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

Olga Nagy, Isabelle Nuez, Rosina Savisaar, Alexandre E. Peluffo, Amir Yassin, Michael Lang, David L Stern, Daniel R Matute, Jean R David, Virginie Courtier-Orgogozo

Correlated Evolution of two Copulatory 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[1,6,*]

[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.

[6]Lead Contact

*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 copulatory organs.

Highlights:
- We identify a gene and 3 substitutions causing genital evolution between species
- The evolved mutations lie in a pleiotropic enhancer
- One mutation decreases genital bristle number and increases leg sex comb tooth number
- This mutation disrupts a binding site for Abd-B in genitals and for another factor in legs

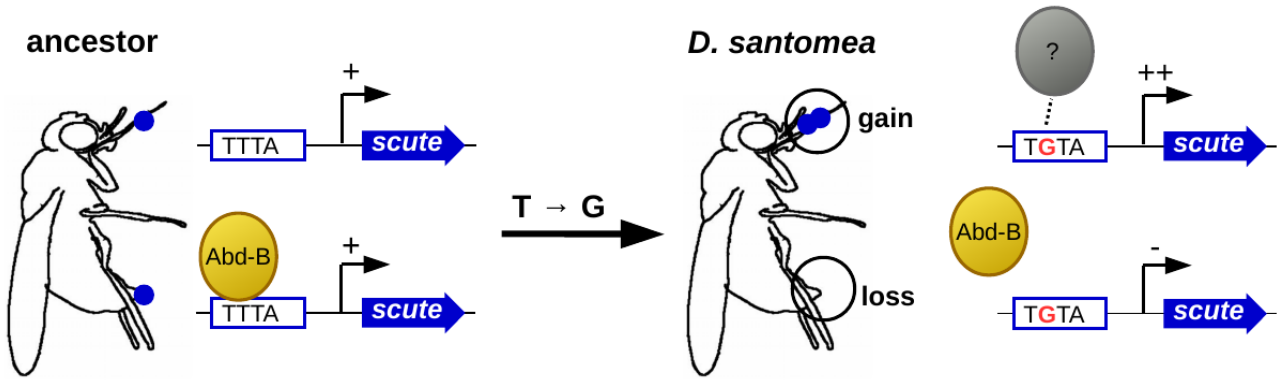
SUMMARY

Diverse traits often covary between species [1–3]. The possibility that a single mutation could contribute to the evolution of several characters between species [3] 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 present evidence 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 suggests that morphological evolution between species can occur through a single nucleotide change affecting several sexually dimorphic traits.

(126 words)

Keywords: pleiotropy, cis-regulatory, sensory organ, Drosophila, copulation, genitalia, hypandrium, sex comb, microevolution, *scute*, *Abd-B*

Graphical Abstract:



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 [10–13]: most 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 [14].

Here we focused on male sexual bristle evolution between *Drosophila yakuba* and *Drosophila santomea*, which diverged approximately 0.5-1 million years ago [15] and can produce fertile F1 females in the laboratory [16], 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 (Figure 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 (Figure 1, see also Extra Tables). No other genital bristle type was noticeably variable in number between *D. yakuba* and *D. santomea* (see Extra Figures).

We performed whole-genome QTL mapping between *D. santomea* and *D. yakuba* and found that the left tip of chromosome X explains 44% of the variance in hypandrial bristle number in each backcross (confidence interval = 7 Mb for the *D. santomea* backcross and 2.6 Mb for the *D. yakuba* backcross, Figure 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*) (Figure 2B-C, see also Extra Tables).

The *AS-C* locus contains four genes, but only two, *achaete* (*ac*) and *scute* (*sc*), are required for bristle formation [17]. Both genes are co-expressed, share *cis*-regulatory elements and act redundantly to specify bristles [18,19]. The elaborate expression pattern of *ac* and *sc* genes prefigures the adult bristle pattern and is controlled by numerous *cis*-regulatory elements [18]. We tested which of the two genes, *ac* or *sc*, contributes to loss of bristles using null mutants in *D. melanogaste*r. All ac^{CAM1} null mutant males had 2 hypandrial bristles (n=15) and sc^{M6} null mutants had none (n=15) (Tables S1-2), indicating that *sc* is required for hypandrial bristle development in *D. melanogaster*.

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 (Figure S1A, see also Extra Figures, Tables S1-2). 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 (Figures 2C and S1B-E, see also Extra Tables). Only one of these lines, *18C05*, increased hypandrial bristle number with *UAS-scute* in a *sc* mutant background or in a sc^{+} background (Figures 2C and S1F-Q, see also Extra Figures). The 2036-bp *18C05* region is located within the 5-kb candidate region identified with *ac-sc* structural mutations (Figure 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. 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 (Figure 3). 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}, Figure 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 (Figure 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 (Figure S2A-B, see also Extra Figures). 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 (Extra Figures). 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 (Figure S2, see also Extra Tables and Extra Figures). 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 (Figure 3, Extra Figures). 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) (Figure 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* (Figure S2C). 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 (Figure 3, GLM-Quasi-Poisson, $p>0.9$ in sc^{M6}). We conclude that at least three fixed substitutions

within a 350-bp region located 49 kb away from *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 [20] but also in male developing forelegs in the presumptive sex comb domain [21](Figure 4A-B,D-F). The *18C05-GFP* reporter constructs drive expression in fewer cells than *sc-GFP* (Fig. 4C), indicating that *sc* expression in the presumptive sex comb domain is also regulated by cis-regulatory regions outside of *18C05*. Sex combs are sensory organs used for grasping the female during copulation [22]. They differ in bristle number between *D. santomea* and *D. yakuba* (Figure 4G-I, see also Extra Figures), and 35% of the species difference is attributed to the X chromosome [23], where *sc* is located. These results prompted us to test whether the mutations contributing to hypandrial bristle evolution also affect sex combs. Significantly more GFP-positive cells were detected in the first tarsal segment at 5h after puparium formation (APF) with *18C05yakT1775G-GFP* than with *18C05yak-GFP* (GLM-Poisson , Chi-squared (20,2) deviance = 9.75, p = 0.033), suggesting that *T1775G* increases *sc* expression in the first tarsal segment. Sex comb tooth number was reduced in sc^{M6} and sc^{6} mutants and significantly rescued with several *18C05-sc* constructs (Figure 4J-K). Analysis of sc^{M6} and sc^{6} 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 (Figure 4J-K). 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 S3). *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 [24]. We found that reducing *Abd-B* expression, using either genetic mutations or RNA interference, resulted in loss of hypandrial bristles (Figure S3 and Table S4), 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 (Figure S4). 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. All our analyses of the effects of individual substitutions have been carried out in *D. melanogaster* background. It is thus possible that the *18C05* enhancer represents only part of the effect of the *sc* locus on bristle divergence.

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 [22,25]. Genitals are the most rapidly evolving organs in animals with internal fertilization [26]. 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 [27]. Second, an amino acid change in the *nath10 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 [28], which posits that certain characters may evolve arbitrarily as a result of selection on other traits due to pleiotropic mutations. In our

case, whether the evolutionary change in sex comb tooth number or in genital bristle number has any effect on fitness is unknown.

We report here the first experimental evidence for a cis-regulatory substitution between species 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 strict tissue-specific modules underlying evolutionary changes in targeted traits [29]. Even though cis-regulatory mutations may affect several tissues, it is probable that they still tend to be less pleiotropic than coding changes. Our results are thus compatible with the idea that cis-regulatory changes tend to have fewer pleiotropic effects than coding changes on average. Enhancer sequences evolve rapidly, with rapid turn over of individual binding sites while maintaining transcriptional output over millions of years by compensatory mutations [30]. 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.

