Era. Trends Ecol. Evol. (Amst.) *32*, 279–290.

3. Paaby, A.B., and Rockman, M.V. (2013). The many faces of pleiotropy. Trends in Genetics *29*, 66–73.

4. Fisher, R.A. (1930). The genetical theory of natural selection (Oxford: Clarendon).

5. Orr, H.A. (2000). Adaptation and the cost of complexity. Evolution *54*, 13–20.

6. Wagner, G.P., and Zhang, J. (2011). The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. Nature Reviews Genetics *12*, 204–213.

7. Stearns, F.W. (2010). One Hundred Years of Pleiotropy: A Retrospective. Genetics *186*, 767–773.

8. Lonfat, N., Montavon, T., Darbellay, F., Gitto, S., and Duboule, D. (2014). Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. Science *346*, 1004–1006.

9. Preger-Ben Noon, E., Sabarís, G., Ortiz, D.M., Sager, J., Liebowitz, A., Stern, D.L., and Frankel, N. (2018). Comprehensive Analysis of a cis -Regulatory Region Reveals Pleiotropy in Enhancer Function. Cell Reports *22*, 3021–3031.

10. Duveau, F., and Félix, M.-A. (2012). Role of pleiotropy in the evolution of a cryptic developmental variation in Caenorhabditis elegans. PLoS Biol. *10*, e1001230.

11. Chang, S.H., Jobling, S., Brennan, K., and Headon, D.J. (2009). Enhanced Edar Signalling Has Pleiotropic Effects on Craniofacial and Cutaneous Glands. PLOS ONE *4*, e7591.

12. Kent, C.F., Daskalchuk, T., Cook, L., Sokolowski, M.B., and Greenspan, R.J. (2009). The Drosophila foraging Gene Mediates Adult Plasticity and Gene–Environment Interactions in Behaviour, Metabolites, and Gene Expression in Response to Food Deprivation. PLOS Genetics *5*, e1000609.

13. Wittkopp, P.J., Haerum, B.K., and Clark, A.G. (2008). Regulatory changes underlying expression differences within and between Drosophila species. Nature Genetics *40*, 346–350.

14. Turissini, D.A., and Matute, D.R. (2017). Fine scale mapping of genomic introgressions within the Drosophila yakuba clade. PLoS Genetics *13*, e1006971.

15. Lachaise, D., Harry, M., Solignac, M., Lemeunier, F., Bénassi, V., and Cariou, M.L. (2000). Evolutionary novelties in islands: Drosophila santomea, a new melanogaster sister species from São Tomé. Proc. Biol. Sci. *267*, 1487–1495.

16. Simpson, P., Woehl, R., and Usui, K. (1999). The development and evolution of bristle patterns in Diptera. Development *126*, 1349–1364.

17. Gómez-Skarmeta, J.L., Rodríguez, I., Martínez, C., Culí, J., Ferrés-Marcó, D., Beamonte, D., and Modolell, J. (1995). Cis-regulation of achaete and scute: shared enhancer-like elements drive their coexpression in proneural clusters of the imaginal discs. Genes Dev *9*, 1869–1882.

18. Marcellini, S., Gibert, J.-M., and Simpson, P. (2005). achaete, but not scute, is dispensable for the peripheral nervous system of Drosophila. Dev. Biol. *285*, 545–553.

19. Jory, A., Estella, C., Giorgianni, M.W., Slattery, M., Laverty, T.R., Rubin, G.M., and Mann, R.S. (2012). A Survey of 6,300 Genomic Fragments for cis-Regulatory Activity in the Imaginal Discs of Drosophila melanogaster. Cell Reports *2*, 1014–1024.

20. Ng, C.S., and Kopp, A. (2008). Sex combs are important for male mating success in Drosophila melanogaster. Behavior genetics *38*, 195.

21. Coyne, J.A., Elwyn, S., Kim, S.Y., and Llopart, A. (2004). Genetic studies of two sister species in the Drosophila melanogaster subgroup, D. yakuba and D. santomea. Genetics Research *84*, 11–26.

22. Foronda, D., Estrada, B., de Navas, L., and Sánchez-Herrero, E. (2006). Requirement of Abdominal-A and Abdominal-B in the developing genitalia of Drosophila breaks the posterior downregulation rule. Development *133*, 117–127.

23. Hurtado-Gonzales, J.L., Gallaher, W., Warner, A., and Polak, M. (2015). Microscale Laser Surgery Demonstrates the Grasping Function of the Male Sex Combs in Drosophila melanogaster and Drosophila bipectinata. Ethology *121*, 45–56.

24. Eberhard, W.G. (1988). Sexual Selection and Animal Genitalia (Harvard University Press).

25. McLean, C.Y., Reno, P.L., Pollen, A.A., Bassan, A.I., Capellini, T.D., Guenther, C., Indjeian, V.B., Lim, X., Menke, D.B., Schaar, B.T., *et al.* (2011). Human-specific loss of regulatory DNA and the evolution of human-specific traits. Nature *471*, 216–219.

26. Mayr, E. (1963). Animal species and evolution (Harvard University Press).

27. Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell *134*, 25–36.

28. Cheng, Y., Ma, Z., Kim, B.-H., Wu, W., Cayting, P., Boyle, A.P., Sundaram, V., Xing, X., Dogan, N., Li, J., *et al.* (2014). Principles of regulatory information conservation between mouse and human. Nature *515*, 371.

29. Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T.T., Mast, J., Sunayama-Morita, T., and Stern, D.L. (2011). Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Res. *21*, 610–617.

30. Broman, K.W., and Sen, S. (2009). A Guide to QTL Mapping with R/qtl 1st ed. (Springer).

31. Broman, K.W., Wu, H., Sen, S., and Churchill, G.A. (2003). R/qtl: QTL mapping in experimental crosses. Bioinformatics *19*, 889–890.

32. Haley, C.S., and Knott, S.A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity (Edinb) *69*, 315–324.

33. Venken, K.J.T., Popodi, E., Holtzman, S.L., Schulze, K.L., Park, S., Carlson, J.W., Hoskins, R.A., Bellen, H.J., and Kaufman, T.C. (2010). A molecularly defined duplication set for the X chromosome of Drosophila melanogaster. Genetics *186*, 1111–1125.

34. Turissini, D.A., McGirr, J.A., Patel, S.S., David, J.R., and Matute, D.R. (2017). The rate of evolution of postmating-prezygotic reproductive isolation in Drosophila. Mol. Biol. Evol.

35. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature methods *9*, 671–675.

36. Taylor, B.J. (1989). Sexually dimorphic neurons in the terminalia of Drosophila melanogaster: I. Development of sensory neurons in the genital disc during metamorphosis. J. Neurogenet. *5*, 173–192.

37. Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., *et al.* (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature *448*, 151–156.

38. Jenett, A., Rubin, G.M., Ngo, T.-T., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B.D., Cavallaro, A., Hall, D., Jeter, J., *et al.* (2012). A GAL4-driver line resource for Drosophila neurobiology. Cell reports *2*, 991–1001.

39. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature methods *6*, 343–345.

40. Bryne, J.C., Valen, E., Tang, M.-H.E., Marstrand, T., Winther, O., da Piedade, I., Krogh, A., Lenhard, B., and Sandelin, A. (2007). JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. Nucleic acids research *36*, D102–D106.

41. Zhu, L.J., Christensen, R.G., Kazemian, M., Hull, C.J., Enuameh, M.S., Basciotta, M.D., Brasefield, J.A., Zhu, C., Asriyan, Y., Lapointe, D.S., *et al.* (2010). FlyFactorSurvey: a database of Drosophila transcription factor binding specificities determined using the bacterial one-hybrid system. Nucleic acids research *39*, D111–D117.

42. Crawley, M.J. (2012). The R book (John Wiley & Sons).

43. Hilbe, J.M. (2014). Modeling Count Data (Cambridge University Press).

44. Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A., and Smith, G.M. (2011). Mixed Effects Models and Extensions in Ecology with R Softcover reprint of hardcover 1st ed. 2009 edition. (New York, NY: Springer).

45. Team, R.C. (2016). R: A language and environment for statistical (Vienna, Austria) Available at: http://www.R-project.org/.

46. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

47. Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. Biometrical journal *50*, 346–363.

48. Bretz, F., Hothorn, T., and Westfall, P. (2010). Multiple comparisons using R (CRC Press).

49. Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian journal of statistics, 65–70.

50. Frangioni, J.V., and Neel, B.G. (1993). Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. Analytical biochemistry *210*, 179–187.

51. Fan, Y.-J., Gittis, A.H., Juge, F., Qiu, C., Xu, Y.-Z., and Rabinow, L. (2014). Multifunctional RNA Processing Protein SRm160 Induces Apoptosis and Regulates Eye and Genital Development in Drosophila. Genetics *197*, 1251–1265.

52. Chatterjee, S.S., Uppendahl, L.D., Chowdhury, M.A., Ip, P.-L., and Siegal, M.L. (2011). The female-specific doublesex isoform regulates pleiotropic transcription factors to pattern genital development in Drosophila. Development *138*, 1099–1109.

53. Skaer, N., Pistillo, D., and Simpson, P. (2002). Transcriptional heterochrony of scute and changes in bristle pattern between two closely related species of blowfly. Developmental biology *252*, 31–45.

54. Casanova, J., Sánchez-Herrero, E., and Morata, G. (1986). Identification and characterization of a parasegment specific regulatory element of the abdominal-B gene of Drosophila. Cell *47*, 627–636.

55. Hopmann, R., Duncan, D., and Duncan, I. (1995). Transvection in the iab-5, 6, 7 region of the bithorax complex of Drosophila: homology independent interactions in trans. Genetics *139*, 815–833.

56. Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223–1237.

57. Estrada, B., and Sánchez-Herrero, E. (2001). The Hox gene Abdominal-B antagonizes appendage development in the genital disc of Drosophila. Development *128*, 331–339.

58. Maroni, G. and S.C. Stamey (1983). Developmental profile and tissue distribution of alcohol dehydrogenase. Drosophila Information Service *59*.

59. Andres, A.J., and Thummel, C.S. (1994). Methods for quantitative analysis of transcription in larvae and prepupae. Methods Cell Biol. *44*, 565–573.

60. 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.

61. 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.

62. 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.

63. Kvon, E.Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J.O., Pagani, M., Schernhuber, K., Dickson, B.J., and Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature *512*, 91–95., 91–95.

ACKNOWLEDGEMENTS

We thank the Tucson Drosophila Species Stock Center, the VDRC, Kyoto and Bloomington Stock Centers for flies. We thank São Tomé authorities for allowing us to collect flies. We thank S. Picard for help with MSG, E. Sánchez-Herrero for *Abd-B* flies, J. Selegue and S.B. Carroll for the *Abd-B* construct, L. Pintard and N. Joly for help in protein purification, R. Mann for EMSA suggestions, F. Mallard and T. Tully for advices on the statistical analyses, J. L. Villanueva-Cañas for help with JASPAR. We thank Q.D. Tran, M. Notin, V. Ludger, A. Matamoro-Vidal, S.

El Ouisi, A. La, F. Foutel-Rodier, C. Guillard-Sirieix and A. Aydogan for their contributions and C. Desplan, D. Petrov, B. Prud'homme, A. Martin and J.A. Lepesant for comments on the manuscript. We acknowledge the ImagoSeine core facility of the Institut Jacques Monod, member of IBiSA and France-BioImaging (ANR-10-INBS-04) infrastructures. We thank the Courtier-Orgogozo team for providing a stimulating environment and technical support.

FUNDING

The research leading to this paper has received funding from the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant Agreement no. 337579) to VCO, from the labex "Who am I?" (ANR-11-LABX-0071) funded by the French government through grant no. ANR-11-IDEX-0005-02 to AEP and from NIH (1R01GM121750) to DRM.

AUTHOR CONTRIBUTIONS

J.R.D. found that *D. santomea* lacks hypandrial bristles and that the trait difference is X-linked, D.L.S. genotyped flies with MSG, A.Y., I.N. and V.C.O. performed the QTL mapping experiment, D.R.M. made the *D. santomea-D. melanogaster* hybrids, I.N. dissected them, O.N. did all other fly crosses and dissected them, O.N., I.N., R.S. and A.E.P. phenotyped >3000 males for hypandrial bristles, O.N. phenotyped all other bristles, O.N. and M.L. did EMSA, O.N. and I.N. constructed the plasmids, O.N. performed immunostainings and microscopy, A.E.P. performed all statistical analyses with feedback from O.N. and M.L., D.R.M. collected wild flies, V.C.O. supervised research, performed bioinformatics sequence analysis and wrote the paper with O.N. All authors provided feedback on the text.

DECLARATION OF INTEREST

The authors declare no competing interests.

Data and materials availability: Sequences were deposited into GenBank (accession numbers MG460736-MG460765). Source data for Bristle Number and QTL mapping analysis are available as Auxiliary Supplementary Materials.

Supplementary Materials:

Materials and Methods

Figures S1-S12

Tables S1-S13

References (25-59)

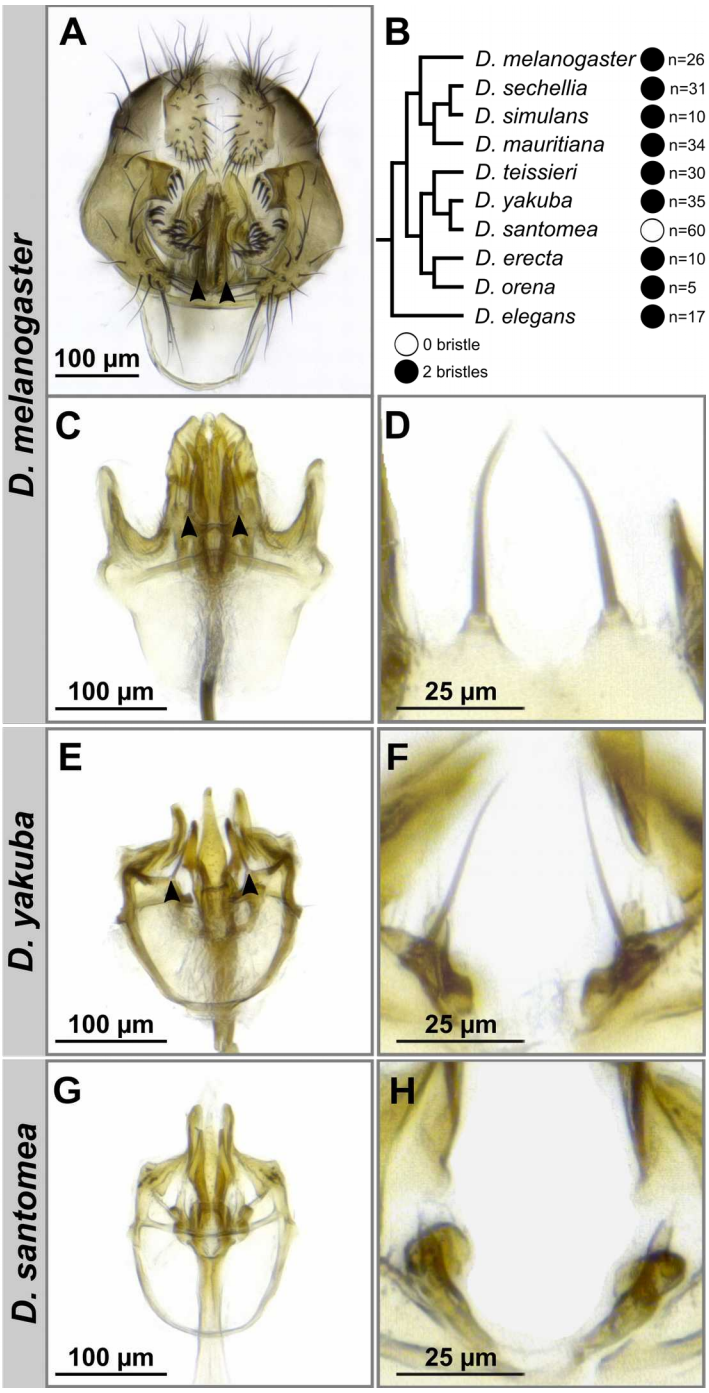

Fig. 1. *D. santomea* lost hypandrial bristles.

(A) *Drosophila melanogaster* male genitalia. (B) Phylogeny of the *Drosophila melanogaster* species subgroup. All species of have two hypandrial bristles (black circles) except *Drosophila santomea*, which lacks hypandrial bristles (white circle). (C-H) Light microscope preparations of ventral genitalia (B,D,F) and hypandrial bristles (D,F,H) in *D. melanogaster* (C-D), *D. yakuba* (E-F) and *D. santomea* (G-H). Hypandrial bristles are indicated with arrowheads on A, C and E. n: number of scored individuals.

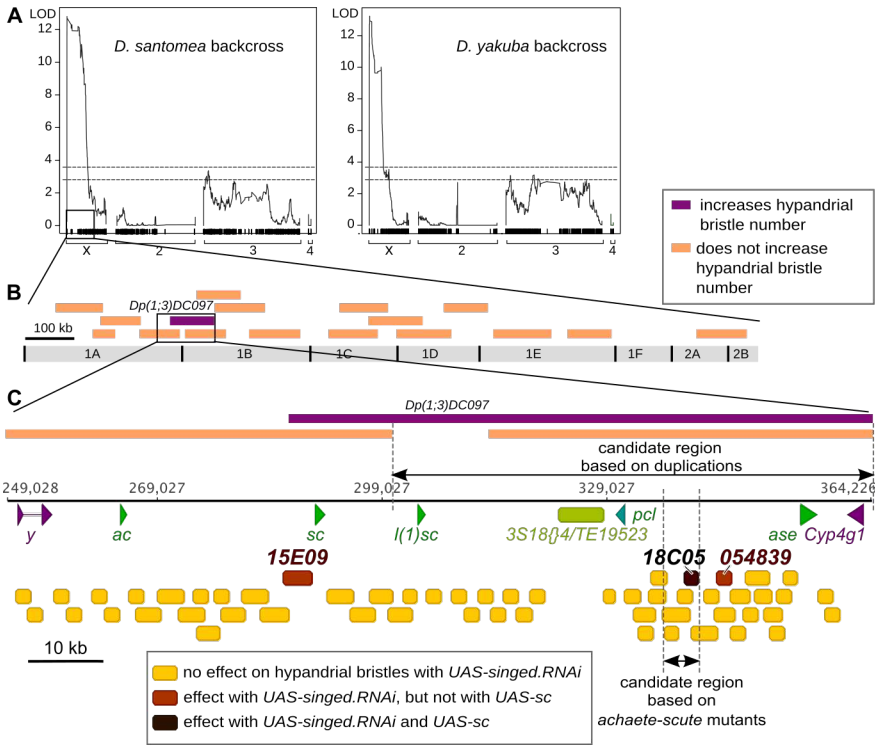

Fig. 2. Mapping of the cis-regulatory element involved in hypandrial bristle evolution. (A) QTL analysis of hypandrial bristle number in a *D. santomea* backcross (left) and a *D. yakuba* backcross (right). On the y-axis are the LOD profiles from a Haley-Knott regression analysis. The x-axis represents physical map position in the *D. yakuba* genome. Ticks represent recombination informative markers. Dotted lines represent the 1% (top) and 5% (bottom) significance thresholds. (B) Schematic representation of the left tip of chromosome X and of 19 duplicated fragments of chromosome X that were tested for their effect on hypandrial bristle number in D. santomea-*D. melanogaster* hybrid males. All duplications had no significant effect (orange) except *Dp(1;3)DC097* (purple), which significantly increased hypandrial bristle number. (C) Genomic organization of the *AS-C* locus in *D. melanogaster*. Arrows 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. The light green box represents the insertion of a *3S18{}4/TF9523* natural transposable element. Boxes indicate cis-regulatory elements whose corresponding *GAL4* reporter lines have been tested. Expression of *UAS-singed.RNAi* with 52 *GAL4* lines (yellow boxes) has no effect while it results in singed hypandrium bristles with *15E09-*, *18C05-* and *054839-GAL4*. Extra hypandrial bristles are found with *UAS-sc* and *18C05-GAL4* (dark brown box) but not with *15E09-* and *054839-GAL4* (light brown boxes).

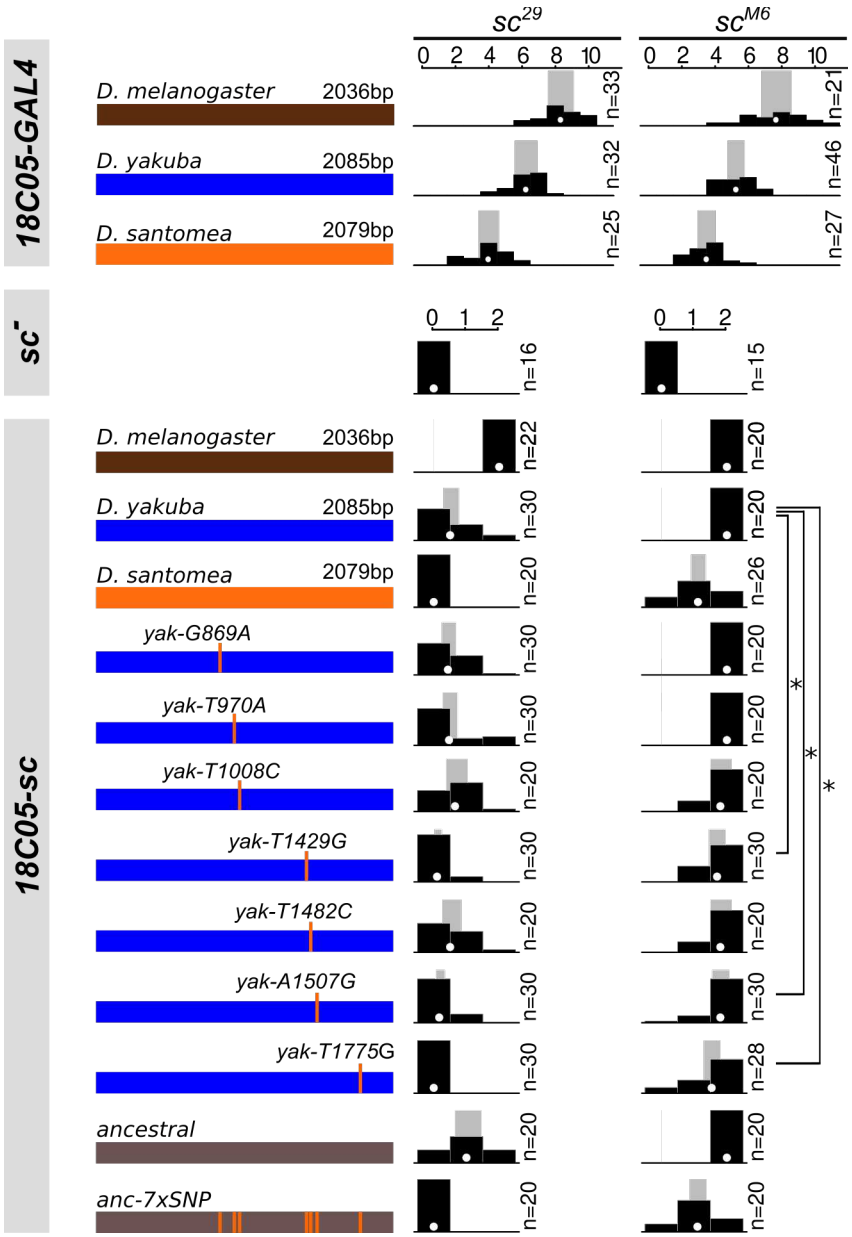

Fig. 3. Three *D. santomea*-specific substitutions in *18C05* contribute to the loss of hypandrial bristles.

Rescue of the hypandrial bristle loss of sc^{29} (left column) and sc^{M6} (right column) *D. melanogaster* mutants by expression of either *GAL4* with *UAS-sc* or *sc* driven by *18C05* sequences from *D. melanogaster* (brown), *D. yakuba* (blue) and *D. santomea* (orange). Seven *D. santomea*-specific substitutions (vertical orange bars) were introduced into either the *D. yakuba* region (blue) or the ancestrally reconstructed *18C05* region (grey). 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. *: $p<0.05$

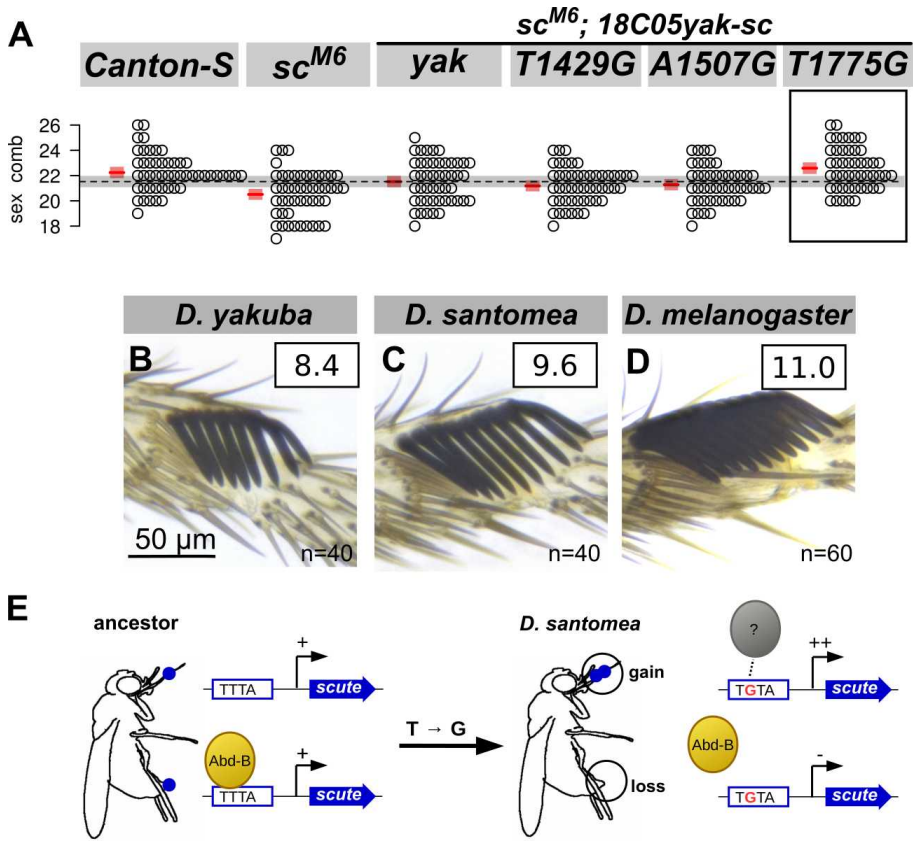

Fig. 4. *D. santomea*-specific substitution T1775G contributes to increase in sex comb tooth number.

(A) Sex comb number in wild-type (Canton-S), sc^{M6} mutants and sc^{M6} mutants rescued with different *18C05-sc* constructs. Each circle represents one male raised at 25°C. Mean (red line) and 95% confidence interval (red rectangle) from a fitted GLM Quasi-Poisson model are shown. For easier comparison, the horizontal dashed line and the surrounding grey line indicate the mean and 95% confidence interval for sc^{M6}*;18C05yak-sc*. (B-D), Leg sex comb in *D. yakuba* (B), *D. santomea* (C) and *D. melanogaster* (D). Average sex comb tooth numbers are shown in squares. n: number of scored individuals. (E) How a single nucleotide change contributes to the evolution of both leg and genital sensory organs in *D. santomea*.

Supplemental Information

Supplemental Information includes:
in this PDF:

- Materials and Methods
- Figs. S1 to S12
- Tables S1 to S13
- Captions for Data S1 to S9

in extra files:

- Data S1 to S9 [QTL-data, Hypandrial bristle number with *18C05-GAL4* constructs, Hypandrial bristle number with *18C05-sc* constructs, Sex comb tooth number with various *18C05yak-sc* constructs, Genital bristle numbers in *D. yakuba* and *D. santomea* strains, Hypandrial bristle number with *18C05 melanogaster-GAL4* constructs, Hypandrial bristle number with chimeric constructs, Sex comb tooth numbers in *D. yakuba* and *D. santomea* strains, Fractional occupancy data]

Materials and Methods

Fly stocks and crosses

The origin of all the fly strains used can be found in Tables S1,3,5-6. All flies were cultured on standard cornmeal–agar medium in uncrowded conditions at 25°C unless stated. We used *Canton-S* as a wild-type *D. melanogaster* strain. Transgenic constructs were integrated into the *attP2* landing site in *D. melanogaster* w^{1118} by BestGene Inc. Hybrid males between *D. yakuba* and *D. santomea* were obtained by collecting 20 virgin females with 20 males from each stocks and crossing them reciprocally in both directions. At least 10 such crosses were made and flipped every 4-5 days for several weeks. For QTL mapping, *D. yakuba yellow*[1] virgin females were crossed *en masse* to *D. santomea* SYN2005 males to generate F1 hybrid females, which were subsequently backcrossed, separately, to both parental strains. Genitalia of backcross males were isolated for dissection and the remaining carcass was stored at -20 °C for subsequent sequencing library preparation.

Genotyping of backcross males for QTL mapping

The carcass of each male was crushed in a 1.5-ml Eppendorf tube with a manual pestle in 180 µl of Qiagen Tissue Lysis buffer. DNA of individual flies was extracted using Qiagen DNeasy Blood & Tissue extraction kit (cat #69506). A Multiplexed Shotgun Genotyping sequencing library was made from 189 *D. santomea* backcross males and for 181 *D. yakuba* backcross males as described previously [29]. The list of barcodes used in this study are provided in Supplemental Data File 1, within the names of the individuals that were sequenced. *D. yakuba* and *D. santomea* parental genome sequences were generated by updating the *D. yakuba* genome sequence dyak-4-chromosome-r1.3.fasta with Illumina paired-end reads from *D. yakuba yellow[1]* and *D. santomea* SYN2005 (sequenced by BGI) using the msgUpdateParentals.pl function of the MSG software package. The resulting updated genome files are dsan-all-chromosome-yak1.3-r1.0.fasta.msg.updated.fasta and dyak-4-chromosome-r1.3.fasta.msg.updated.fasta. Ancestry was estimated for all backcross progeny using MSG software (github.com/YourePrettyGood/msg). Ancestry files were reduced to only those markers informative for recombination events using the script pull_thin_tsv.py (github.com/dstern/pull_thin). Markers were considered informative when the conditional probability of being homozygous differed by more than 0.05 from their neighboring markers.

QTL mapping

QTL mapping was performed using the R/qtl package version 1.4 [30,31]. The thinned posterior genotype probabilities were imported into R/qtl using the R function read.cross.msg.1.5.R (github.com/dstern/read_cross_msg). QTL mapping was performed independently on each backcross population. We performed genome scans with a single QTL model (“scanone”) using the Haley-Knott regression method [32] performs well with genotype information at a large number of positions along the genome. The genome-wide 5% and 1% significance levels were determined using 1,000 permutations. One QTL peak above the 1% significance level was found for both backcrosses. To check for additional QTL, we built a QTL model with this single QTL using the “fitqtl” function and scanned for additional QTL using the “addqtl” function. A second QTL was found on chromosome 3 for both backcrosses. When introduced into a new multiple QTL model, refined and fitted to account for possible interactions, a third significant QTL was

found. Based on the full three-QTL model, no additional significant QTL were found with the function "addqtl": the highest LOD score for a fourth QTL reached only 1.8 and 1.2 for the *D. yakuba* backcross and the *D. santomea* backcross, respectively. Various three-QTL models with different interactions between loci were assessed. Positive significant interaction was detected between the QTL on chromosome 1 and both QTLs on chromosome 3. The interaction between the two QTLs on chromosome 3 was not significant. For the three-QTL model with interactions between the QTL on chromosome 1 and both QTLs on chromosome 3, we computed the LOD score of the full model and the estimated effects of each locus. The 2-LOD intervals were calculated using the "lodint" function with parameter drop of 2. Analysis of F1 hybrid males is consistent with a large effect of the X chromosome on hypandrial bristle number: male F1 hybrids carrying a *D. yakuba* X chromosome have on average 1.9 hypandrial bristles (n=34) while reciprocal hybrid males possessing the *D. santomea* X chromosome have none (n=29) (Table S2).Note that few informative markers are found on the right arm of chromosome 2, suggesting the presence of an inversion between parental lines. In both backcrosses the large-effect QTL is estimated to cause a decrease of 0.9±0.1 bristles between a *D. yakuba* hemizygote and a *D. santomea* hemizygote male (Data S1). The QTL peak is at position 46,886 and 221,928 for the *D. santomea* and *D. yakuba* backcross, respectively. The *AS-C* locus is at position 179,000-290,000.