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, F. Schweisguth for *GFP-sc* flies, J. Selegue and S.B. Carroll for the *Abd-B* construct, L. Pintard and N. Joly for help in protein purification, R. Mann, S. Feng and G. Rice for EMSA and immunostaining 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, C. Nobre, G. Verebes, 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.

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 (John Murray).
2. Saltz, J.B., Hessel, F.C., and Kelly, M.W. (2017). Trait Correlations in the Genomics Era. Trends Ecol. Evol. *32*, 279–290.
3. Paaby, A.B., and Rockman, M. V. (2013). The many faces of pleiotropy. Trends Genet. *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 (N. Y). *54*, 13–20.
6. Wagner, G.P., and Zhang, J. (2011). The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. Nat. Rev. Genet. *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 (80-.). *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 Rep. *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 Genet. *5*, e1000609.

13. Endler, L., Gibert, J.M., Nolte, V., and Schlötterer, C. (2018). Pleiotropic effects of regulatory variation in tan result in correlation of two pigmentation traits in Drosophila melanogaster. Mol. Ecol. *27*(16), 3207-3218.

14. Stern, D.L., and Orgogozo, V. (2008). The loci of evolution: How predictable is genetic evolution ? Evolution *62*, 2155–2177.

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

16. 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. Proceedings. Biol. Sci. *267*, 1487–1495.

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

18. 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.

19. 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.

20. 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 Rep. *2*, 1014–1024.

21. Tanaka, K., Barmina, O., Sanders, L.E., Arbeitman, M.N., and Kopp, A. (2011). Evolution of sex-specific traits through changes in HOX-dependent doublesex expression. PLoS Biol. *9*, e1001131.

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

23. 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. Genet. Res. (Camb). *84*, 11–26.

24. 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.

25. 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.

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

27. 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.

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

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

30. 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.

31. 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.

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

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

34. 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.

35. 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.

36. 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.

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

38. 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.

39. 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.

40. 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.

41. 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 Rep. *2*, 991–1001.

42. 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. Nat. Methods *6*, 343–345.

43. 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 Res. *36*, D102–D106.

44. 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 Res. *39*, D111–D117.

45. Culi, J., and Modolell, J. (1998). Proneural gene self-stimulation in neural precursors: an essential mechanism for sense organ development that is regulated byNotch signaling. Genes Dev. *12*, 2036–2047.

46. Jeong, S., Rokas, A., and Carroll, S.B. (2006). Regulation of Body Pigmentation by the Abdominal-B Hox Protein and Its Gain and Loss in Drosophila Evolution. Cell *125*, 1387–1399.

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

48. 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.

49. 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.

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

51. 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.

52. 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.

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

54. 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.

55. Maroni, G., and Stamey, S.C. (1983). Developmental profile and tissue distribution of alcohol dehydrogenase. Drosoph. Inf. Serv. *59*.

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

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

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

59. 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. (New York, NY: Springer).

60. Team, R.C. (2016). R: A language and environment for statistical.

61. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv Prepr. arXiv1406.5823.

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

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

64. Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scand. J. Stat., 65–70.

65 Corson, F., Couturier, L., Rouault, H., Mazouni, K., & Schweisguth, F. (2017). Self-organized Notch dynamics generate stereotyped sensory organ patterns in Drosophila. Science, *356*(6337), eaai7407.

66. Pfeiffer, B.D., Jenett, A., Hammonds, A.S., Ngo, T.-T.B., Misra, S., Murphy, C., Scully, A., Carlson, J.W., Wan, K.H., Laverty, T.R., et al. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. Proc. Natl. Acad. Sci. U. S. A. 105, 9715–9720.

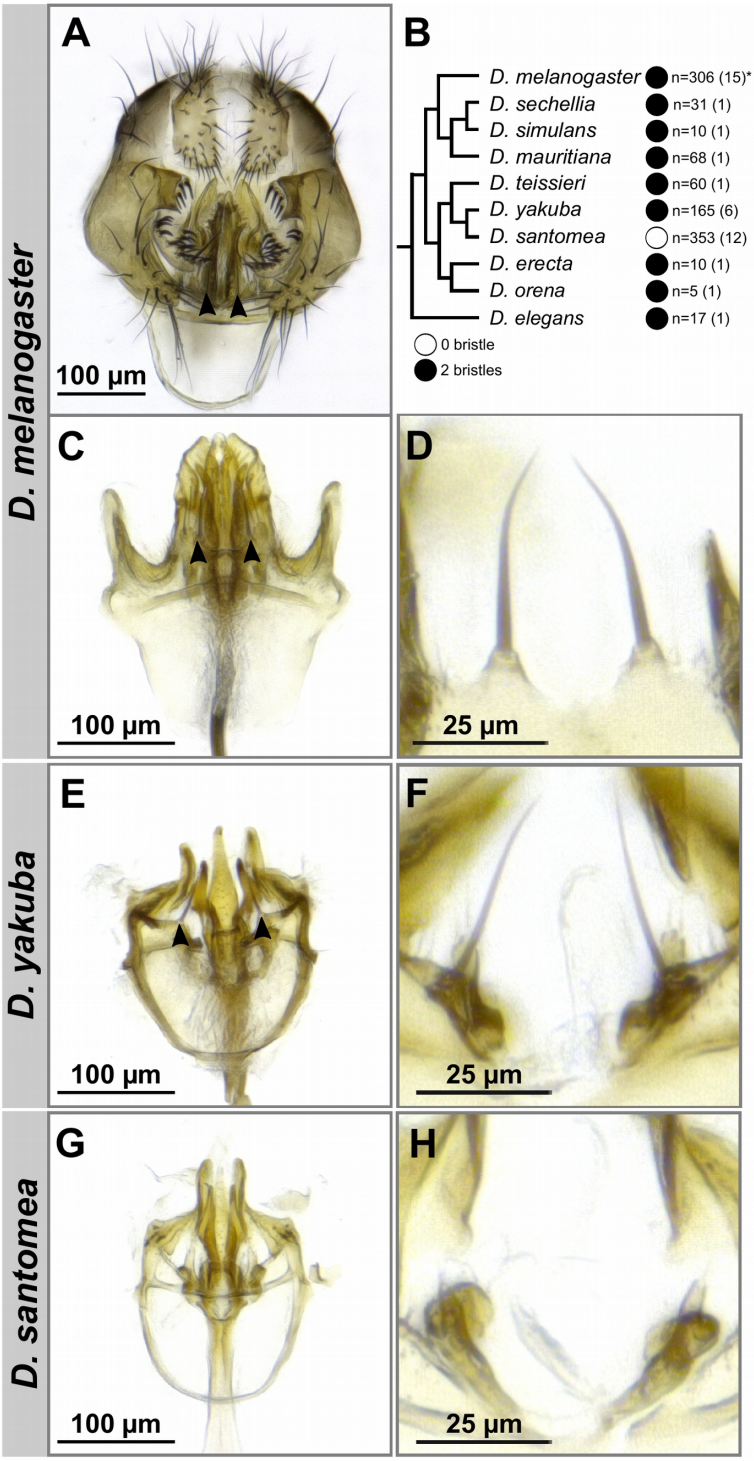


Figure 1. *D. santomea* lost hypandrial bristles.
(A) *Drosophila melanogaster* male genitalia. (B) Phylogeny of the *Drosophila melanogaster* species subgroup. All species of have two hypandrial bristles (black circles) except *Drosophila santomea*, which lacks hypandrial bristles (white circle). n: number of scored males, with the number of scored strains in parentheses. Asterisk indicates that 4 males out of 306 had three hypandrial bristles. (C-H) Light microscope preparations of ventral genitalia (C,E,G) 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.