Duplication Mapping in *D. santomea-D. melanogaster* hybrids

We used a set of *D. melanogaster* duplication lines to test overlapping parts of chromosome X for their effect on hypandrial bristle number [33]. Each line contains a fragment of the chromosome X inserted into the same attP docking site on chromosome 3L using ΦC31 integrase, allowing direct comparison between fragments. Each duplication was used to screen for complementation of the loss of function allele(s) from *D. santomea*. We exploited the fact that rare *D. santomea-D. melanogaster* hybrid males can be produced by crossing *D. melanogaster* females carrying a compound X chromosome with *D. santomea* males [34]. The resulting hybrid males carry a *D. santomea* X chromosome. We first created a *D. melanogaster* stock whose genotype is *TM3, Sb[1] Ser[1]/Nup98-96[339]* by crossing *Nup98-96[339]/TM3, Sb[1]* with *Df(3R)D605/TM3, Sb[1] Ser[1]*. We then performed three successive crosses at room temperature in glass vials: (a) *C(1)RM, y[1] w[1] f[1] ; +/+ × +/+; TM3, Sb[1] Ser[1]/Nup98-96[339]*, (b) *C(1)RM, y[1] w[1] f[1] ; TM3, Sb[1] Ser[1]/+ × +/Y; Dp(1,3)/Dp(1,3)*, (c) *C(1)RM, y[1] w[1] f[1] ; TM3, Sb[1] Ser[1]/Dp(1,3) D. melanogaster* females × *D. santomea* males. The same procedure was followed for 21 duplication lines and progeny was obtained for 17 of them. Hybrid males from the last cross were sorted in two pools, the [Sb^-, Ser^-] males who carried the duplication and the [Sb^+, Ser^+] males which were used as controls which carried no duplication but the balancer chromosome *TM3 Sb[1] Ser[1]*. In *D. melanogaster/D. santomea* hybrids, dominant markers are not always fully penetrant. A few progeny males exhibited [Sb^+, Ser^-] or [Sb^-, Ser^+] phenotypes; they were considered as control individuals carrying the balancer chromosome *TM3, Sb[1] Ser[1]*. Males were stored in ethanol until dissection. Duplication mapping narrowed down the causal region to a 84.6 kb region (*DC097*) of the *achaete-scute* complex (*AS-C*) (Fig. 2.B-C, Table S2, GLM-Poisson, Chisq(17,478)=398.44, $p = 10^{-4}$).

Examination of Hypandrial Bristle Phenotypes

Male genitalia were cut with forceps and then hypandria were dissected with fine needles or forceps Dumont #5 (112525-20, Phymep) in a drop of 1x PBS. For *D. melanogaster* in order to

see the hypandrial bristles better we removed the aedeagus by holding the aedeagal apodem with forceps and gently pushing the hypandrium upwards with an other forceps until it separated. Hypandria were mounted in DMHF (Dimethyl Hydantoin Formaldehyde, Entomopraxis). Before dissection, males were sometimes stored at -20°C in empty Eppendorf tubes or in glycerol:acetate:ethanol (1:1:3) solution. For analysis of non-hypandrial bristles, males were stored at -20°C in glycerol:acetate:ethanol (1:1:3) solution. We never stored these males in empty tubes because we found that such a storage procedure can break and remove external bristles (but, as far as we know, hypandrial bristles were not affected by such a procedure, maybe because hypandrial bristles are relatively internal and protected by the epandrium). Furthermore, we never observed a single socket devoid of shaft on the male hypandrium, indicating that hypandrial bristles cannot be accidentally cut or lost with our experimental protocol. 3D projection images of the preparations were taken at 500X magnification with the Keyence digital microscope VHX 2000 using optical zoom lens VH-Z20R/W.

Examination of Other Bristles

Since genitalia are the most rapidly evolving organs in animals with internal fertilization [32], we compared the number of genital bristles between two strains of *D. yakuba* and two strains of *D. santomea*. We found no difference between *D. yakuba* and *D. santomea* in any genital bristles except for anal plate and clasper bristles, where a slightly significant interspecific variation was detected (Fig. S1.). The loss of hypandrial bristles in *D. santomea* is thus the major change in genital bristles between *D. santomea* and *D. yakuba*. Genitalia were dissected in 1X PBS, hypandria were removed and the epandria were mounted in 99% glycerol. Gentle pressure was applied on the cover-slip with forceps to flatten the preparations in order to see all bristles. Pictures were taken at a 500X magnification with a digital microscope VHX 2000 (Keyence) using lens VH-Z20R/W. Bristles were counted on the images.

For sex comb preparations, prothoracic legs were dissected at the coxa with forceps Dumont #5 and were mounted in DMHF (Dimethyl Hydantoin Formaldehyde, Entomopraxis). Images of the sex combs were taken at 1000x magnification with the Keyence digital microscope as written above. Sex comb teeth were counted on the images. Distances between the sex comb and the central bristle were measured as the distance between the middle of the socket of the central bristle and the most anterior basis of the sex comb. Measurements were performed in Image J [35].

Analysis of *scute* coding sequence

The *scute* coding sequence (CDS) of *D. melanogaster* iso-1 was retrieved from FlyBase. We blasted the updated genome sequences of *D. yakuba yellow[1]* and *D. santomea* SYN2005 (see above) with *D. melanogaster scute* coding region and retrieved only one locus in each species. The *scute* coding region was then annotated with Geneious and no intron was found, as in *D. melanogaster*.

Screening *as-GAL4* lines for expression in the hypandrium

The *as-GAL4* lines were ordered from VDRC and Bloomington Stock Center (Table S6). Two lines were not available (*GMR1509* and *VT054822)* so we created new transgenic lines for these regions, named *GMR15X09-GAL4* and *VT054822b-GAL4* (see below). Because screens are easier on adults than on genital discs, and also because the exact developmental stage and location of hypandrial bristle development are unknown [36], we decided to look *for* GAL4-

triggered phenotypes in adult males. As a readout *of* GAL4 expression, we tested vari*ous* UAS lin*es (UAS-mCD8*-G*FP, UAS-ye*llow i*n a ye*llow mutant backgrou*nd, UAS-sc.*RN*Ai, UAS-achaete.*RN*Ai, UAS-forked.*RN*Ai, UAS-singed.*RNAi) (lines are listed in Table S1 and S5) together w*ith DC-GAL4*, which drives expression in the dorso-central thoracic bristl*e*s [37]. To enhance RNAi potency we also used U*AS-Dicer-2* (3*5)*. With U*AS-mCD8-GFP* and U*AS-yellow* the change in fluorescence or color was hardly visible. The most penetrant bristle phenotype was obtained with U*AS-Dicer-2 UAS-singed.RNAi1*^{105747} at 29 °C (Table S5). Therefore this line was chosen for screening all the a*s-GAL4* constructs.

Five *as-GAL4* males of each *as-GAL4* line were crossed to five *Dcr2; UAS-singed*^{105747}*.RNAi/CyO* virgin females. Crosses were kept at 29 ºC. The non-curly males (*Dcr2; UAS-singed*^{105747}*.RNAi/+*; *+/as-GAL4)* were collected for dissection and kept at -20 °C. Hypandrium dissection and image acquisition were performed as indicated above. For each *as-GAL4* line at least 5 genitalia were examined (Table S6).

To test whether the *15E09*, *18C05* and *054839* enhancer-*GAL4* drive expression in the hypandrial bristle region in absence of *sc*, we crossed five *sc*^{29}*; UAS-scute (III)* females with five males of each respective *GAL4* line, as well as five *sc*^{M6}*/FM7 ; UAS-scute (III)* females with five males of each respective *GAL4* line. Of the three GAL4 lines, only *18C05* could induce hypandrial bristles with *UAS-sc* in a *sc* mutant background. The *18C05-GAL4* line produced approximately 10 bristles, where normally only two develop, which may reflect the amplification of gene expression that is inherent to the *UAS-GAL4* system. These results suggest that only *18C05* drives sufficiently strong expression in the hypandrial region to alter bristle patterning.

Cloning of *enhancers* into pBPGUw and pBPSUw

Enhancers were cloned into the *GAL4* reporter vector pBPGUw using the same strategy as in [37,38]. Enhancer sequences were amplified by Phusion® High Fidelity Polymerase (New England Biolabs) in two steps reaction using the primers and templates listed in Table S7, S8 and S9. PCR products and vectors were purified by Nucleospin Gel and PCR Clean-Up Kit (Machery-Nagel). Clones were purified by E.Z.N.A.® Plasmid Mini Kit I (Omega Bio-tek). All *GAL4* constructs were cloned using the Gateway® system (ThermoFisher Scientific). The enhancer fragments were first ligateded into *Kpn*I and *Hind*III restriction enzyme site of the vector pENTR/D-TOPO (Addgene) (Table S8). Recombination into the destination vector pBPGUw was performed using LR clonase II enzyme mix (Invitrogen) and products were transformed into One Shot® TOP10 (Invitrogen) competent cells. Recombinant clones were selected by ampicillin resistance on Amp-LB plates (100 µg/ml)

The pBPSUw vector was constructed by replacing the *GAL4* cassette of pBPGUw by *scute* CDS. The *scute CDS* was amplified from *D. melanogaster iso-1* with Scute-CDS-Rev and Scute-CDS-For primers and ligated into pGEM-T Easy (Promega). The *sc-CDS* insert was cut out using *Kpn*I and *Hind*III and cloned into *Kpn*I and *Hind*III sites in pBPGUw, thus replacing *GAL4*. The vector was named pBPSUw where “S” stands for *scute. 18C05* sequences from *D. melanogaster, D yakuba and D. santomea* were cloned into pBPSUw and tested in rescuing hypandrial bristles in *sc* mutants as written above. We found that *18C05* from *D. melanogaster* rescued two hypandrial bristles in both *sc*^{29} and *sc*^{M6} mutants. *D. santomea 18C05* enhancer rescued fewer hypandrial bristles on average than the *D. yakuba 18C05* region (Fig. 3., bristle number for *D. yakuba 18C05* in *sc*^{29} is significantly different from 0 (Exact-Poisson, $p < 10^{-16}$) and bristle number for *D. santomea 18C05* in *sc*^{M6} is significantly different from 2 (Exact-Poisson, p =0.0008)).

The *18C05* full length sequences were amplified by PCR from *D. melanogaster iso-1*(BL2057), *D. melanogaster T-7*, *D. yakuba Ivory Coast* and *D. santomea SYN2005* with the primers listed in Table S7. The PCR products were cloned into *pBPSUw* as described above. Three different *D. melanogaster 18C05* sequences were tested with *UAS-sc* in the hypandrium in sc^{29} and sc^{M6}. *GMR-18C05* (BL2057) was obtained from the Janelia Farm collection [38] and *18C05_BL2057* and *18C05_T7* were cloned in this study. Hypandrial bristle number was found to be significantly higher for *GMR-18C05* than for *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}). 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'. All the *18C05* constructs we made were inserted in the same orientation, 5'-3'. *GMR-18C05* and *18C05_BL2057* are the same sequences (from *D. melanogaster* Bloomington Stock Center Strain #2057), but cloned in opposite directions. *18C05_T7* contains the *18C05* sequence of *D. melanogaster T.7* strain. Comparing bristle number between *GMR-18C05-GAL4* and *18C05_BL2057-GAL4* shows that the orientation of the cis-regulatory region has an effect on bristle number.
The *18C05-chimera-pBPSUw* constructs were cloned using Gibson Assembly [39] by fusing together different lengths of *18C05* sequences from *D. yakuba Ivory Coast* and *D. santomea SYN2005*. The different chimeras are listed in Table S9. Cloning primers were designed using NEBuilder Tools (http://nebuilder.neb.com/). Primer sequences and templates used in PCR are listed in Tables S7 and S9. To assemble the *18C05* fragments in pBPSUw (Table S10), the vector was linearized by *Aat*II and *Fse*I restriction enzymes (New England Biolabs Inc.). After digestion thermosensitive alkaline phosphatase (FastAP, ThermoFisher Scientific) was added to the reaction to prevent self-ligation of the plasmid. PCR products and the linearized plasmid were isolated from 1% agarose gels and spin column purified. Gibson Assembly was performed as in [39], except that the assembly reactions were incubated at 37 °C for 10 minutes and then 3 hours at 50 °C in a PCR machine. 2 µl of assembly mixtures were transformed into NEB® 10-beta (New England Biolabs Inc.) competent cells and ampicillin-resistant colonies were selected on 100 µg/ml Amp-LB plates. The Gibson Assembly Master-mix was prepared according to [39], its components were purchased from Sigma-Aldrich.

The *18C05-yakubaSNP-pBPSUw* constructs were cloned by Gibson Assembly as described above, except for *18C05yakT1008C* and *18C05yakT1482C* sequences, which were synthesized and cloned by GenScript® (Table S11). The *18C05-ancestral* sequences were synthesized and cloned by GenScript® into pBPSUw *Aat*II and *Fse*I sites, except for the *18C05_AncG869A*, *18C05_AncT670G* and *18C05_Anc-7SNP* sequences, which were cloned by us by Gibson Assembly into pBPSUw *Aat*II and *Fse*I sites using the 18C05_Ancestral_Gibson_forward and 18C05_Ancestral_Gibson_reverse primers (Tables S7 and S10). All transgenic constructs were integrated into the *attP2* landing site in *D. melanogaster* w^{1118} by BestGene Inc.

Genomic DNA preparations for sequencing the *18C05* region

Genomic DNA was isolated with Zymo Research Quick-DNA™ Miniprep Plus Kit from 3 males and 3 females from the *D. yakuba*, *D. santomea* and *D. teisseri* lines listed in Table S12. *18C05* sequences were amplified with San-Yak_lines_sequencing-For and San-Yak_lines_sequencing-Rev primers (Table S7) using Phusion® High Fidelity Polymerase (New England Biolabs).

Sequence Analysis

Geneious software was used for cloning design and DNA sequence analysis. Nucleotide positions are given according to the alignment of *D. yakuba* Ivory Coast *18C05* sequence with *D. santomea* SYN2005 *18C05* sequence. The *18C05ancestral* sequence of *D. yakuba* and *D. santomea* was reconstructed in Geneious based on the *18C05* sequence alignment of *Drosophila* lines listed in Table S12. Manual parsimony reconstruction of all the ancestral nucleotides was unambiguous, except for one position (766, indel polymorphism), where the sequence is absent in the *simulans* complex and in *D. santomea*, while it is present in *D. teissieri* and polymorphic in *D. yakuba*. For this position we chose *D. teissieri* as the ancestral sequence. The *18C05* sequences of *D. melanogaster* subgroup species were retrieved by BLAST from the NCBI website. Transcription Factor (TF) binding sites in *18C05* were predicted using the JASPAR CORE Insecta database (http://jaspar.genereg.net [40]). 25-60 bp sequences of *18C05* were scanned with all JASPAR matrix models with 50-95% Relative Profile Score Thresholds to test for sensitivity and selectivity [40] (Table S12). For TFs which were absent in JASPAR (Scute), we used Fly Factor Survey to analyze their putative binding affinities to the probe. As *sc* cis-regulatory region is known to contain binding sites for Scute itself [41], we looked for Scute binding sites in *18C05*, *15E09* and *054839*. Two putative Scute binding sites (consensus motif *CAYCTGY*, Fly Factor Survey [41] were found in *15E09* and *054839* but not in *18C05*.

Statistical Analyses

Since bristle number is a classical type of count data, we performed statistical analysis using generalized linear models (GLM) and generalized linear mixed models (GLMM) where bristle number, the response variable, is assumed to follow a Poisson distribution [42–44]. All statistical analyses were performed using R 3.4 [45]. GLM were fitted with the function glm() ("stats" core package 3.5.0) and GLMM with the function glmer() ("lme4" package 1.1-14 [46] with the parameter "family" taken to be "Poisson". We tested differences in bristle number by comparing two wild-type stocks of *D. yakuba* with two wild-type stocks of *D. santomea*. We tested the difference between species, using a GLMM of the Poisson type (GLMM-Poisson) where the number of bristles was the response variable, species was a fixed effect to test and stock a random effect. For all other analyses, we tested differences in bristle number between genotypes using GLM of the Poisson type (GLM-Poisson) where the response variable was bristle number and genotype, a categorical variable, was the fixed effect. When we noticed important differences between residual deviance and residual degrees of freedom, we also fitted a quasi-likelihood model of the type "quasi-Poisson" (GLM-Quasi-Poisson) which allows for a model of the Poisson type, but where the variance can differ from the mean and is estimated based on a dispersion parameter (see for example [44] p. 225). For each model, in order to retain the model that fitted best to the data, analysis of deviance was performed using the anova.glm() with "test = Chisq" for GLM-Poisson and "test = F" for GLM-Quasi-Poisson. When needed, we performed multiple comparisons using the glht() function and the "Holm" adjustment parameter ("multcomp" package 1.4-7 [47]) which performs multiple comparisons between fitted GLM parameters and yields adjusted p-values corrected according to the Holm-Bonferroni method [48,49] also performed an exact Poisson test (R function "poisson.test") to test sample mean to a reference value assuming a Poisson distribution. Mean and 95% confidence intervals were directly extracted from the fitted GLM and transformed using exp(coef()) and exp(confint.default()).