385

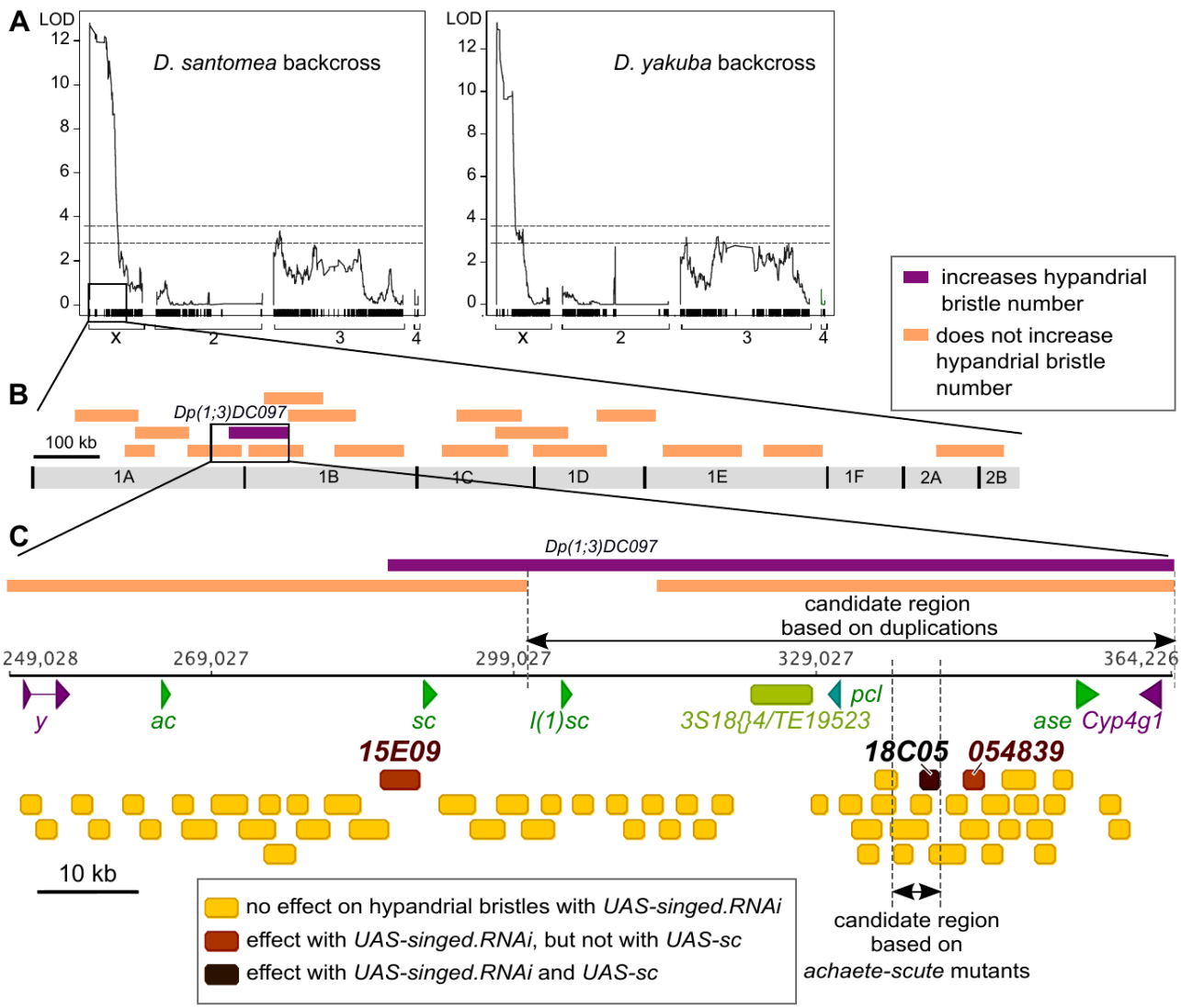


Figure 2. Mapping of the cis-regulatory element involved in hypandrial bristle evolution. (A) QTL analysis of hypandrial bristle number in a *D. santomea* backcross (left) and a *D. yakuba* backcross (right). On the y-axis are the LOD profiles from a Haley-Knott regression analysis. The x-axis represents 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).

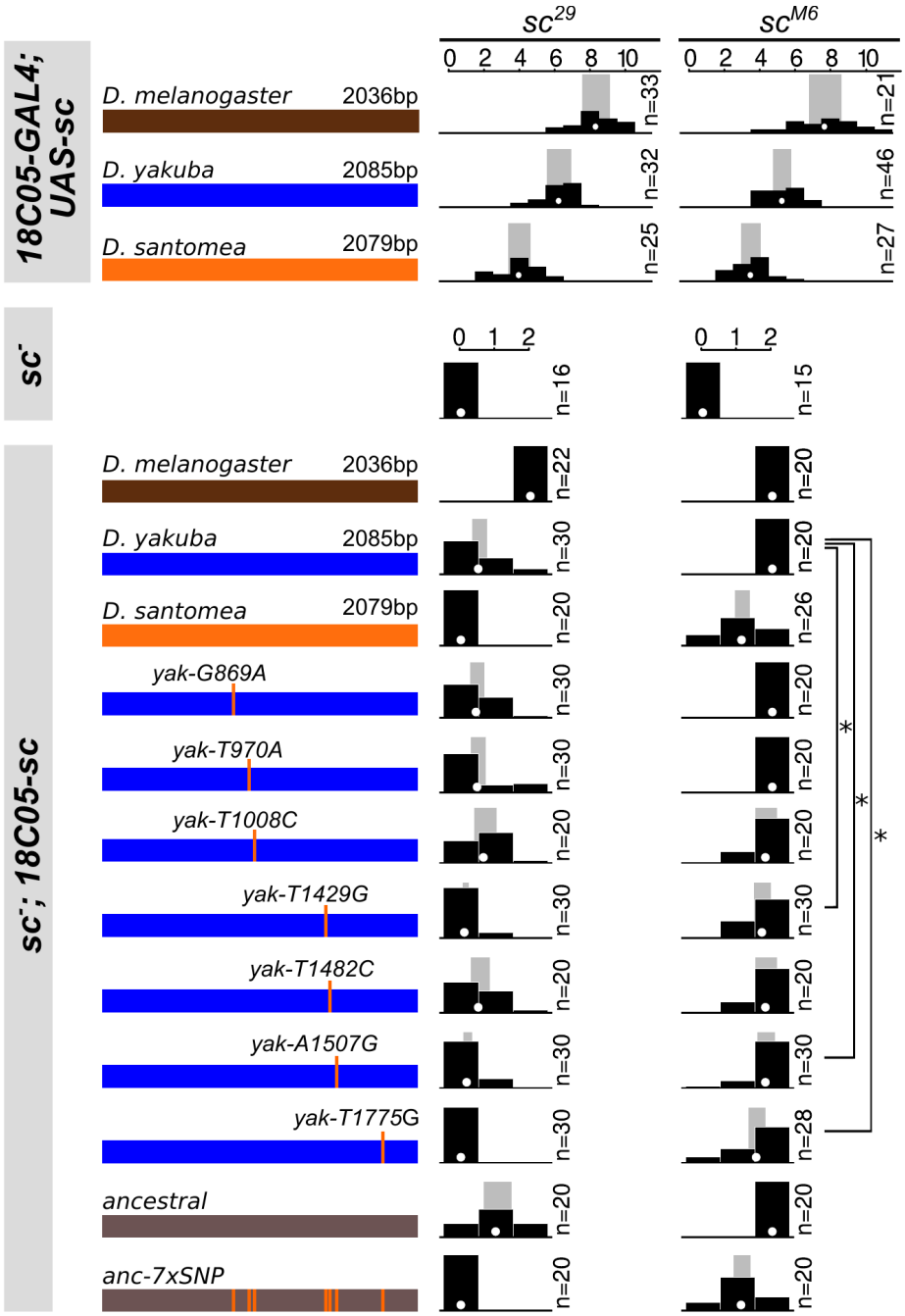


Figure 3. Three *D. santomea*-specific substitutions in *18C05* contribute to the loss of hypandrial bristles.

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. Note that for a given rescue construct, *18C05-GAL4 UAS-sc* produces more hypandrial bristles than *18C05-sc*, probably due to the amplification of gene expression caused by the *GAL4/UAS* system. n: number of scored individuals. *: $p<0.05$

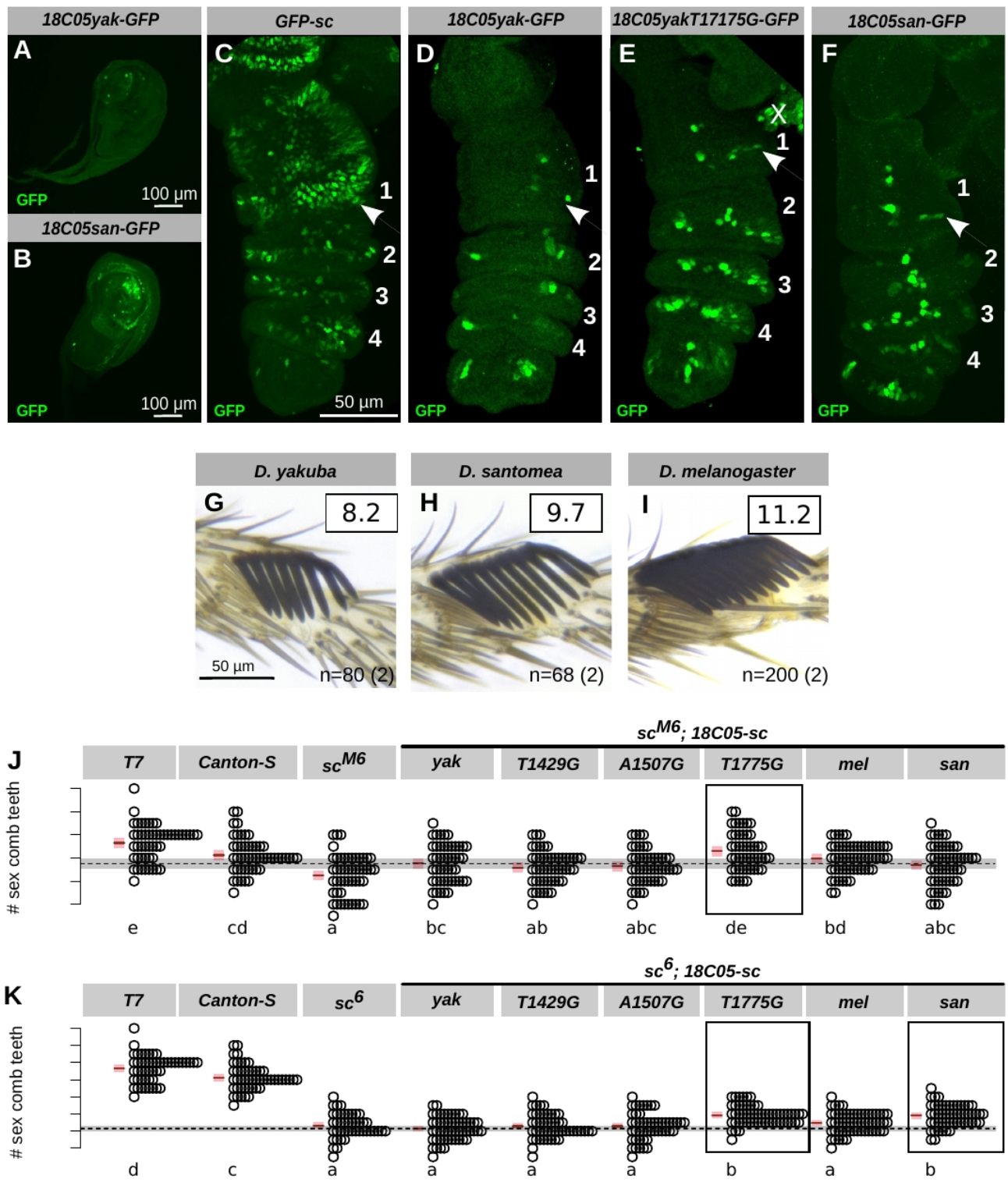


Figure 4. *D. santomea*-specific substitution *T1775G* contributes to increase in sex comb tooth number.

(A-F) GFP staining (green) in T1 leg discs of late L3 larvae (A-B) and in 5h APF pupal legs (C-F) in *D. melanogaster* containing *18C05* reporter transgenes or *sc-GFP.* Genotype is indicated on top of each panel. Tarsal segments are numbered. Arrowheads point to the presumptive sex comb regions. "X" indicates non-leg tissue. (G-I) Leg sex comb in *D. yakuba* (G), *D. santomea* (H) and *D. melanogaster* (I). Average sex comb tooth numbers per leg are shown in squares. n: number of scored individuals, with the number of scored strains in parentheses. (J-K) Sex comb tooth number in wild-type (T7 and Canton-S), sc^{M6} (J) and sc^{6} (K) mutants rescued with different *18C05-sc*

constructs. Each circle represents one male raised at 25°C. Mean (brown line) and 95% confidence interval (pink rectangle) from a fitted GLM Quasi-Poisson model are shown. Letters indicate the results of all-pairwise comparisons after Holm-Bonferroni correction. Two genotypes are significantly different from each other ($p < 0.05$) when they do not share a letter. For easier comparison, the horizontal dashed line and the surrounding grey line indicate the mean and 95% confidence interval for *sc⁻;18C05yak-sc*. Transgenic constructs with sex comb tooth number significantly different from *18C05yak-sc* are shown in boxes in J-K. On average *D. santomea* males have about 1 extra tooth per sex comb compared to *D. yakuba* (G-H). The substitution *T1775G* produces on average 0.5 extra sex comb tooth per leg, which is more than expected. It is possible that the *D. melanogaster* background, where all our rescue constructs were tested, amplifies the effect of the tested substitutions, especially since *D. melanogaster* males have more sex comb teeth than *D. santomea* or *D. yakuba*.

STAR*METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Virginie Courtier-Orgogozo (virginie.courtier@normalesup.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The origin of all the fly strains used can be found in KRT Table, Table S1 and Extra Tables. 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.

METHOD DETAILS

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 [31]. The list of barcodes used in this study can be found in Mendeley (http://dx.doi.org/10.17632/xjvz2m8z6r.1) or in the Extra Data files, 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 [32,33]. 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 [34] which 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) (Extra Tables). 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 (Extra Data Files). 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 [35]. 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 [36]. 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*) (Figure 2.B-C, see also Extra Tables, 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 [26], 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 (Extra Figures). 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 with Image J [37].