For EMSA data, response curves were compared between yak probe and san probe using an ANCOVA after natural log transformation. The unlabeled san 450x responses were compared between yak probe and san probe using a one-sided Mann-Whitney U-test.

Abd-B homeodomain (Abd-B-HD) purification and EMSA

The Abd-B-HD-pGEX-4T-1 plasmid (kindly provided by Sangyun Jeong) was transformed into BL21 (DE3) chemically competent cells. Protein expression was induced by 0.1 mM IPTG (isopropyl-β-D-thiogalactopyranoside, Sigma Aldrich). Recombinant protein was purified from 500 ml of bacterial culture as described in Frangioni [50] except that proteins were eluted into 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10mM reduced glutation (Sigma-Aldrich, G-4251) and 5% glycerol. Concentrations and purity of the protein were determined by SDS-PAGE and Qubit 2.0 Fluorometer (Life Technologies). Protein aliquots of 20 µl were snap-frozen in liquid nitrogen and stored at -80 °C.

The HPLC-purified biotinylated and non-labelled oligonucleotides (Sigma-Aldrich) were used in PCR to obtain 54 bp probes *yak* and *san* (*san=yakT1775G*) from *18C05yak-pBPSUw* and *18C05yakT1775G-pBPSUw* plasmid templates. Oligonucleotides are listed in Table S7. PCR products were column-purified.

We then used electrophoretic mobility shift assay (EMSA) to test whether the purified Abd-B homeodomain (ABD-B-HD) can bind directly to a 54-bp fragment of *18C05* with the T1775G site at position 13 containing either T (*yak* probe) or G (*san* probe). In each binding reaction, 20 fmol of probes were mixed with the purified ABD-B HD ranging from 0-1.25 µg (0 µg, 0.75 µg, 1 µg and 1.25 µg) in binding buffer containing 10mM TRIS pH 7.5, 50 mM Kcl, 0.5 mM DTT, 6.25 mM $MgCl_2$, 0.05 mM EDTA, 50 ng/µl Salmon Sperm DNA (Sigma Aldrich) and 9.00% Ficoll 400 (Sigma Aldrich). The competition assay was performed by adding 9 pmol of unlabeled probes (450-fold excess) to the binding reaction. The reaction mixtures were incubated at 22 °C for 30 min and run on a non-denaturing 6% polyacrylamide gel in 0.5X TBE (Eurofins).

Labeling reactions were carried out with LightShift™ Chemiluminiscent EMSA Kit (ThermoFisher Scientific) according to the provider instructions with the following modifications: after electrophoresis, gels were blotted overnight in 20X SSC using the TurboBlotter Kit (GE Healthcare Life Sciences) and cross-linking of the probe to the membrane UV-light was performed at 254 nm and 120 mJ/cm2 (UV stratalinker® 2400, STRATAGENE). Chemiluminescence stained membranes were exposed to a CDD camera (FUJIFILM, LAS-4000) for 50x 10 sec exposition time increments. The last images were used for quantification and were never saturated according to LAS 4000 software.

To quantify the binding affinity of Abd-B-HD to the probes, the fractional occupancy (ratio of bound/(free+bound) probe) was calculated for three replicate experiments (Fig. S12. E) using the intensity values of the bands measured in Image [35]. The mean fractional occupancy was significantly lower with *D. santomea* probes than with *D. yakuba* probes (ANCOVA, $F(1,15)=10.58$, $p = 0.005$). We found that ABD-B-HD binds both *D. yakuba* and *D. santomea* DNA (Fig. S12. B-C). ABD-B-HD binding to the *D. yakub*a probe always resulted in a stronger shift than to the *D. santomea* probe. Furthermore, the *D. santomea* cold probe did not compete as efficiently as the *D. yakuba* cold probe to prevent formation of the *D. yakuba* DNA-ABD-B-HD complex (U-test, p=0.05).

Abd-B RNAi and clonal analysis

To test whether *Abd-B* is required for hypandrial bristle development, we reduced *Abd-B* expression using either genetic mutations or RNAi. Two *UAS.Abd-B-RNAi lines (#51167* and *#26746*) were crossed with 3 different *GAL4* lines, *GMR18C05-GAL4*, *NP5130-GAL4* and *NP6333-GAL4.* Crosses were kept at 29 °C and the hypandrium phenotype was examined in 10-50 F1 males (Table S13). Using the genitalia *GAL4* drivers *esg-GAL4*^{NP5130} [51] and *NP6333 [52]* to express *Abd-B.RNAi*^{51167}, we obtained 20 males out of 100 with developed hypandrium, among which two aberrant hypandrial bristle phenotypes were found, either bristle size reduction or bristle loss (Fig. S11. A-F, n=9/11 for *NP5130*, n=8/9 for *NP6333*, Table S13). Smaller bristles might arise from a delay in *sc* expression during development [53]. Since *Abd-B* null mutations are lethal [54], we produced mitotic mutant clones for two null mutations, *Abd-B*^{M1} [54] and *Abd-B*^{D18} [55]. *Abd-B* mutant mitotic recombinant clones were induced by the FLP/FRT system [56] using *Abd-B*^{M1} and *Abd-B*^{D18} null mutations. To induce clones, ten *yw hsflp122; FRT82B hs-CD2 y*$^{+}$ *M(3) w*^{123}/TM2 virgin females were crossed to ten *y; FRT82B Abd-B*^{M1} *red[1] e[11] ro[1] ca[1]/TM6B* or *y; FRT82B Abd-B*^{D18}*/TM3* males (stocks were kindly provided by Ernesto Sánchez-Herrero). Crosses were flipped every 24 hours and F1 progeny were heat-shocked at 38 °C for 1 hour at different stages of larval development: 24-48, 48-72, 72-98 and 96-120 hours after egg laying [57]. From both crosses, a total of 82 F1 males (Table S13) with the genotype of *yw hsflp122; FRT82B hs-CD2 y*$^{+}$ *M(3) w*^{123}*/FRT82B Abd-B*^{M1} *red[1] e[11] ro[1] ca[1]* and *yw hsflp122; FRT82B hs-CD2 y*$^{+}$ *M(3) w*^{123}*/FRT82B Abd-B*^{D18} were examined. Hypandria were mounted and bristle clones were screened as described above. Most of the resulting males showed extreme transformation of the genitalia (Fig. S11. G-H, O-P) but 12 males out of 82 had analyzable hypandrium (twelve males for *Abd-B*^{M1} and two males for *Abd-B*^{D18}). Among them, 6 males were devoid of one or both hypandrial bristles (Table S13, Fig. S11. I-N). When hypandrial bristles were present, most of them were heterozygous for the *Abd-B* mutation according to the visible markers associated with somatic recombination. Together, our results suggest that *Abd-B* is required for hypandrial bristle development.

Immunostaining

Larvae were fed on freshly prepared Formula 4-24® Instant Drosophila Medium, Blue (Carolina) and staged by the presence of blue staining in their gut [58]. Larvae were chosen with the most clear gut, indicating a developmental stage of 1-6 hours before pupa formation [59]. Head and genital parts of the larvae were cut and fixed in 4% PFA in PBS pH 7.4 for 20 minutes at room temperature, washed three times for 5 minutes in PBS containing 0.1% Tween20 and then permeabilized in TNT buffer (TRIS-NaCl buffer containing 0.5% Triton X-100) for 10 minutes. Samples were washed in 5% BSA in TNT for up to 5 hours at room temperature and then incubated with rabbit anti-GFP primary antibodies (Thermofisher #A6455) diluted in 1:1000 in TNT overnight at 4 °C and rinsed in TNT three times for 10 minutes at room temperature. Then, samples were washed in 5% BSA in TNT for up to 5 hours at room temperature and incubated with donkey Anti-Rabbit Dylight® 488 (Thermofisher) secondary antibodies diluted 1:200 in TNT overnight at 4 °C. After washing the preparations in TNT for 5 minutes DNA was stained in 1µg/µl DAPI solution (Sigma-Aldrich) for 30 minutes at room temperature. The preparations were finally washed in TNT three times for 5 minutes and the imaginal discs were dissected in PBS and mounted in Vectashield® H-1000. Images were acquired using Spinning Disc CSU-W1.

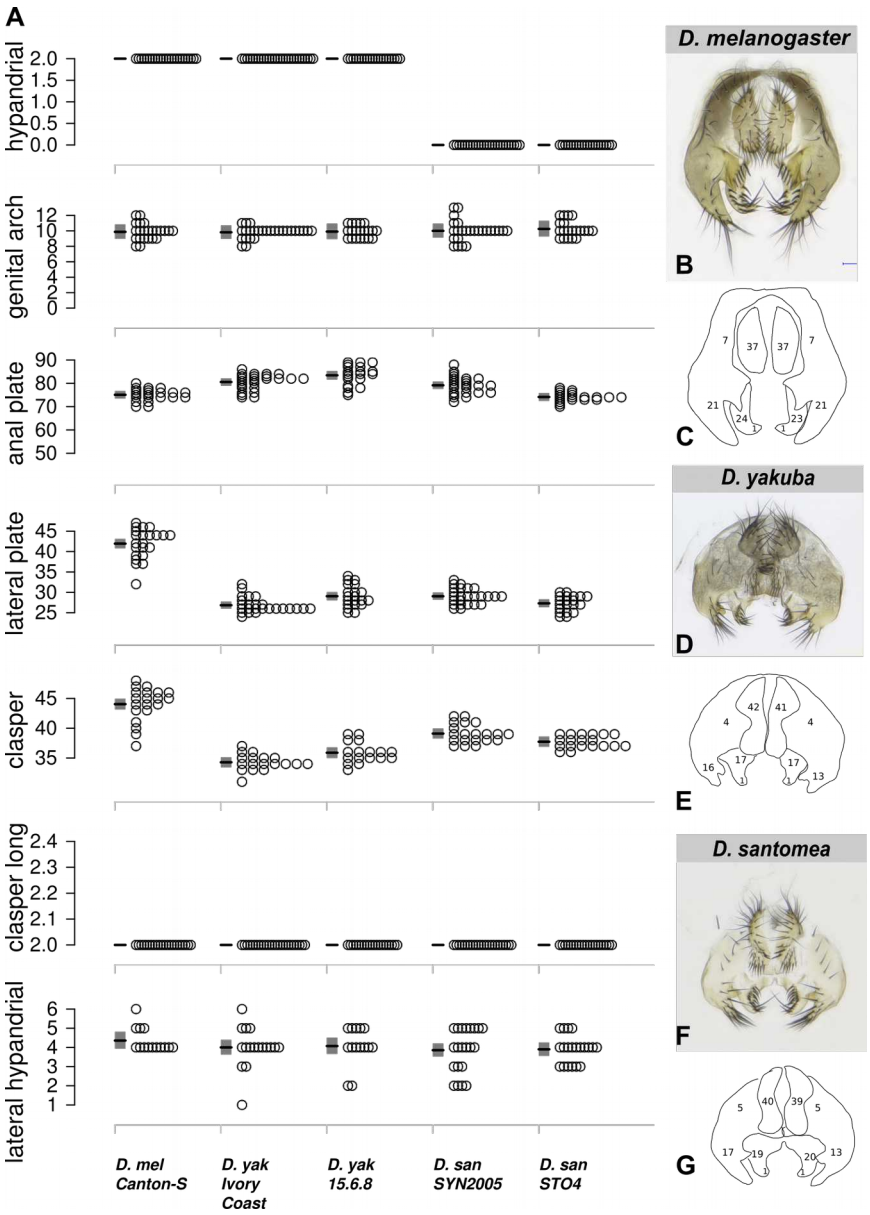

Fig. S1. Genital bristle number in *D. yakuba* and *D. santomea*.

(A) The bristle number of various bristle types is indicated for several strains of *D. melanogaster*, *D. yakuba* and *D. santomea*. Each grey dot represents one individual raised at 25°C. Mean (black line) and 95% confidence interval (grey rectangle) from a fitted GLM Quasi-Poisson model are shown. *D. yakuba* (yak) has fewer clasper bristles than *D. santomea* (san) (mean_yak= 35, mean_san = 38, GLMM-Poisson, Chisq(1) = 5. 01, p = 0.025) and more anal plate bristles than *D. santomea* (mean_yak = 82, mean_san = 77, GLMM-Poisson, Chisq(1) =3.87 p = 0.049). Dissected external genitalia from *D. melanogaster* Canton-S (B), *D. yakuba* Ivory Coast (D) and *D. santomea* SYN2005 (F) and their schematic representations (C,E,G) indicating the number of bristles in each genitalia part.

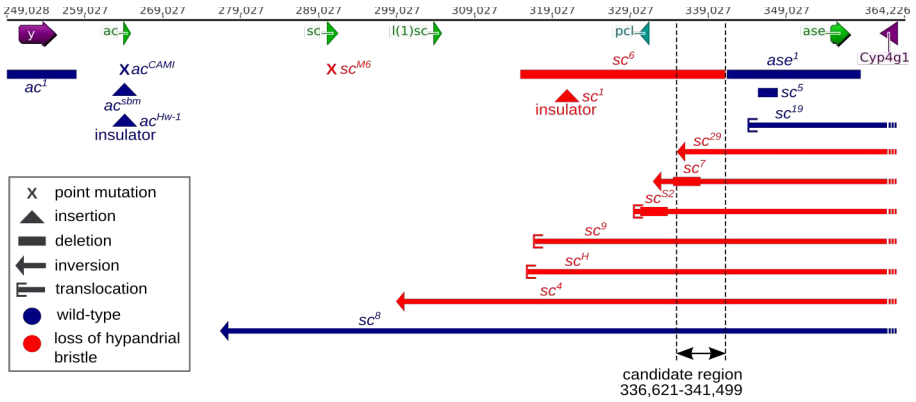

Fig. S2. Analysis of *achaete-scute* mutants in *D. melanogaster* reveals a 5-kb candidate region driving expression in hypandrial bristles.

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.

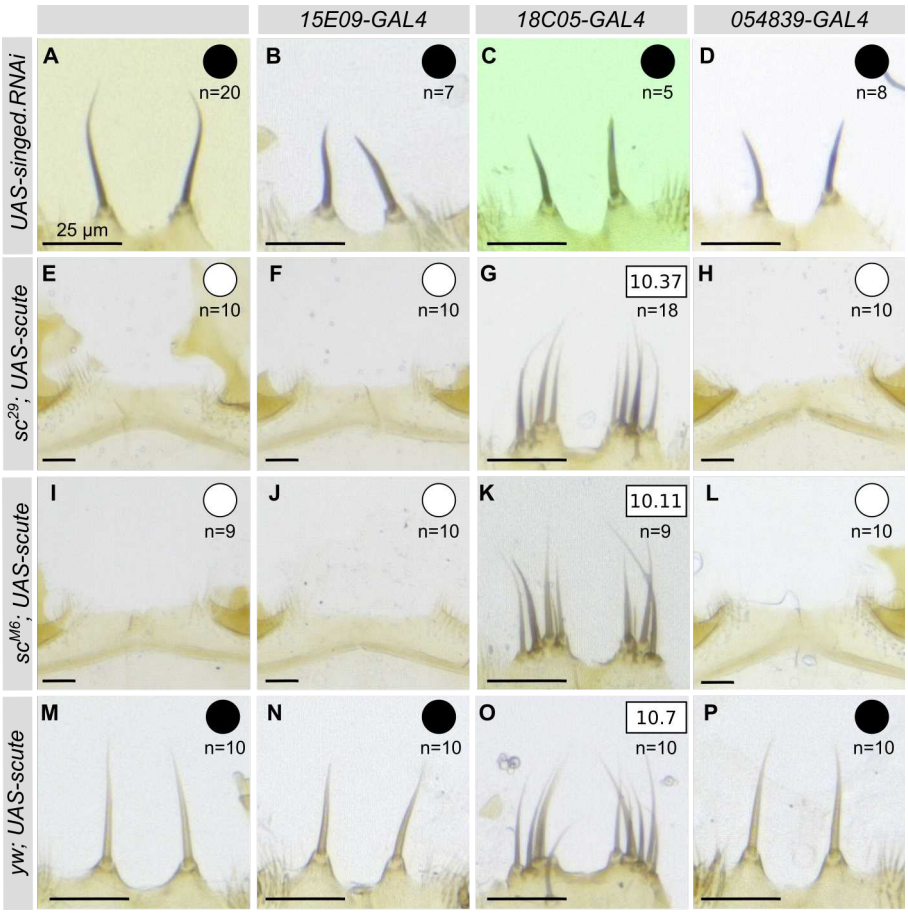

Fig. S3. Three cis-regulatory elements of the *achaete-scute* complex (*AS-C*) govern expression in the hypandrium in *D. melanogaster*.