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 [38] and Bloomington Stock Center (Extra Tables). 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 [39], 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-GFP, 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 KRT Table, see also Extra Tables) together w*ith DC-GAL4*, which drives expression in the dorso-central thoracic bristl*e*s [40]. To enhance RNAi potency we also used *UAS-Dicer-2* [40]. 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*^{05747} at 29 °C (Extra Tables). 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 (Extra Tables).

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.

<u>Cloning of *enhancers* into pBPGUw and pBPSUw</u>

Enhancers were cloned into the *GAL4* reporter vector pBPGUw using the same strategy as in [38,41]. Enhancer sequences were amplified by Phusion® High Fidelity Polymerase (New England Biolabs) in two steps reaction using the primers and templates (Extra Tables). 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) (Extra Tables). 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 (Figure 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 described in Extra Tables. 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 [41] 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 [42] by fusing together different lengths of *18C05* sequences from *D. yakuba Ivory Coast* and *D. santomea SYN2005*. The different chimeras are described in Extra Tables. Cloning primers were designed using NEBuilder Tools (http://nebuilder.neb.com/). Primer sequences and templates used in PCR are listed at Extra Tables. To assemble the *18C05* fragments in pBPSUw (Extra Tables), 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 [42], 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 [42], 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® (Extra Tables). 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 (Extra Tables).

All transgenic constructs were integrated into the *attP2* landing site in *D. melanogaster* w^{1118} by BestGene Inc. The *T1775G* substitution affects nucleotide position 447,055 in the Dm6 reference assembly.

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 the summary of the alignment of *18C05* sequences available at https://doi.org/10.6084/m9.figshare.6972707. *18C05* sequences were amplified with San-Yak_lines_sequencing-For and San-Yak_lines_sequencing-Rev primers (Extra Figures) 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 multiple *Drosophila* species available at https://doi.org/10.6084/m9.figshare.6972707 . 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 [43]). 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 [43] (Table S3). For TFs which were absent in JASPAR (Scute), we used Fly Factor Survey [44] to analyze their putative binding affinities to the probe. As *sc* cis-regulatory region is known to contain binding sites for Scute itself [45], we looked for Scute binding sites in *18C05*, *15E09* and *054839*. Two putative Scute binding sites (consensus motif *CAYCTGY*, Fly Factor Survey [44] were found in *15E09* and *054839* but not in *18C05*. Given the present results, we cannot exclude the involvement of *15E09* and *054839* in the evolution of hypandrial bristle evolution in *D. santomea*. In this paper, we decided to focus on the *18C05* enhancer, whose effect could be studied in a *sc* mutant background.

<u>Abd-B homeodomain (Abd-B-HD) purification and EMSA</u>

The Abd-B-HD-pGEX-4T-1 plasmid [46] (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 [47] 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 Extra Tables. 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 (Invitrogen) 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 (Figure S4E) using the intensity values of the bands measured in ImageJ [37]. 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 (Figure S4B-D). 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 S4). Using the genitalia *GAL4* drivers *esg-GAL4*^{NP5130} [48] and *NP6333* [49] 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 (Figure S3A-F, n=9/11 for *NP5130*, n=8/9 for *NP6333*, Table S4). Smaller bristles might arise from a delay in *sc* expression during development [50]. Since *Abd-B* null mutations are lethal [51], we produced mitotic mutant clones for two null mutations, *Abd-B*^{M1} [51] and *Abd-B*^{D18} [52]. *Abd-B* mutant mitotic recombinant clones were induced by the FLP/FRT system [53] 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 [54]. From both crosses, a total of 82 F1 males (Table 4) 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 (Figure S4G-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 (Figure S4I-N, Table S4). 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

For leg disc stainings the 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 [55]. Larvae were chosen with the most clear gut, indicating a developmental stage of 1-6 hours before pupa formation [56]. Head parts of the larvae were cut and fixed in 4% PFA in PBS pH 7.4 for 20 minutes at room temperature. For pupal leg preparations the anterior part of the pupae were cut and fixed in 4% PFA in PBS pH 7.4 for 50 minutes at room temperature. Following fixation, samples were 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 and pupal legs were dissected in PBS and mounted in Vectashield® H-1000. Images were acquired using Spinning Disc CSU-W1. Number of GFP-positive cells were counted in the z-stack using ImageJ [37] in a blind fashion regarding the genotypes using randomized file names.

QUANTIFICATION AND STATISTICAL ANALYSIS

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 [57–59]. All statistical analyses were performed using R 3.4 [60]. GLM were fitted with the function glm() ("stats" core package 3.5.0) and GLMM with the function glmer() ("lme4" package 1.1-14 [61] 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 [59] 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 [62]) which performs multiple comparisons between fitted GLM parameters and yields adjusted p-values corrected according to the Holm-Bonferroni method [63,64] 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.

DATA AVAILABILITY

Sequences were deposited into GenBank (accession numbers MG460736-MG460765). Source data for Bristle Number, QTL mapping analysis, EMSA and immunostaining are available in BioRxiv and at Mendeley: http://dx.doi.org/10.17632/xjvz2m8z6r.1. Additional Data Figures and Data Tables are available in BioRxiv and at Figshare: https://doi.org/10.6084/m9.figshare.6972707 and https://doi.org/10.6084/m9.figshare.6972740, respectively.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit anti-GFP primary antibody	Thermofisher	#A6455
donkey Anti-Rabbit Dylight® 488	Thermofisher	#SA5-10038
Chemicals, Peptides, and Recombinant Proteins		
Dimethyl Hydantoin Formaldehyde	Entomopraxis	N/A
Paraformaldehyde	Sigma-Aldrich	#158127-5G
Dapi	Sigma-Aldrich	#D9542-1MG
Vectashield® H-1000	Vector Laboratories	#H-1000
reduced glutation	Sigma-Aldrich	#G-4251
Abdominal-B-HD protein	This paper	N/A
Critical Commercial Assays		
Qiagen DNeasy Blood & Tissue extraction kit	Qiagen	#69506
Nucleospin Gel and PCR Clean-Up Kit	Machery-Nagel	#740609
E.Z.N.A.® Plasmid Mini Kit I	Omega Bio-tek	#D6942-01
Quick-DNA™ Miniprep Plus Kit	Zymo Research	#D4069
LightShiftTM Chemiluminiscent EMSA Kit	ThermoFisher Scientific	#20148
TurboBlotter Kit	GE Healthcare Life Sciences	#10416314
Deposited Data		
Raw and analyzed data	This paper	Mendeley: http://dx.doi.org/10.17632/xjvz2m8z6r.1 (DOI: 10.17632/xjvz2m8z6r.1)
Genitalia Bristle Number	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972707.v1
Alignment of *18C05* sequence from *D. santomea* SYN2005 and *D. yakuba* (Ivory Cost)	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972707.v1
Twelve substitutions are fixed in *D. santomea* 18C05	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972707.v1
Sex comb tooth number in *D. yakuba* and *D. santomea* and *D. melanogaster scute* mutants	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972707.v1
Hypandrial bristle number in pure species and F1 hybrids	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972740.v3
Test of various *UAS*-reporter constructs with *DC-GAL4*	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972740.v3

Achaete-scute GAL4 lines and their hypandrial bristle phenotype with *Dcr2*; *UAS-singed.RNAi*^{10574}	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972740.v3
Cloning strategy	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972740.v3
18C05_D. san_A1200.4	This paper	GenBank #MG460738
18C05_D. san_car1490.3	This paper	GenBank #MG460742
18C05_D. san_BS14.1	This paper	GenBank #MG460740
18C05_D. san_51.7.3_1560	This paper	GenBank #MG460737
18C05_D. san_OBAT_1200.13	This paper	GenBank #MG460744
18C05_D. san_Quija650.37	This paper	GenBank #MG460747
18C05_D. san_Quija650.22	This paper	GenBank #MG460746
18C05_D. san_Quija650.14	This paper	GenBank #MG460745
18C05_D. san_C1350.14	This paper	GenBank #MG460741
18C05_D. san_B1300.13	This paper	GenBank #MG460739
18C05_D. san_Rain42	This paper	GenBank #MG460748
18C05_D. san_Field3.4	This paper	GenBank #MG460743
18C05_D. san_STO4	This paper	GenBank #MG460749
18C05_D. san_SYN2005	This paper	GenBank #MG460750
18C05_D. yak_15.6.8	This paper	GenBank #MG460759
18C05_D. yak_LP1	This paper	GenBank #MG460761
18C05_D. yak_2.22.1	This paper	GenBank #MG460756
18C05_D. yak_4.23.1	This paper	GenBank #MG460757
18C05_D. yak_4.32.1	This paper	GenBank #MG460758
18C05_D. yak_Tai18E2	This paper	GenBank #MG460765
18C05_D. yak_Ivory Coast	This paper	GenBank #MG460760
18C05_D. yak_Tai18E2 (NCBI)	This paper	GenBank #MG460765
18C05_D. yak_PB 3.1.3	This paper	GenBank #MG460763
18C05_D. yak_PB 3.4.1	This paper	GenBank #MG460764
18C05_D. yak_PB1.4.21	This paper	GenBank #MG460762
18C05_D. yak_5.3.1	This paper	GenBank #MG460755

18C05_D. teis_(Mt. Selinda)	This paper	GenBank #MG460753
18C05_D. teis_(SDSC#14021-0257.01)	This paper	GenBank #MG460754
18C05_D. mel_BL2057	This paper	GenBank #MG460736
18C05_D. sim_w501	This paper	GenBank #MG460752
18C05_D. sim_M252	This paper	GenBank #MG460751
Experimental Models: Organisms/Strains		
D.melanogaster 3870	Tony Long UC Irvine, RVC 3. Collected from Riverside, California, USA in 1963	N/A
D.melanogaster 3844	Tony Long UC Irvine, BS1. Collected from Barcelona, Spain in 1954	San Diego Stock Center #14021-0231.60
D.melanogaster 3841	Tony Long UC Irvine, BOG1. Collected from Bogota, Colombia in 1962	San Diego Stock Center #14021-0231.59
D.melanogaster 3852	Tony Long UC Irvine, KSA2. Collected in 1963	San Diego Stock Center #14021-0231.64
D.melanogaster 3864	Tony Long UC Irvine, KI2. Collected from Israel in 1954	San Diego Stock Center # 14021-0231.68
D.melanogaster T.7	Tony Long UC Irvine, Collected from Taiwan in 1968	San Diego Stock Center #14021-0231.07
D.melanogaster T.4	Tony Long UC Irvine, Collected from Kuala Lumpur, Malaysia in 1962	San Diego Stock Center #14021-0231.04
D.melanogaster 3875	Tony Long UC Irvine, VAG1. Collected from Athens, Greece in 1965	San Diego Stock Center #14021-0231.69
D.melanogaster 3886	Tony Long UC Irvine, Wild 5B. Collected from Red Top Mountain, Georgia in 1966	N/A
D.melanogaster T.1	Tony Long UC Irvine, Collected from Ica, Peru in 1956	San Diego Stock Center #14021-0231.04
D.melanogaster 3839	Tony Long UC Irvine, BER1. Collected from Bermudas in 1954.	San Diego Stock Center # 14021-0231.58
D.melanogaster 3846	Tony Long UC Irvine, CA1. Collected from Cape Town, South Africa.	San Diego Stock Center #14021-0231.62

D.melanogaster Sam	Tony Long UC Irvine, DSPR line. originally from TFC Mackay Sam; ry506	N/A
D.melanogaster iso-1 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]	Bloomington Stock Center	Bloomington Stock Center #2057
D.melanogaster Canton-S	Roger Karess	Kyoto DGGR #105666
D.melanogaster *dor[4]/C(1)RM, y[1] w[1] f[1]*	Bloomington Stock Center	Bloomington Stock Center #35
D.melanogaster *Nup98-96[339]/TM3, Sb[1]*	Bloomington Stock Center	Bloomington Stock Center #4951
D.melanogaster *Df(3R)D605/TM3, Sb[1] Ser[1]*	Bloomington Stock Center	Bloomington Stock Center #823
D.melanogaster DC002 *w1118; Dp(1;3)DC002, PBac{DC002}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30213
D.melanogaster DC003 *w1118; Dp(1;3)DC003, PBac{DC003}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30214
D.melanogaster DC004 *w1118; Dp(1;3)DC004, PBac{DC004}VK00033/TM6C, Sb1*	Bloomington Stock Center	Bloomington Stock Center #30215
D.melanogaster DC006 *w1118; Dp(1;3)DC006, PBac{DC006}VK00033/TM6C, Sb1*	Bloomington Stock Center	Bloomington Stock Center #30217
D.melanogaster DC097 *w1118; Dp(1;3)DC097, PBac{DC097}VK00033/TM6C, Sb1*	Bloomington Stock Center	Bloomington Stock Center #31440
D.melanogaster DC098 *w1118; Dp(1;3)DC098, PBac{DC098}VK00033*	Bloomington Stock Center	Bloomington Stock Center #31441
D.melanogaster DC007 *w1118; Dp(1;3)DC007, PBac{DC007}VK00033/TM6C, Sb1*	Bloomington Stock Center	Bloomington Stock Center #30218
D.melanogaster DC008 *w1118; Dp(1;3)DC008, PBac{DC008}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30745
D.melanogaster DC009 *w1118; Dp(1;3)DC009, PBac{DC009}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30219
D.melanogaster DC012 *w1118; Dp(1;3)DC012, PBac{DC012}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30222
D.melanogaster DC099 *w1118; Dp(1;3)DC099, PBac{DC099}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30749
D.melanogaster DC013 *w1118; Dp(1;3)DC013, PBac{DC013}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30746
D.melanogaster DC014 *w1118; Dp(1;3)DC014, PBac{DC014}VK00033*	Bloomington Stock Center	Bloomington Stock Center #31434
D.melanogaster DC400 *w1118; Dp(1;3)DC400, PBac{DC400}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30795
D.melanogaster DC019 *w1118; Dp(1;3)DC019, PBac{DC019}VK00033*	Bloomington Stock Center	Bloomington Stock Center #30223
D.melanogaster DC436 *w1118; Dp(1;3)DC436, PBac{DC436}VK00033/TM6C, Sb1*	Bloomington Stock Center	Bloomington Stock Center #33487
D.melanogaster DC401 w1118; Dp(1;3)DC401, PBac{DC401}VK00033	Bloomington Stock Center	Bloomington Stock Center #30796
D.melanogaster DC-GAL4 *yw ; DC-GAL4 , UAS-GFP /TM6B*	V. Stamataki (Pat Simpson lab)	N/A
UAS-forked.RNAi^{33200}	Vienna Stock Center	VDRC #33200