Expression of *UAS-singed.RNAi* results in singed hypandrium bristles compared to wild-type (A) with *15E09-* (B), *18C05-* (C) and *054839-GAL4* (D). Expression of *UAS-sc* (E-P) in wild-type (M-P), sc^{29} (E-H) or sc^{M6} (I-L) mutant background produces extra bristles with 1*8C05-GAL4* (G,K,O) but not in absence of a *GAL4* reporter line (E,I,M) nor with *15E09-* (F,J,N) and *05439-GAL4* (H,L,P). White and black circles indicate zero or two hypandrial bristles, respectively. Average bristle numbers are shown in squares (G,K,O). n: number of scored individuals. (A-P) Scale bars indicate 25 µm.

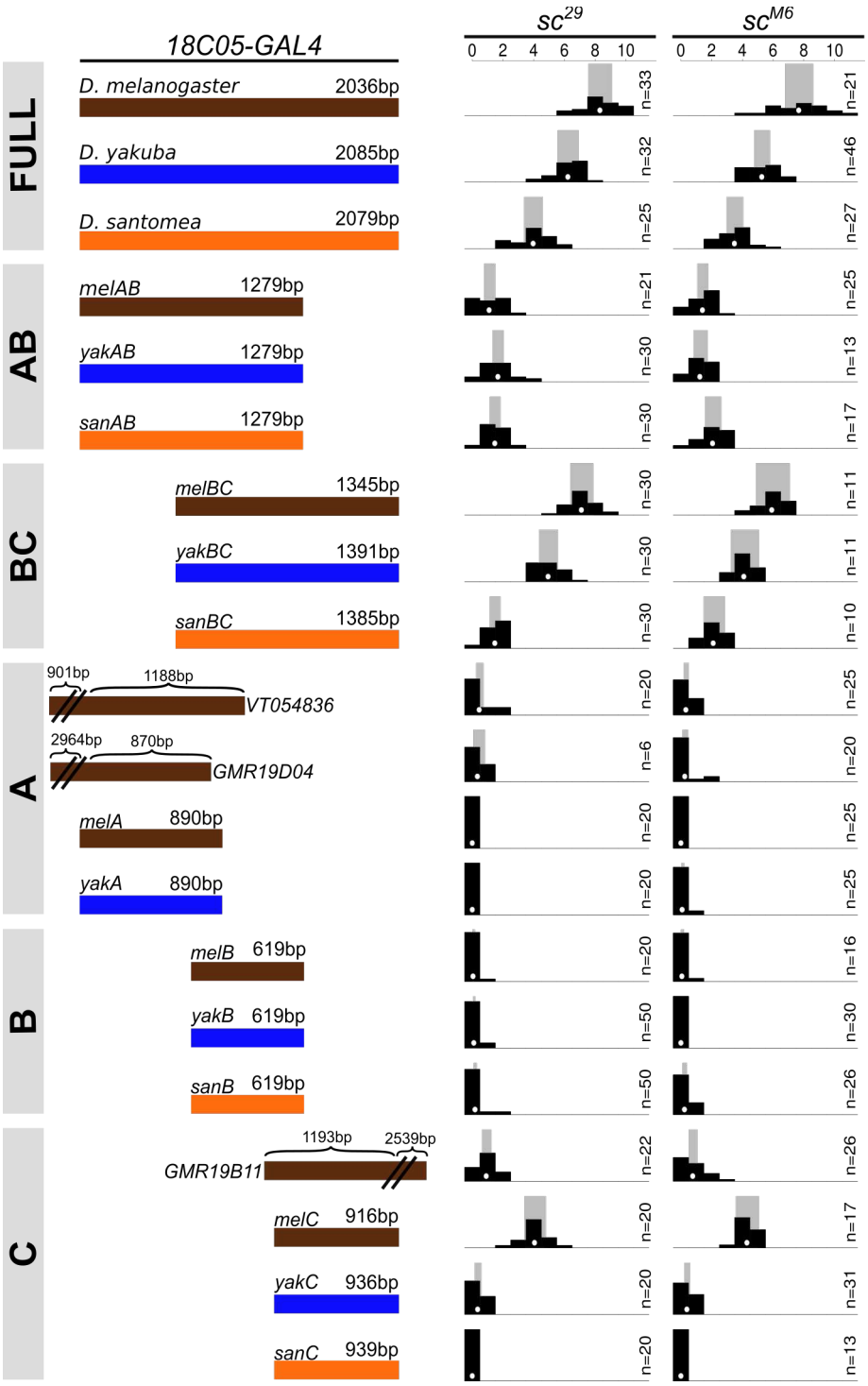

Fig. S4. Full-length *18C05* is required for complete expression in the hypandrial bristles. Rescue of the hypandrial bristle loss of sc^{29} (left column) and sc^{M6} (right column) *D. melanogaster* mutants by expression of *GAL4* with *UAS-sc* driven by various genomic regions from *D. melanogaster* (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, $p < 10\text{-}16$), whereas addition of DNA 3' of *18C05* reduces the number of bristles (*GM19B11* vs. *melC*; GLM-Quasi-Poisson, $F(19, 509) = 161.7$, $p < 10\text{-}16$ for sc^{29}; $F(19, 415) = 125.9$, $p < 10\text{-}16$ for sc^{M6}).

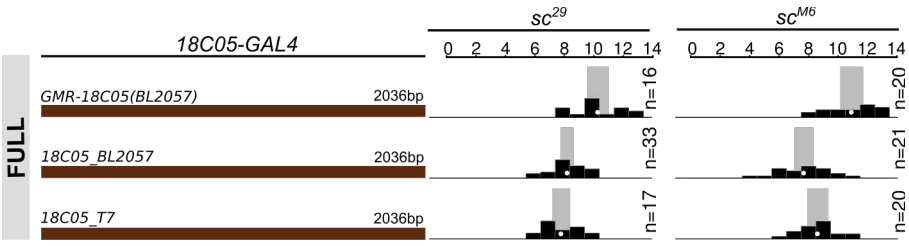

Fig. S5. Comparison of *18C05 melanogaster-GAL4* constructs.

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 Farm collection (*35*). *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'.

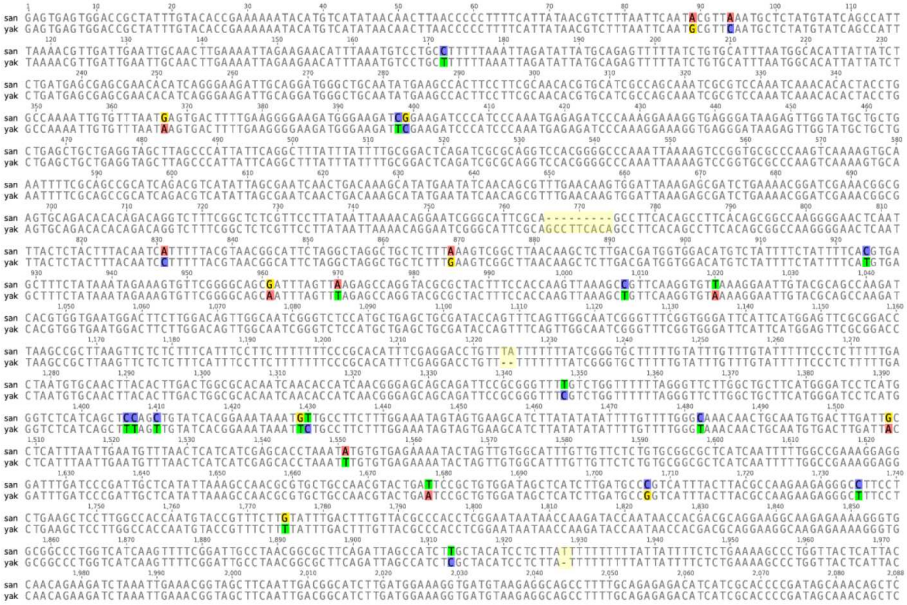

Fig. S6. Alignment of *18C05* sequence from *D. santomea* SYN2005 (san) and *D. yakuba* Ivory Coast (yak).

	GenBank Acc No	Nucleotide position	89	94	172	367	398	399	766	831	869	924	961	970	1008	1020	1224	1348	1406	1407	1410	1429	1430	1482	1507	1551	1678	1707	1735	1775	1913	1928
		Ancestral	**A**	**A**	**C**	**G**	**C**	**G**	**CAGCCTTCA**	**A**	**G**	**T**	**G**	**T**	**T**	**A**	**TA**	**T**	**C**	**C**	**C**	**T**	**T**	**T**	**A**	**A**	**T**	**C**	**C**	**T**	**C**	**del**
D. santomea	MG460738	A1200.4	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460742	car1490.3	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460740	BS14.1	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	del
	MG460737	51.7.3_1560	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	A
	MG460744	OBAT_1200.13	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460747	Quija650.37	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460746	Quija650.22	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460745	Quija650.14	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	A
	MG460741	C1350.14	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460739	B1300.13	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460748	Rain42	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460743	Field3.4	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	A
	MG460749	STO4	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
	MG460750	SYN2005	A	A	C	G	C	G	del	A	A	C	G	A	C	T	TA	T	C	C	C	G	T	C	G	A	T	C	C	G	T	T
D. yakuba	MG460759	15.6.8	A	A	C	A	T	C	CAGCCCTCA	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460761	LP1	G	C	T	A	C	G	del	A	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460756	2.22.1	G	C	T	A	T	C	del	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460757	4.23.1	A	A	C	A	T	C	del	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460758	4.32.1	A	A	C	A	T	C	CAGCCTTCA	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460765	Tai18E2	A	A	C	A	G	G	del	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460760	Ivory Coast	G	C	T	A	T	C	CAGCCTTCA	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460765	Tai18E2 (NCBI)	A	A	C	G	C	G	CAGCCTTCA	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	del	C	T	T	C	del
	MG460763	PB 3.1.3	A	A	C	A	T	C	CAGCCTTCA	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460764	PB 3.4.1	A	A	C	A	C	G	del	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460762	PB1.4.21	A	A	C	A	T	C	del	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
	MG460755	5.3.1	A	A	C	A	C	G	del	C	G	T	A	T	T	A	del	C	T	T	T	T	C	T	A	T	A	G	T	T	C	del
NCBI	MG460753	*D. teis*_(Mt. Selinda)	A	A	C	G	C	G	CAGCCCTCA	A	G	C	G	T	T	T	TA	T	C	C	C	T	T	T	A	A	T	C	C	T	C	del
	MG460754	*D. teis_(SDSC#14021-0257.01)*	A	A	C	G	C	G	CAGCCCTCA	A	G	C	G	T	T	T	TA	T	C	C	C	T	T	T	A	A	T	C	C	T	C	del
	MG460736	*D. mel*_BL2057	A	A	T	G	del	del	del	A	G	T	G	T	nd	A	TA	T	C	C	C	T	T	T	A	A	T	A	C	T	C	del
	MG460752	*D. sim*_w501	A	A	T	G	del	del	del	A	del	T	G	T	T	A	TA	T	C	C	C	T	T	T	A	A	T	A	C	T	T	del
	MG460751	*D. sim*_M252	A	A	T	G	del	del	del	A		T	G	T	T	A	TA	T	C	C	C	T	T	T	A	A	T	A	C	T	T	del
		*D. sim*_New Caledonia	A	A	nd	G	del	del									TA	T	C	C	C	T	T	T	A	A	T	A	C	T	T	
		*D. sim*_MD106TS	A	A	T	G	del	del																								
		*D. sim*_SIM6							del	A	del	T	G	T	nd	nd	TA	T	C	C	C	T	T	T	A	A	T	A	C	T	T	del
		*D. sim*_SIM4								A	del	T	G	T	nd	nd	TA	T	C	C	C	T	T	T	A	A	T	A	C	T		
		*D. elegans*_contig_6741				G	del	del																				A	C	T	C	del
		D. erecta	A	del?	C/del	G	del	del	del	A	G	A	G	T	T	A	TA	T	C	C	C	T	G	T	A	A	T	C	C	T	C	del
		*D. sec*_Rob3c	A	A	T	G	del	del	del	A	del	T	G	T	nd	nd	TA	T	C	C	C	T	T	T	A	A	T	A	C	T	T	del

Fig. S7. Twelve substitutions are fixed in *D. santomea 18C05*.

Summary of the alignment of *18C05* sequences from different lines of 7 species. Only divergent sites within and between *D. yakuba* and *D. santomea* are shown (divergence to *D. teissieri* is not shown). Nucleotide positions are indicated in the first row. The *D. santomea*-specific substitutions are labelled in orange when they are fixed in the other *D. melanogaster* subgroup species and in yellow when they are variable in at least one of the other *D. melanogaster* subgroup species. The reconstructed ancestral sequence of *D. santomea* and *D. yakuba* is shown at the top. Del: deletion. NCBI: sequences were retrieved from NCBI BLAST. GenBank accession numbers are shown in the first column for the sequences obtained in this study.

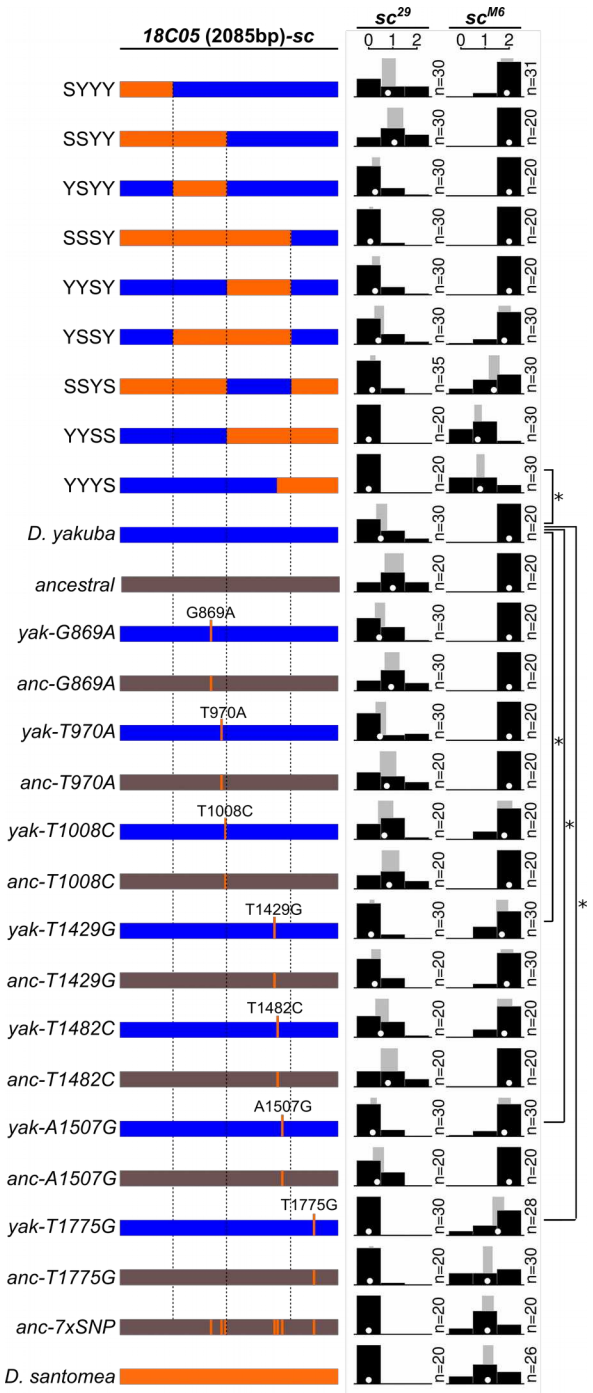

Fig. S8. Three *D. santomea*-specific substitutions in *18C05* cause the loss of hypandrial bristles

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}, YSYY 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).