D.melanogaster UAS-singed.RNAi^{105747}	Vienna Stock Center	VDRC #105747
D.melanogaster UAS-forked.RNAi^{24632}	Vienna Stock Center	VDRC #24632
D.melanogaster UAS-ac.RNAi^{100647}	Vienna Stock Center	VDRC #100647
D.melanogaster UAS-singed.RNAi^{32579}	Vienna Stock Center	VDRC #32579
D.melanogaster UAS-forked.RNAi^{103813}	Vienna Stock Center	VDRC #103813
D.melanogaster UAS-sc.RNAi^{105951}	Vienna Stock Center	VDRC #105951
D.melanogaster yw; UAS-y *y[1] w[1118]; P{w[+mC]=UAS-y.C}MC1*	Vienna Stock Center	Bloomington Stock Center #3043
D.melanogaster yw;UAS-y TM3/pnr-GAL4	Mark Rebeiz.	N/A
D.melanogaster UAS-mCD8-GFP GFP transgene on second chromosome	Veronique Brodu	N/A
D.melanogaster w,UAS-Dcr2 ; Pin/CyO	Bloomington Stock Center	Bloomington Stock Center #24644
D.melanogaster UAS-scute	Bloomington Stock Center	Bloomington Stock Center #51672
D.melanogaster GFP-sc *GFP* inserted at the *scute* locus by CRISPR-mediated homologous recombination, which produces Scute protein with GFP sequence fused at the N terminus.	F. Schweisguth, [65]	N/A
D.melanogaster yw;UAS-Abd-B.RNAi^{51167}	Bloomington Stock Center	Bloomington Stock Center #51167
D.melanogaster yw; UAS-Abd-B.RNAi^{26746}	Bloomington Stock Center	Bloomington Stock Center #26746
D.melanogaster yw; NP5130-GAL4	Kyoto DGGR	Kyoto DGGR #109126
D.melanogaster yw; NP6333-GAL4	Kyoto DGGR	Kyoto DGGR #113920
D. simulans	Collected by J. R. David from Marrakech, Morocco in 2010	N/A
D. mauritiana	Collected by J. R. David from Mauritius Island in 1985	N/A
D. sechellia GFP *w[1] ; pBac(3xP3-EGFPafm)::MCS::(pW8 mini-white)*	San Diego Species Stock Center	San Diego Stock Center #14021-0248.32
D. yakuba Ivory Coast	D. L. Stern, Collected from Ivory Coast in 1955	San Diego Stock Center #14021-0261.00
D. yakuba 15.6.8, Isofemale stock	Collected by D. R. Matute in São Tomé at altitude 110m in 2009	N/A
D. yakuba yellow[1]	San Diego Species Stock Center	San Diego Species Stock Center #14021-0261.05
D. yakuba 4.23.1, Isofemale stock	Collected by D. R. Matute in São Tomé at altitude 1070m in 2009	N/A
D. yakuba LP1, Isofemale stock	Collected by D. R. Matute in São Tomé at altitude 0m in 2009	N/A
D. yakuba 2.22.1, Isofemale stock	Collected by D. R. Matute	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1250m in 2009 by D. Matute

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

D. santomea Rain42, Isofemale stock	Collected by D. R. Matute	Given by D. Matute. Isofemale stock, collected in São Tomé at altitude 1240m in 2009 by D. Matute
D. santomea Field3.4, Isofemale stock	Collected by D. R. Matute in São Tomé at altitude 1250m in 2009	N/A
D. teissieri	Collected by J. R. David in Mt Selinda, Zimbabwe in 1970	N/A
D. teissieri #14021-0257.01	San Diego Stock Center	San Diego Stock Center # 14021-0257.01
D. orena, Isofemale stock	Collected by J. R. David in 1975 in Cameroon	N/A
D. erecta	Collected by D. Lachaise in La Lopé, Gabon in 2005	N/A
D. elegans	B. Prud'homme, Collected in Hong-Kong.	San Diego Stock Center # 14027-0461.03.
D. melanogaster sc^{M6} *sc[M6]/FM7i, P{w[+mC]=ActGFP}JMR3*	Bloomington Stock Center	#52668
D. melanogaster ac^{CAMI} *y[1] P{w[+mW.hs]=GawB}CG32816[NP6014] ac[cami]*	Bloomington Stock Center	#36540
D. melanogaster sc^{6} *sc[6] w[a]*	Bloomington Stock Center	#108
D. melanogaster ase^{1} *Df(1)ase-1, sc[ase-1] pn[1]/C(1)DX, y[1] f[1]*	Bloomington Stock Center	#104
D. melanogaster sc^{5} *y[1] sc[5]*	Bloomington Stock Center	#178
D. melanogaster ac^{1} *y[1] ac[1] w[1118]; P{w[+mC]=GAL4-ac.13}1*	Bloomington Stock Center	#8715
D. melanogaster sc^{1} *y[1] sc[1]*	Bloomington Stock Center	#176
D. melanogaster ac^{sbm} *ac[sbm]*	Given by P. Simpson	N/A
D. melanogaster ac^{Hw-1} *Df(1)sc10-1, sc[10-1]/y[1] ac[Hw-1]*	Bloomington Stock Center	#109
D. melanogaster sc^{29} *In(1)sc[29], sc[29] w[a] eag[sc29]*	Bloomington Stock Center	#1442
D. melanogaster ac^{1} sc^{1} *y[1] ac[1] sc[1] pn[1]*	Bloomington Stock Center	#4596
D. melanogaster sc^{H} *C(1)DX, y[1] f[1]; T(1;4)sc[H], sc[H]*	Bloomington Stock Center	#4055
D. melanogaster sc^{9} *In(1)sc[9], sc[9] w[a] f[1] Bx[1]*	DGRC Kyoto Stock Center	#102028
D. melanogaster sc^{S2} *T(1;2)sc[S2], y[+] sc[S2]: cn[1] M(2)53[1]/+; CyO*	Bloomington Stock Center	#3333
D. melanogaster sc^{7} *Df(1)B/In(1)sc[7], In(1)AM, sc[7] ptg[4]*	Bloomington Stock Center	#723
D. melanogaster ac^{3} sc^{10-1} *In(1)ac[3], sc[10-1] ac[3] w[1] sable[1]/FM7i, P{w[+mC]=ActGFP}JMR3*	Bloomington Stock Center	#36541
D. melanogaster sc^{4} *In(1)sc[4], y[1] sc[4] ABO-X[1]*	Bloomington Stock Center	#789