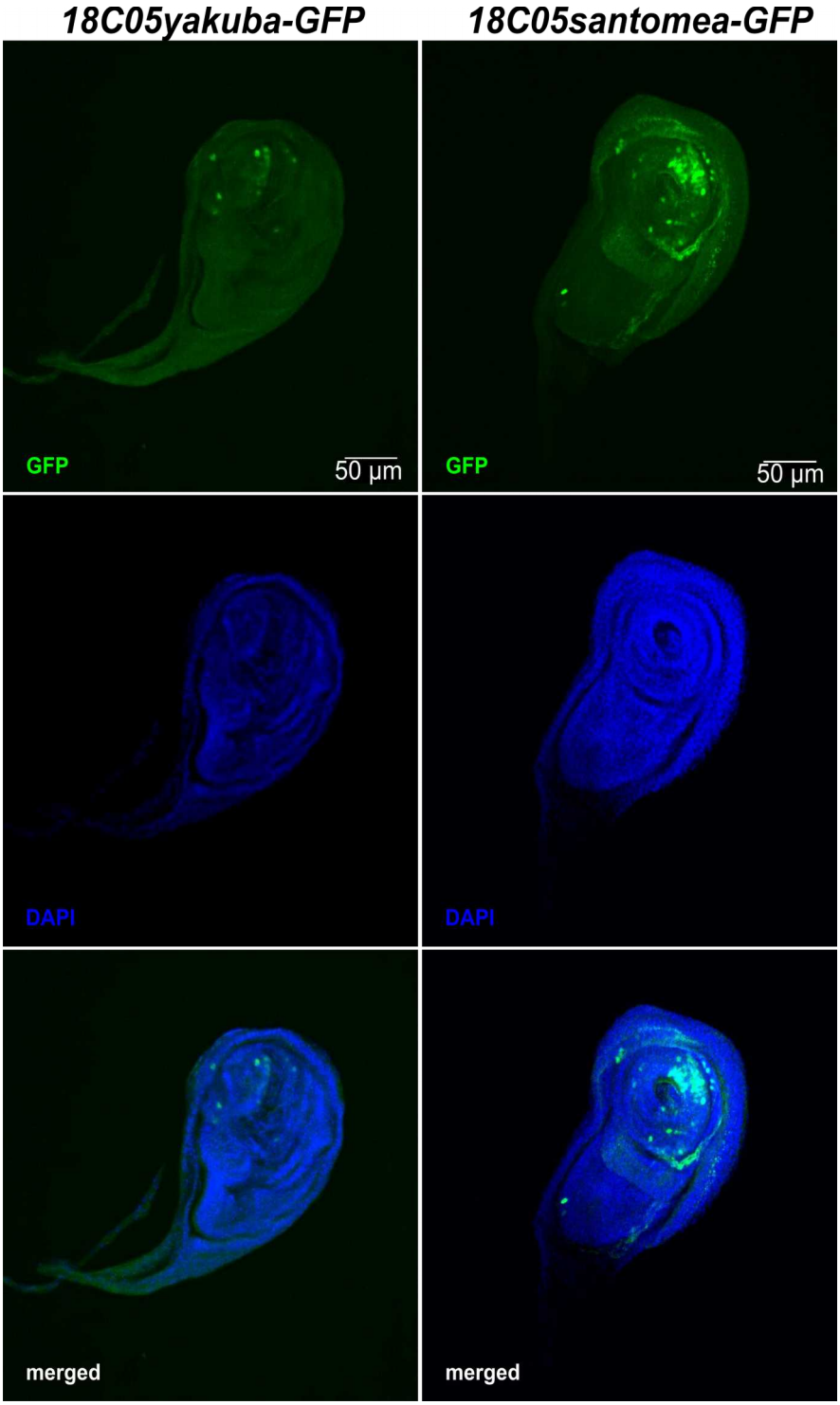

Fig. S9. *18C05* drives expression in T1 leg discs.

Immunostaining of T1 leg discs from *D. melanogaster* late L3 larvae containing either *18C05yakuba-GFP* (left column) or *18C05santomea-GFP* (right column) transgenes. GFP is labeled in green, DNA is shown in blue.

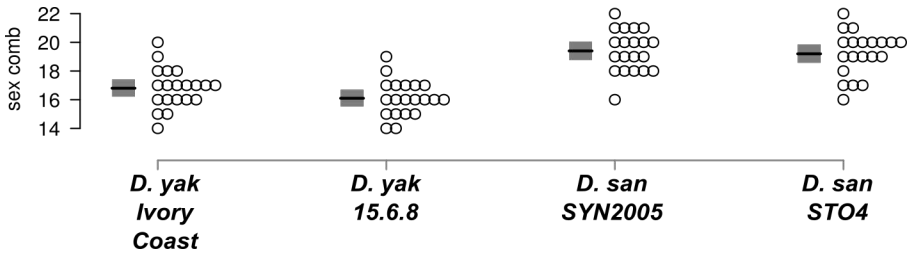

Fig. S10. Sex comb tooth number in *D. yakuba* and *D. santomea*.

Each circle represents one individual raised at 25° C. Mean (black line) and 95% confidence interval (grey rectangle) from a fitted GLM Quasi-Poisson model are shown. Sex comb tooth number is significantly different between *D. yakuba* and *D. santomea* (GLMM-Poisson, Chisq(1)=2.76, p = 0.007).

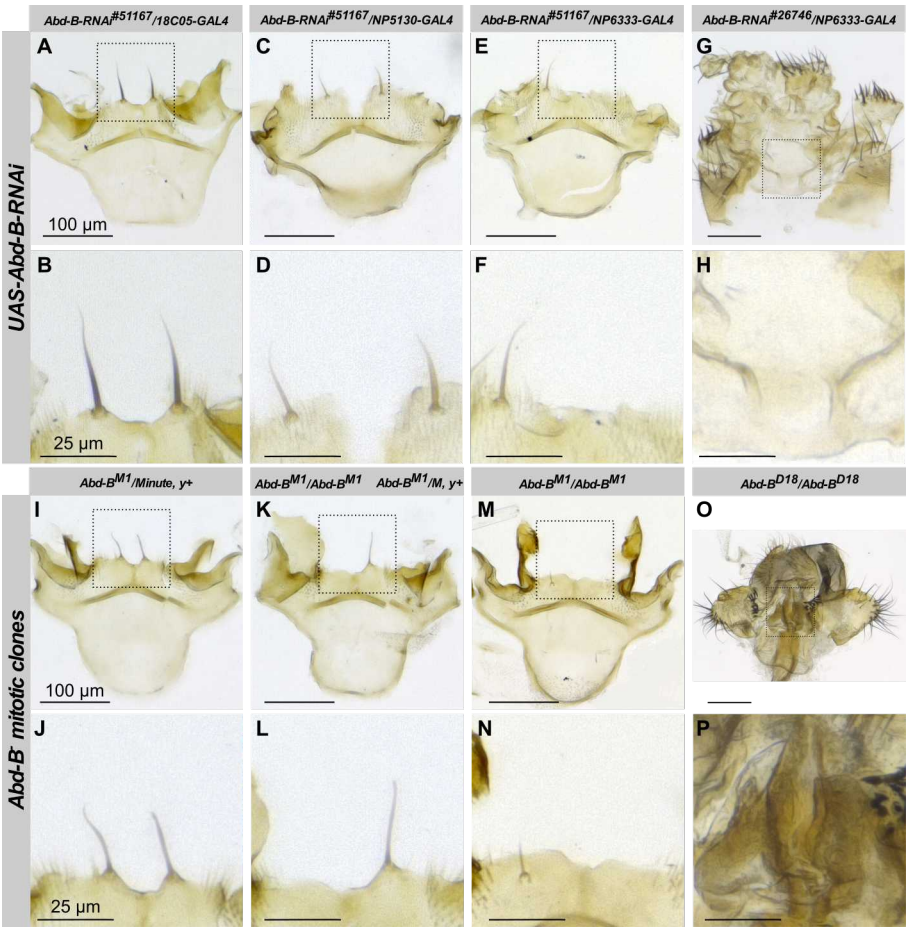

Fig. S11. Hypandrial bristle development affected in *Abd-B-RNAi* and *Abd-B⁻ clones*.

Wild-type hypandrial bristles. (B) In *Abd-B-RNAi*$^{\#51167}$*/GMR18C05-GAL4* males the hypandrial bristles are thiner. (C) In *Abd-B-RNAi*$^{\#51167}$*/NP5130-GAL4* hypandrial bristles are thiner and shorter. (D) In *Abd-B-RNAi*$^{\#51167}$*/NP6333-GAL4* hypandrial bristles are thiner or lost. (I-N) In *Abd-B*$^{-}$ clones hypandrial 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).

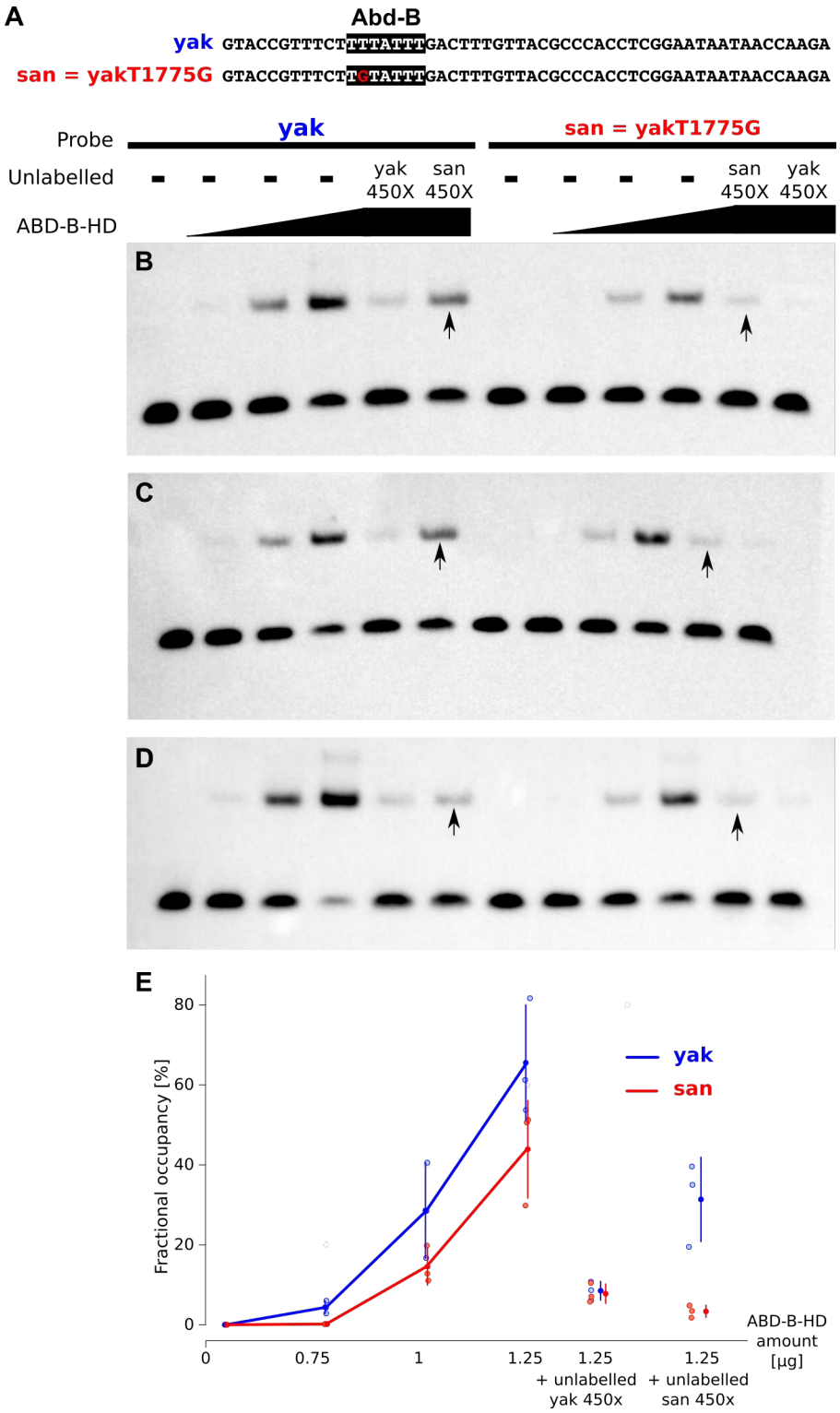

Fig. S12. *D. santomea*-specific substitution T1775G affects Abd-B binding.

(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) Threee paralel 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.

Table S1. Fly lines used.

See Table S4 for *GAL4* lines others than *DC-GAL4* and Table S5 for *achaete-scute* mutant lines.

Species Name	**Strain Name**	**Origin**
D. melanogaster	3870	Given by T. Long. RVC 3. Collected from Riverside, California, USA in 1963
D. melanogaster	3844	Given by T. Long. San Diego Stock Center #14021-0231.60 BS1. Collected from Barcelona, Spain in 1954
D. melanogaster	3841	Given by T. Long. San Diego Stock Center #14021-0231.59 BOG1. Collected from Bogota, Colombia in 1962
D. melanogaster	3852	Given by T. Long. San Diego Stock Center #14021-0231.64 KSA2. Collected in 1963
D. melanogaster	3864	Given by T. Long. San Diego Stock Center # 14021-0231.68 KI2. Collected from Israel in 1954
D. melanogaster	T.7	Given by T. Long. San Diego Stock Center #14021-0231.07 Collected from Taiwan in 1968
D. melanogaster	T.4	Given by T. Long. San Diego Stock Center #14021-0231.04 Collected from Kuala Lumpur, Malaysia in 1962
D. melanogaster	3875	Given by T. Long. San Diego Stock Center #14021-0231.69 VAG1. Collected from Athens, Greece in 1965
D. melanogaster	3886	Given by T. Long. Wild 5B. Collected from Red Top Mountain, Georgia in 1966
D. melanogaster	T.1	Given by T. Long. San Diego Stock Center #14021-0231.04 Collected from Ica, Peru in 1956
D. melanogaster	3839	Given by T. Long. San Diego Stock Center # 14021-0231.58 BER1. Collected from Bermudas in 1954.
D. melanogaster	3846	Given by T. Long. San Diego Stock Center #14021-0231.62 CA1. Collected from Cape Town, South Africa.
D. melanogaster	Sam	Given by T. Long. DSPR line. originally from TFC Mackay Sam; *ry506*
D. melanogaster	iso-1	Bloomington Stock Center #2057 (ordered) y[1]; Gr22b[iso-1] Gr22d[iso-1] cn[1] CG33964[iso-1] bw[1] sp[1]; LysC[iso-1] MstProx[iso-1] GstD5[iso-1] Rh6[1]
D. melanogaster	Canton-S	Kyoto DGGR #105666 (Given by Roger Karess)
D. melanogaster	*dor[4]/C(1)RM, y[1] w[1] f[1]*	Bloomington Stock Center #35 (ordered)
D. melanogaster	*Nup98-*	Bloomington Stock Center #4951 (ordered)

	96[339]/TM3, Sb[1]	
D. melanogaster	*Df(3R)D605/TM3, Sb[1] Ser[1]*	Bloomington Stock Center #823 (ordered)
D. melanogaster	DC002	Bloomington Stock Center #30213 (ordered) w1118; Dp(1;3)DC002, PBac{DC002}VK00033
D. melanogaster	DC003	Bloomington Stock Center #30214 (ordered) w1118; Dp(1;3)DC003, PBac{DC003}VK00033
D. melanogaster	DC004	Bloomington Stock Center #30215 (ordered) w1118; Dp(1;3)DC004, PBac{DC004}VK00033/TM6C, Sb1
D. melanogaster	DC006	Bloomington Stock Center #30217 (ordered) w1118; Dp(1;3)DC006, PBac{DC006}VK00033/TM6C, Sb1
D. melanogaster	DC097	Bloomington Stock Center #31440 (ordered) w1118; Dp(1;3)DC097, PBac{DC097}VK00033/TM6C, Sb1
D. melanogaster	DC098	Bloomington Stock Center #31441 (ordered) w1118; Dp(1;3)DC098, PBac{DC098}VK00033
D. melanogaster	DC007	Bloomington Stock Center #30218 (ordered) w1118; Dp(1;3)DC007, PBac{DC007}VK00033/TM6C, Sb1
D. melanogaster	DC008	Bloomington Stock Center #30745 (ordered) w1118; Dp(1;3)DC008, PBac{DC008}VK00033
D. melanogaster	DC009	Bloomington Stock Center #30219 w1118; Dp(1;3)DC009, PBac{DC009}VK00033
D. melanogaster	DC012	Bloomington Stock Center #30222 (ordered) w1118; Dp(1;3)DC012, PBac{DC012}VK00033
D. melanogaster	DC099	Bloomington Stock Center #30749 (ordered) w1118; Dp(1;3)DC099, PBac{DC099}VK00033
D. melanogaster	DC013	Bloomington Stock Center #30746 (ordered) w1118; Dp(1;3)DC013, PBac{DC013}VK00033
D. melanogaster	DC014	Bloomington Stock Center #31434 (ordered) w1118; Dp(1;3)DC014, PBac{DC014}VK00033
D. melanogaster	DC400	Bloomington Stock Center #30795 (ordered) w1118; Dp(1;3)DC400, PBac{DC400}VK00033
D. melanogaster	DC019	Bloomington Stock Center #30223 (ordered) w1118; Dp(1;3)DC019, PBac{DC019}VK00033
D. melanogaster	DC436	Bloomington Stock Center #33487 (ordered) w1118; Dp(1;3)DC436, PBac{DC436}VK00033/TM6C, Sb1
D. melanogaster	DC401	Bloomington Stock Center #30796 (ordered) w1118; Dp(1;3)DC401, PBac{DC401}VK00033
D. melanogaster	*DC-GAL4*	*yw ; DC-GAL4 , UAS-GFP /TM6B*

		Given by V. Stamataki (Pat Simpson lab).
D. melanogaster	*UAS-forked.RNAi*^{33200}	VDRC #33200 (ordered)
D. melanogaster	*UAS-singed.RNAi*^{105747}	VDRC #105747 (ordered)
D. melanogaster	*UAS-forked.RNAi*^{24632}	VDRC #24632 (ordered)
D. melanogaster	*UAS-ac.RNAi*^{100647}	VDRC #100647 (ordered)
D. melanogaster	*UAS-singed.RNAi*^{32579}	VDRC #32579 (ordered)
D. melanogaster	*UAS-forked.RNAi*^{103813}	VDRC #103813 (ordered)
D. melanogaster	*UAS-sc.RNAi*^{105951}	VDRC #105951 (ordered)
D. melanogaster	*yw; UAS-y*	Bloomington Stock Center #3043 y[1] w[1118]; P{w[+mC]=UAS-y.C}MC1
D. melanogaster	*yw;UAS-y TM3/pnr-GAL4*	Given by M. Rebeiz. y w; UAS-y TM3 Ser / pnr-GAL4
D. melanogaster	*UAS-mCD8-GFP*	Given by V. Brodu. *GFP* transgene on second chromosome
D. melanogaster	*w,UAS-Dcr2 ; Pin/CyO*	Bloomington Stock Center #24644 (ordered)
D. melanogaster	*UAS-scute*	Bloomington Stock Center #51672 (ordered)
D. melanogaster	*yw;UAS-Abd-B.RNAi*^{51167}	Bloomington Stock Center #51167 (ordered)
D. melanogaster	*yw; UAS-Abd-B.RNAi*^{26746}	Bloomington Stock Center #26746 (ordered)
D. melanogaster	*yw; NP5130-GAL4*	Kyoto Drosophila Stock Center (ordered)
D. melanogaster	*yw; NP6333-GAL4*	Kyoto Drosophila Stock Center (ordered)
D. simulans		Collected from Marrakech, Morocco by J. David in 2010
D. mauritiana		Collected from Mauritius Island in 1985
D. sechellia	GFP	San Diego Stock Center #14021-0248.32 w[1] ; pBac(3xP3-EGFPafm)::MCS::(pW8 mini-white)
D. yakuba	Ivory Coast	Given by D. Stern. San Diego Stock Center #14021-0261.00 Collected from Ivory Coast in 1955
D. yakuba	15.6.8	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 110m in 2009 by D. Matute
D. yakuba	*yellow[1]*	San Diego Species Stock Center #14021-0261.05

D. yakuba	4.23.1	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1070m in 2009 by D. Matute
D. yakuba	LP1	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 0m in 2009 by D. Matute
D. yakuba	2.22.1	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1250m in 2009 by D. Matute
D. yakuba	PB1.4.21	Given by D. Matute. Isofemale stock, collected in Bioko at altitude 1300m in 2009 by D. Matute
D. santomea	SYN2005	Given by D. Matute. Mix of six isofemale lines collected by J. Coyne at the field station Bom Sucesso (elevation 1,150 m) in 2005
D. santomea	STO.4	Given by D. Stern. San Diego Stock Center #14021-0271.00 Collected in São Tomé in 1998
D. santomea	Quija 650.22	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 650m in 2009 by D. Matute.
D. santomea	Quija 650.37	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 650m in 2005 by Lucio Primo Monteiro under the supervision of Daniel Lachaise.
D. santomea	Quija 650.14	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 650m in 2005 by Lucio Primo Monteiro under the supervision of Daniel Lachaise.
D. santomea	BS14.1	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1150m in 2009 by D. Matute
D. santomea	CAR1490.3	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1490m in 2009 by D. Matute
D. santomea	B1300.13	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1300m in 2009 by D. Matute
D. santomea	OBAT1200.3	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1200m in 2009 by D. Matute
D. santomea	A1200.4	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1200m in 2009 by D. Matute
D. santomea	C1350.14	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1350m in 2009 by D. Matute
D. santomea	Rain42	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1240m in 2009 by D. Matute
D. santomea	Field3.4	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1250m in 2009 by D. Matute
D. teissieri		Collected in Mt Selinda, Zimbabwe in 1970

D. teissieri	#14021-0257.01	San Diego Stock Center # 14021-0257.01
D. orena		Isofemale stock collected by J. David in 1975 in Cameroon
D. erecta		Collected in La Lopé, Gabon by D. Lachaise in 2005
D. elegans		Given by B. Prud'homme. San Diego Stock Center # 14027-0461.03. Collected in Hong-Kong.

Table S2. Hypandrial bristle number in pure species and F1 hybrids

[0]Flies raised at 25°C,[1]Flies raised at 29°C, [2]Flies raised at 18°C, [3]Flies raised at 21°C, [4]Flies raised at 14°C, [5]Flies collected directly from the wild.

	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**
D. melanogaster					
3870[0]	0	0	20	0	20
3844[0]	0	0	21	0	21
3841[0]	0	0	21	0	21
3852[0]	0	0	21	0	21
3864[0]	0	0	22	0	22
T.7[0]	0	0	20	0	20
T.4[0]	0	0	19	1	20
3875[0]	0	0	26	0	26
3886[0]	0	0	19	1	20
T.1[0]	0	0	19	1	20
3839[0]	0	0	20	0	20
3846[0]	0	0	20	0	20
Sam[0]	0	0	20	0	20
iso-1[0]	0	6	7	1	15
Canton-S[0]	0	0	20	0	20
D. sechellia *GFP*[0]	0	0	31	0	31
D. mauritiana[0]	0	0	34	0	34
D. mauritiana[2]	0	0	34	0	34
D. simulans[5]	0	0	10	0	10
D. yakuba					
Ivory Coast[0]	0	0	32	0	32
Ivory Coast[2]	0	0	30	0	30
4.23.1[1]	0	0	20	0	20
PB1.4.21[1]	0	0	25	0	25

LP1[1]	0	0	35	0	35
2.22.1[1]	0	0	13	0	13
15.6.8[1]	0	0	10	0	10
D. santomea					
SYN2005[0]	60	0	0	0	60
SYN2005[2]	30	0	0	0	30
STO.4[0]	29	0	0	0	29
Quija 650.37[0]	30	0	0	0	30
Quija 650.14[0]	35	0	0	0	35
BS14.1[2]	22	0	0	0	22
CAR1490.3[2]	21	0	0	0	21
B1300.13[2]	20	0	0	0	20
OBAT1200.3[2]	22	0	0	0	22
A1200.4[2]	20	0	0	0	20
C1350.14[2]	20	0	0	0	20
Rain42[2]	23	0	0	0	23
Field3.4[2]	21	0	0	0	21
D. teissieri[0]	0	0	30	0	30
D. teissieri[2]	0	0	30	0	30
D. orena[3]	0	0	5	0	5
D. erecta[4]	0	0	10	0	10
D. elegans[0]	0	0	17	0	17
D. santomea/D. yakuba F1 hybrids[0] (*D. santomea* females x *D. yakuba* males)	0	2	32	0	34
D. santomea/D. yakuba F1 hybrids[0] (*D. yakuba* females x *D. santomea* males)	29	0	0	0	29
***D. melanogaster/D. santomea* hybrids**					
san/DC002 (1A1-1A1)	19	4	1	0	24
san/DC003 (1A1-1A1)	15	7	2	0	24
san/DC004 (1A1-1A3)	18	6	0	0	24
san/DC006 (1A3-1A8)	18	3	3	0	24
san/DC097 (1A7-1B4)	**8**	**9**	**7**	0	24
san/DC098 (1B2-1B8)	19	3	2	0	24

san/DC007 (1B1-1B5)	18	5	1	0	24
san/DC008 (1B4-1B10)	21	2	1	0	24
san/DC009 (1B9-1B13)	18	4	2	0	24
san/DC012 (1C3-1C5)	17	5	2	0	24
san/DC099 (1C3-1C4)	21	3	0	0	24
san/DC013 (1C4-1D2)	16	8	0	0	24
san/DC014 (1D1-1D2)	20	4	0	0	24
san/DC400 (1D2-1E1)	19	3	2	0	24
san/DC019 (1E1-1E3)	18	5	1	0	24
san/DC436 (1E3-1E5)	22	0	2	0	24
san/DC401 (2A2-2B1)	21	3	0	0	24
san/TM3 (control)	68	18	2	0	95

Table. S3. *achaete-scute* mutant lines used.

Coordinates are for *D. melanogaster* reference genome iso-1 (FB2013_03, version 3). Based on restriction sites we estimated that point 0 of [60,61] 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.

Name	**Origin and full genotype**	**Description**
sc^{M6}	Bloomington Stock Center #52668 *sc[M6]/FM7i, P{w[+mC]=ActGFP}JMR3*	Nonsense Glu114X mutation in *scute*.
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 *achaete* 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 *achaete* ATG.
$ac^{Hw\text{-}1}$	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 *achaete* 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[sc29]	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	Inversion. Left breakpoint is between 318656

	In(1)sc[9], sc[9] w[a] f[1] Bx[1]	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 sc structural gene [62]. Minimal extent of the deletion is 332225- 336481.
ac^{3} $sc^{10\text{-}1}$	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 *scute* ($sc^{10\text{-}1}$).
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 *scute* 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 S4. Hypandrial bristle number in *achaete-scute* mutants

	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)					
sc^{M6}	15	0	0	0	15
ac^{CAMI}	0	0	15	0	15
Deletions (Cis-Regulatory)					
sc^{6}	16	0	0	0	16
ase^{1}	0	3	12	1	16
sc^{5}	0	0	17	0	17
ac^{1}	0	0	10	0	10
Insertions					
sc^{1}	17	0	0	0	17
ac^{sbm}	1	0	14	0	15
$ac^{Hw\text{-}1}$	0	0	15	0	15
Complex changes (inversions, translocations, etc.)					
sc^{29}	16	0	0	0	16
ac^{1} sc^{1}	15	0	0	0	15
sc^{H}	6	0	0	0	6
sc^{9}	15	0	0	0	15
sc^{S2}	5	0	0	0	5
sc^{7}	14	0	0	0	14
ac^{3} $sc^{10\text{-}1}$	14	0	0	0	14
sc^{4}	13	0	0	0	13
sc^{8}	0	0	23	0	23
ac^{1} sc^{19}	0	0	8	0	8

Table S5. Test of various *UAS*-reporter constructs with *DC-GAL4*

The number of adults (males and females) with *singed*/*forked*/absent dorsocentral thoracic bristles is shown for various *UAS-RNAi* reporter lines. The phenotype was recorded individually for each of the four dorsocentral bristles: AL: anterior left, AR: anterior right, PL: posterior left, PR: posterior right. *Dcr2*: *UAS-Dicer-2*, WT: wild-type bristle phenotype. Crosses were performed at 25°C and at 29°C for each genotype.

Bristles displaying a mutant phenotype		**AL AR PL PR**	**AL AR PL**	**AL AR PR**	**AL PL PR**	**AR PL PR**	**AL AR**	**PL PR**	**AL PL**	**AR PR**	**AL PR**	**AR PL**	**AL**	**AR**	**PL**	**PR**	**WT**	**Total Number of Flies**
Dcr2; singed^{105747}	25 °C	22	20	22			12		1	7			1	3			8	88
	29 °C	69	11	13			3							1			1	98
Dcr2; forked^{103813}	25 °C	15	18	22			14	2	2	2			3	6	5	6	44	139
	29 °C	123	33	36			15		1	1						1	5	215
Dcr2; forked^{33200}	25 °C		48										3	3			46	100
	29 °C	64	28	22			61		1				3	2			3	183
Dcr2; forked^{24632}	25 °C													1			154	155
	29 °C								1						1		175	177
Dcr2; singed^{32579}	25 °C																151	151
	29 °C																217	217
Dcr2; ac^{100647}	25 °C																136	136
	29 °C																258	258
Dcr2; sc^{105951}	25 °C																77	77
	29 °C																113	113

Table S6. *Achaete-scute GAL4* lines and their hypandrial bristle phenotype with *Dcr2; UAS-singed.RNAi*^{105747}.

Lines are ordered according to their left coordinates (Release 5). The origin of each *GAL4* line is indicated by a letter code: V: Vienna Drosophila Research Center [63], J: Janelia Farm [38], S: This Study. The orientation of the *scute* cis-regulatory region within the reporter construct is indicated by + or -. N: no data, I: determined by PCR in this study. In bold are the 3 *GAL4* lines which produce a hypandrial bristle mutant phenotype. For short, *GMR15E09*, *GMR18C05* and *VT054839* are named *15E09*, *18C05* and *054839* in the main text, respectively. Comparison of the *VT085439* cis-regulatory element from *D. santomea* and *D. yakuba* revealed no difference when tested in *GAL4* reporter lines with *Dcr2; UAS-singed.RNAi*^{105747} (last two columns).

Name	Origin	Orientation	Coordinates	Number of Flies with 0 singed bristles	Number of Flies with 1 singed bristles	Number of Flies with 2 singed bristles	Total Number of Flies
VT054793	V	-	X: 252,196-250,097	5	0	0	5
VT054794	V	+	X: 251,577-253,760	5	0	0	5
VT054795	V	+	X: 255,238-257,329	5	0	0	5
VT054796	V	+	X: 256,878-259,056	5	0	0	5
VT054798	V	-	X: 262,456-260,258	5	0	0	5
VT054799	V	+	X: 261,978-264,142	5	0	0	5
GMR14C10	J	N	X: 265,138-267,274	5	0	0	5
GMR15B10	J	N	X: 266,123-269,675	5	0	0	5
GMR15C11	J	N	X: 269,057-272,751	5	0	0	5
GMR15X09	S	+	X: 271,792-275,220	5	0	0	5
VT054805	V	+	X: 273,824-275,941	5	0	0	5
GMR15A01	J	N	X: 277,456-280,802	5	0	0	5
GMR14C12	J	N	X: 276,489-278,649	5	0	0	5
GMR15A04	J	N	X: 277,456-280,802	6	0	0	6
GMR15C10	J	N	X: 280,274-283,966	5	0	0	5
GMR15E07	J	N	X: 282,747-286,633	5	0	0	5
GMR15E09	**J**	**-**	**X: 289,793-285,861**	**0**	**0**	**6**	**6**
GMR13D04	J	N	X: 291,606-295,262	5	0	0	5
GMR13C08	J	N	X: 294,631-298,184	5	0	0	5
GMR12H02	J	N	X: 297,331-300,499	5	0	0	5

GMR13B12	J	N	X: 299,747-303,236	6	0	0	6
VT054820	V	-	X: 303,848-301,726	6	0	0	6
VT054821	V	-	X: 306,956- 304,793	6	0	0	6
VT054822b	S	N	X :308,704- 306,743	7	0	0	7
VT054823	V	+	X: 308,224- 310,380	9	0	0	9
VT054824	V	-	X: 312,108- 310,003	7	0	0	7
VT054825	V	+	X: 311,767- 313,924	6	0	0	6
VT054826	V	-	X: 315,628- 313,523	6	0	0	6
VT054827	V	+	X: 315,231- 317,337	7	0	0	7
VT054828	V	-	X: 319,149- 316,931	15	0	0	15
VT054829	V	-	X: 320,924- 318,733	12	0	0	12
VT054831	V	+	X: 328,588- 330,249	8	0	0	8
VT054832	V	-	X: 333,580- 331,369	6	0	0	6
GMR18G09	J	N	X: 332,537-335,588	5	0	0	5
VT054833	V	-	X: 335,283- 333,110	7	0	0	7
GMR18E07	J	N	X: 334,502-337,059	5	0	0	5
VT054834	V	+	X: 334,894- 337,095	6	0	0	6
GMR19D04	J	+/I	X: 336,372-340,205	5	0	0	5
VT054835	V	+	X: 336,673-338,747	7	0	0	7
VT054836	V	-	X: 340,523- 338,435	5	0	0	5
GMR18C05	**J**	**-/I**	**X: 341,372- 339,336**	**0**	**0**	**6**	**6**
GMR19B11	J	+/I	X: 340,180-343,911	6	0	0	6
VT054838	V	+	X: 341,960- 344,040	5	0	0	5
GMR18G07	J	+/I	X: 343,240-346,246	6	0	0	6
VT054839	**V**	**+**	**X: 343,684- 345,799**	**0**	**0**	**5**	**5**
GMR18F05	J	N	X: 345,467-348,263	5	0	0	5
VT054840	V	-	X: 347,574- 345,470	5	0	0	5
VT054841	V	-	X: 349,308- 347,221	7	0	0	7
GMR19A06	J	N	X: 347,473-350,875	5	0	0	5
VT054842	V	-	X: 351,081- 348,769	5	0	0	5
GMR18E10	J	N	X: 349,909-352,526	5	0	0	5
VT054843	V	-	X: 352,807- 350,712	7	0	0	7
GMR20B05	J	N	X: 351,638-353,720	5	0	0	5

VT054845	V	+	X: 357,132- 359,309	5	0	0	5
VT054846	V	-	X: 360,191- 358,098	8	0	0	8
VT054839mel -BL2057	S	+	Does not apply.	1	2	2	5
VT054839yak	S	+	Does not apply.	7	4	7	18
VT054839san	S	+	Does not apply.	7	8	5	20

Table S7. Primers used.

All primers were purchased from Sigma Aldrich.

Name	Sequence
15X09FF	CACCTT TCG CGA TGC AGG GAG AAA GTG A
15X09RR	ATG CTG CGC AGC GTG AGA AAA TGC A
VT054822bisF	CACCAAC GGA TTG GCG TGT TAA CCA GCG A
VT054822bisR	AAG CGG AAG AGG TAT CGG TGT CCT T
VT054839FF	CACCGGCAACTTTCAGCTCAGTTATGTCA
VT054839RR	CAT CGT CGC AGC GGG AAT ATG AAT T
18C05FDsanyak	CACCGAGTGAGTGGACCGCTATTTGTACA
18C05Rsanyak	GAGCTGTTTGCTATCGGGGTGCGA
18C05FDmel	CACCGAATGAGTGGACCGCTATTTGTACA
18C05Rmel	CAAGCCGTTTTCCTTTGAGGTGTCT
18C05bDF	CACCCGAGTGCAGACACACAGACAGGTCT
18C05dDF	CACCGTTCGCGGACCTAAGCCGCTTAAGT
18C05aR	TTG AGT TCC CCT TGG CCG CTG TGA A
18C05cR	TCT GCT GCT CCC GTT GAT GGT GTT G
18C05bF2	CACCTG AAA ACG GAT CGA AAC GGC GAG T
18C05cR2	GGT GTT GAT TGT GCG CCA GTC AAG T
18C05aR2mel	AAGCAGCTAAGCCTAGAACCCTGTTAC
Scute-CDS-FOR	GGTACCTGTTGATCGTTMTCCGGAA
Scute-CDS-FOR	AAGCTTAGTCACTGCTCCTGCCATAG
VT054838 forward	CTCCCCAGACTAACCCACTC
VT054838 reverse	AGCACAAAAAGGGCAGGAAC
GAL4 forward	CGAACAAGCATGCGATATTT
GAL4 reverse	GCTGCCGAGTCAATCGATAC
VT054839 forward	GGTGCATCCATCCATTCCT
VT054839 reverse	GCGGGAATATGAATTTTATGTAGA
Gal4 lines sequencing forward	ACAAGTTTGTACAAAAAAGCAGGCT
VT054836 forward	GACCTATGCACCAGCCATTT
VT054836 reverse	CGCACCGGACTTTTAATTTG

GMR15E09 forward	GGTGGACTCCGTTTCAGGT
GMR15E09 reverse	GCCGAACGGCAATAAAGTAA
PBPGUw-GIBSON-For	ATAGGGGTTCCGCGCACAT
PBPGUw-GIBSON-Rev	CTTTTATACCGCTGCGCTCGAT
San18C05_1-1070-Rev	ATTGCCAACTGTCCAAGAAGTC
Yak18C05_1021-2085-For	AAAGGAATTGTACGCAGCCAAG
San18C05_1-1609-Rev	TTCGGCCAAAAATTGATGAG
Yak18C05_1571-2085-For	TGGCATTTGTTGTTCTCTGTG
Yak18C05_1-550-Rev	TAATTTGGGCCCCGTGGA
San18C05_500-2079-For	TTTATTTATTTTGCGGACTCAGA
San18C05_1-550-Rev	ACCGGACTTTTAATTTGGGCCCCGTGGA
Yak18C05_551-2085-For	GCCCAAATTAAAAGTCCGGTGCGCCCAA
Yak18C05_1-1550-Rev_1	AGGTAATCAAGTCACATTGCAGTTGTTTGC
San18C05_1489-2079-For	GCAATGTGACTTGATTACCTCATTTAATTGAATG
Yak_1078-2086_San-SNP970-Rev	acctggctctTAACTAAATTGCTGCCCCGAAC
Yak_1-1078_San-SNP970-For	caatttagttAAGAGCCAGGTACGCGCTA
Yak_1535-2086_San-SNP1429-Rev	ccaaagaaggcaaCATTTATTTCCGTGATACAACTAAAGCTG
Yak_1-1535_San-SNP1429-For	ggaaataaatGCTGCCTTCTTTGGAAATAGTAG
Yak_1613-2086-San-SNP1507-Rev	aattaaatgaggCAATCAAGTCACATTGCAGTTGTTTACC
Yak_1-1613-San-SNP1507-For	gtgacttgattGCCTCATTTAATTGAATGTTTAACTC
Yak_1880-2086-San-SNP1775-Rev	caaagtcaaataCAAGAAACGGTACATTGGTGGC
Yak_1-1880-San-SNP1775-For	taccgtttcttGTATTTGACTTTGTTACGCCC
San-Yak_lines_sequencing-For	AATGGGTCACCCCGTGTA
San-Yak_lines_sequencing-Rev	GGTCTCTCAGGGTTTTCAAGC
EMSA_Yak_1775-For	GTACCGTTTCTTTTATTTGACTTTG
EMSA_San_T1775G-For	GTACCGTTTCTTGTATTTGACTTTG
EMSA_1775-Rev	TCTTGGTTATTATTCCGAGGTG
18C05_Ancestral_Gibson-for	TTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATTCGTTTATCACAAGTTTG
18C05_Ancestral_Gibson-rev	TCGATCCCCGGGCGAGCTCGGCCGGCCGTTTATCACCACTTTG

Table S8. PCR fragments cloned into pENTR/D-TOPO

Fragment	Template	Forward primer	Reverse primer
15X09	*D. melanogaster* iso-1	15X09FF	15X09RR
VT054822b	*D. melanogaster* iso-1	VT054822bisF	VT054822bisR
VT054839yak	*D. yakuba* Ivory Coast	VT054839FF	VT054839RR
VT054839san	*D. santomea* SYN2005	VT054839FF	VT054839RR
18C05_T7	*D. melanogaster* T7	18C05FDmel	18C05Rmel
18C05_BL2057	*D. melanogaster* iso-1	18C05FDmel	18C05Rmel
18C05Yakfull	*D. yakuba* Ivory Coast	18C05FDsanyak	18C05Ryaksan
18C05Sanfull	*D. santomea* SYN2005	18C05FDsanyak	18C05Ryaksan
18C05AmelBL	*D. melanogaster* iso-1	18C05FDmel	18C05aR2mel
18C05BmelBL	*D. melanogaster* iso-1	18C05bF2	18C05cR2
18C05CmelBL	*D. melanogaster* iso-1	18C05dDF	18C05Rmel
18C05ABmelBL	*D. melanogaster* iso-1	18C05FDmel	18C05cR
18C05BCmelBL	*D. melanogaster* iso-1	18C05bDF	18C05Rmel
18C05Asan	*D. santomea* SYN2005	18C05FDsanyak	18C05aR
18C05Bsan	*D. santomea* SYN2005	18C05bDF	18C05cR2
18C05Ayak	*D. yakuba* Ivory Coast	18C05FDsanyak	18C05aR
18C05Byak	*D. yakuba* Ivory Coast	18C05bF2	18C05cR2
18C05Cyak	*D. yakuba* Ivory Coast	18C05dDF	18C05Ryaksan
VT054839mel-BL2057	*D. melanogaster* iso-1	VT054839FF	VT054839RR
VT054839yak	*D. yakuba* Ivory Coast	VT054839FF	VT054839RR
VT054839san	*D. santomea* SYN2005	VT054839FF	VT054839RR

Table S9. PCR fragments used for Gibson assembly into pBPSUw

#	Fragment name	Template	Forward primer	Reverse primer
1.	San_1-1500	18C05_san_full-pBPSUw	PBPGUw-GIBSON-For	San18C05_1-1609bp-Rev
2.	Yak_1500-2085	18C05_yak_full-pBPSUw	Yak18C05_1571-2085-For	PBPGUw-GIBSON-Rev
3.	Yak_1-1000	18C05_yak_full-pBPSUw	PBPGUw-GIBSON-For	San18C05_1-1070-Rev
4.	San_1000-2085	18C05_san_full-pBPSUw	Yak18C05_1021-2085-For	PBPGUw-GIBSON-Rev
5.	San_1-1000+Yak1001-1500	San1-1000+Yak1001-2085-pBPGUw	PBPGUw-GIBSON-For	San18C05_1-1609-Rev
6.	San_1500-2079	18C05_san_full-pBPSUw	Yak18C05_1571-2085-For	PBPGUw-GIBSON-Rev
7.	San_1-500	18C05_san_full-pBPSUw	PBPGUw-GIBSON-For	San18C05_1-550-Rev
8.	Yak_500-2085	18C05_yak_full-pBPSUw	Yak18C05_551-2085-For	PBPGUw-GIBSON-Rev
9.	Yak_1-1500	18C05_yak_full-pBPSUw	PBPGUw-GIBSON-For	Yak18C05_1-1550-Rev_1
10.	San_1500-2079	18C05_san_full-pBPSUw	San18C05_1489-2079-For	PBPGUw-GIBSON-Rev
11.	San_1-1000	18C05_san_full-pBPSUw	PBPGUw-GIBSON-For	San18C05_1-1070-Rev
12.	Yak_1001-2085	18C05_yak_full-pBPSUw	Yak18C05_1021-2085-For	PBPGUw-GIBSON-Rev
13.	San_1001-1500+Yak1501-2085	San1-1500+Yak1501-2085-pBPSUw	Yak18C05_1021-2085-For	PBPGUw-GIBSON-Rev
14.	Yak_1-500	18C05_yak_full-pBPSUw	PBPGUw-GIBSON-For	Yak18C05_1-550-Rev
15.	San_501-1500	18C05_san_full-pBPSUw	San18C05_500-2079bp-For	San18C05_1-1609-Rev
16.	San_501-1000	18C05_san_full-pBPSUw	San18C05_500-2079-For	San18C05_1-1070-Rev
17.	SNP-970/1	18C05_yak_full-pBPSUw	PBPGUw-GIBSON-For	Yak_1078-2086_San-SNP970-Rev
18.	SNP-970/2	18C05_yak_full-pBPSUw	Yak_1-1078_San-SNP970-For	PBPGUw-GIBSON-Rev
19.	SNP-1429/1	18C05_yak_full-	PBPGUw-GIBSON-For	Yak_1535-2086_San-

		pBPSUw		SNP1429-Rev
20.	SNP-1429/2	18C05_yak_full-pBPSUw	Yak_1-1535_San-SNP1429-For	PBPGUw-GIBSON-Rev
21.	SNP-1507/1	18C05_yak_full-pBPSUw	PBPGUw-GIBSON-For	Yak_1613-2086-San-SNP1507-Rev
22.	SNP-1507/2	18C05_yak_full-pBPSUw	Yak_1-1613-San-SNP1507-For	PBPGUw-GIBSON-Rev
23.	SNP-1775/1	18C05_yak_full-pBPSUw	PBPGUw-GIBSON-For	Yak_1880-2086-San-SNP1775-Rev
24.	SNP-1775/2	18C05_yak_full-pBPSUw	Yak_1-1880-San-SNP1775-For	PBPGUw-GIBSON-Rev

Table S10. Scheme of Gibson assembly for chimeric constructs

See Table S9. for description of the PCR fragments.

Construct Name	Assembled PCR Fragments
18C05_SSSY	1+2
18C05_YYSS	3+4
18C05_SSYS	5+6
18C05_SYYY	7+8
18C05_YYYS	9+10
18C05_SSYY	11+12
18C05_YYSY	3+13
18C05_YSYY	14+16+12
18C05_YSSY	14+15+2
18C05yakT970A	17+18
18C05YyakT1429G	19+20
18C05yakA1507G	21+22
18C05yakt1775G	23+24

Table S11. List of *18C05yakuba-* and *18C05ancestral-SNP-pBPSUw* constructs synthesized by GenScript

Name	Construction
18C05_Anc-pBPSUw	Ordered at GenScript
18C05_AncT1008C-pBPSUw	Ordered at GenScript, nucleotide T at position 1008 replaced to C
18C05_AncT1429G-pBPSUw	Ordered at GenScript, nucleotide T at position 1429 replaced to G
18C05_AncT1482C-pBPSUw	Ordered at GenScript, nucleotide T at position 1482 replaced to C

18C05_AncA1507G-pBPSUw	Ordered at GenScript, nucleotide A at position 1507 replaced to C
18C05_AncT1775G-pBPSUw	Ordered at GenScript, nucleotide T at position 1775 replaced to C
18C05_yakG869A-pBPSUw	Ordered at GenScript, nucleotide G at position 869 replaced to A
18C05_yakT1008C-pBPSUw	pUC57 was ordered at GenScript and subcloned into pBPSUw using Gibson Assembly, nucleotide T at position 1008 replaced to C
18C05_yakT1482C-pBPSUw	pUC57 was ordered at GenScript and subcloned into pBPSUw using Gibson Assembly, nucleotide T at position 1482 replaced to C
18C05_AncG869A-pBPSUw	pUC57 was ordered at GenScript and subcloned into pBPSUw using Gibson Assembly, nucleotide G at position 869 replaced to A
18C05_AncT670G-pBPSUw	pUC57 was ordered at GenScript and subcloned into pBPSUw using Gibson Assembly, nucleotide T at position 670 replaced to G
18C05_Anc-SNPall-pBPSUw	pUC57 was ordered at GenScript and subcloned into pBPSUw using Gibson Assembly, nucleotide substitutions at the indicated positions: T670G, G869A, T1008C, T1429G, T1482C, A1507G, T1775G

Table S12. JASPAR binding scores for 6 loci of interest within the *18C05* region

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.

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 S13. Hypandrial bristle phenotypes in *ABD-B RNAi* lines and mitotic clones

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.

Genotype	**Number of dissected flies**	**Number of recognizable hypandrium**	**0 bristle**	**1 bristle**	**2 bristles**	**3 bristles**	**2 misplaced bristle**
Abd-B.RNAi*^{26746}*/NP6333-GAL4	50	0					
Abd-B.RNAi*^{26746}*/NP5130-GAL4	50	0					
Abd-B.RNAi*^{51167}*/NP6333-GAL4	50	9		1(t,s)	6(t,s)	1(t,s)	1(L)
Abd-B.RNAi*^{51167}*/NP5130-GAL4	50	11			9(t,s)		1(L), 1(R)
Abd-B.RNAi*^{51167}*/GMR18C05-GAL4	10	10			7(n), 3(t)		
yw hsflp122; FRT82B hs-CD2 y+ M(3) w*^{123} */FRT82B Abd-B*^{M1} *red[1] e[11] ro[1] ca[1]	30	10	2	2(M,y^{+})	2(M,y^{+}), 1(M,y^{-}), 1(M,y^{+}; M,y^{-})		2(R,M,y^{-})
***yw hsflp122; FRT82B hs-CD2 y+ M(3) w*^{123} */FRT82B Abd-B*^{D18}**	52	2	1	1(y^{+})			

Data S1. QTL-data (separate file)
Data for Fig. 2.

Data S2. Hypandrial bristle number with *18C05-GAL4* constructs (separate file)
Data for Fig. 3.

Data S3. Hypandrial bristle number with *18C05-sc* constructs (separate file)
Data for Fig. 3., Fig S3., Fig S4.

Data S4. Sex comb tooth number with various *18C05yak-sc* constructs (separate file)
Data for Fig. 4.

Data S5. Genital bristle numbers in *D. yakuba* and *D. santomea* strains (separate file)
Data for Fig. S1.Type or paste caption here.

Data S6. Hypandrial bristle number with *18C05 melanogaster-GAL4* constructs (separate file)
Data for Fig. S5.

Data S7. Hypandrial bristle number with chimeric constructs (separate file)
Data for Fig. S8.

Data S8. Sex comb tooth numbers in *D. yakuba* and *D. santomea* strains (separate file)
Data for Fig. S10.

Data S9. Fractional occupancy data (separate file)
Data for Fig. S12. EMSA shift intensity values measured by ImageJ and the calculated fractional occupancy.