D. melanogaster sc^{8} *T(1;3)sc[260-15], sc[260-15]/FM6 B[1] dm[1] sc[8] y[31d]*	Bloomington Stock Center	#842
D. melanogaster ac^{1} sc^{19} *Df(1)sc[19]/y[1] ac[1]; Dp(1;2)sc[19]/In(2L)Cy, S[2] Cy[1]*	Bloomington Stock Center	#3822
D. melanogaster VT054793-GAL4	Vienna Drosophila Research Center	VT054793
D. melanogaster VT054794-GAL4	Vienna Drosophila Research Center	VT054794
D. melanogaster VT054795-GAL4	Vienna Drosophila Research Center	VT054795
D. melanogaster VT054796-GAL4	Vienna Drosophila Research Center	VT054796
D. melanogaster VT054798-GAL4	Vienna Drosophila Research Center	VT054798
D. melanogaster VT054799-GAL4	Vienna Drosophila Research Center	VT054799
D. melanogaster GMR14C10-GAL4	Janelia Research Campus	GMR14C10
D. melanogaster GMR15B10-GAL4	Janelia Research Campus	GMR15B10
D. melanogaster GMR15C11-GAL4	Janelia Research Campus	GMR15C11
D. melanogaster GMR15X09-GAL4	This paper	N/A
D. melanogaster VT054805-GAL4	Vienna Drosophila Research Center	VT054805
D. melanogaster GMR15A01-GAL4	Janelia Research Campus	GMR15A01
D. melanogaster GMR14C12-GAL4	Janelia Research Campus	GMR14C12
D. melanogaster GMR15A04-GAL4	Janelia Research Campus	GMR15A04
D. melanogaster GMR15C10-GAL4	Janelia Research Campus	GMR15C10
D. melanogaster GMR15E07-GAL4	Janelia Research Campus	GMR15E07
D. melanogaster GMR15E09-GAL4	Janelia Research Campus	GMR15E09
D. melanogaster GMR13D04-GAL4	Janelia Research Campus	GMR13D04
D. melanogaster GMR13C08-GAL4	Janelia Research Campus	GMR13C08
D. melanogaster GMR12H02-GAL4	Janelia Research Campus	GMR12H02
D. melanogaster GMR13B12-GAL4	Janelia Research Campus	GMR13B12
D. melanogaster VT054820-GAL4	Vienna Drosophila Research Center	VT054820
D. melanogaster VT054821-GAL4	Vienna Drosophila Research Center	VT054821
D. melanogaster VT054822b-GAL4	This paper	N/A
D. melanogaster VT054823-GAL4	Vienna Drosophila Research Center	VT054823
D. melanogaster VT054824-GAL4	Vienna Drosophila Research Center	VT054824
D. melanogaster VT054825-GAL4	Vienna Drosophila Research Center	VT054825
D. melanogaster VT054826-GAL4	Vienna Drosophila Research Center	VT054826

D. melanogaster VT054827-GAL4	Vienna Drosophila Research Center	VT054827
D. melanogaster VT054828-GAL4	Vienna Drosophila Research Center	VT054828
D. melanogaster VT054829-GAL4	Vienna Drosophila Research Center	VT054829
D. melanogaster VT054831-GAL4	Vienna Drosophila Research Center	VT054831
D. melanogaster VT054832-GAL4	Vienna Drosophila Research Center	VT054832
D. melanogaster GMR18G09-GAL4	Janelia Research Campus	GMR18G09
D. melanogaster VT054833-GAL4	Vienna Drosophila Research Center	VT054833
D. melanogaster GMR18E07-GAL4	Janelia Research Campus	GMR18E07
D. melanogaster VT054834-GAL4	Vienna Drosophila Research Center	VT054834
D. melanogaster GMR19D04-GAL4	Janelia Research Campus	GMR19D04
D. melanogaster VT054835-GAL4	Vienna Drosophila Research Center	VT054835
D. melanogaster VT054836-GAL4	Vienna Drosophila Research Center	VT054836
D. melanogaster GMR18C05-GAL4	Janelia Research Campus	GMR18C05
D. melanogaster GMR19B11-GAL4	Janelia Research Campus	GMR19B11
D. melanogaster VT054838-GAL4	Vienna Drosophila Research Center	VT054838
D. melanogaster GMR18G07-GAL4	Janelia Research Campus	GMR18G07
D. melanogaster VT054839-GAL4	Vienna Drosophila Research Center	VT054839
D. melanogaster GMR18F05-GAL4	Janelia Research Campus	GMR18F05
D. melanogaster VT054840-GAL4	Vienna Drosophila Research Center	VT054840
D. melanogaster VT054841-GAL4	Vienna Drosophila Research Center	VT054841
D. melanogaster GMR19A06-GAL4	Janelia Research Campus	GMR19A06
D. melanogaster VT054842-GAL4	Vienna Drosophila Research Center	VT054842
D. melanogaster GMR18E10-GAL4	Janelia Research Campus	GMR18E10
D. melanogaster VT054843-GAL4	Vienna Drosophila Research Center	VT054843
D. melanogaster GMR20B05-GAL4	Janelia Research Campus	GMR20B05
D. melanogaster VT054845-GAL4	Vienna Drosophila Research Center	VT054845
D. melanogaster VT054846-GAL4	Vienna Drosophila Research Center	VT054846
D. melanogaster VT054839mel-BL2057-GAL4	This paper	VT054839mel-BL2057
D. melanogaster VT054839yak-GAL4	This paper	N/A
D. melanogaster VT054839san-GAL4	This paper	N/A

E. coli One Shot® TOP10	Invitrogen	#C404003
E. coli NEB10-beta	New England Biolabs	#C3019H
E. coli BL21 (DE3)	Nicolas Joly	N/A
Oligonucleotides		
Primers	This paper	Figshare: https://doi.org/10.6084/m9.figshare.6972740.v3
Recombinant DNA		
pBPGUw	Addgene, [66]	#17575
pBPSUw: GAL4 cassette of pBPGUw replaced by *scute* CDS.	This paper	N/A
15X09-pBPGUw	This paper	N/A
VT054822b-pBPGUw	This paper	N/A
VT054839yak-pBPGUw	This paper	N/A
VT054839san-pBPGUw	This paper	N/A
18C05_T7-pBPGUw	This paper	N/A
18C05_BL2057-pBPGUw	This paper	N/A
18C05Yakfull-pBPGUw	This paper	N/A
18C05Sanfull-pBPGUw	This paper	N/A
18C05AmelBL-pBPGUw	This paper	N/A
18C05BmelBL-pBPGUw	This paper	N/A
18C05CmelBL-pBPGUw	This paper	N/A
18C05ABmelBL-pBPGUw	This paper	N/A
18C05BCmelBL-pBPGUw	This paper	N/A
18C05Asan-pBPGUw	This paper	N/A
18C05Bsan-pBPGUw	This paper	N/A
18C05Ayak-pBPGUw	This paper	N/A
18C05Byak-pBPGUw	This paper	N/A
18C05Cyak-pBPGUw	This paper	N/A
VT054839mel-BL2057-pBPGUw	This paper	N/A
VT054839yak-pBPGUw	This paper	N/A
VT054839san-pBPGUw	This paper	N/A
18C05_SSSY-pBPSUw	This paper	N/A
18C05_YYSS-pBPSUw	This paper	N/A
18C05_SSYS-pBPSUw	This paper	N/A
18C05_SYYY-pBPSUw	This paper	N/A
18C05_YYYS-pBPSUw	This paper	N/A
18C05_SSYY-pBPSUw	This paper	N/A
18C05_YYSY-pBPSUw	This paper	N/A
18C05_YSYY-pBPSUw	This paper	N/A
18C05_YSSY-pBPSUw	This paper	N/A
18C05yakT970A-pBPSUw	This paper	N/A
18C05YyakT1429G-pBPSUw	This paper	N/A
18C05yakA1507G-pBPSUw	This paper	N/A
18C05yakt1775G-pBPSUw	This paper	N/A
18C05_Anc-pBPSUw	This paper	N/A
18C05_AncT1008C-pBPSUw	This paper, GeneScript	N/A
18C05_AncT1429G-pBPSUw	This paper, GeneScript	N/A
18C05_AncT1482C-pBPSUw	This paper, GeneScript	N/A
18C05_AncA1507G-pBPSUw	This paper, GeneScript	N/A
18C05_AncT1775G-pBPSUw	This paper, GeneScript	N/A
18C05_yakG869A-pBPSUw	This paper, GeneScript	N/A
18C05_yakT1008C-pBPSUw	This paper, GeneScript	N/A

18C05_yakT1482C-pBPSUw	This paper, GeneScript	N/A
18C05_AncG869A-pBPSUw	This paper, GeneScript	N/A
18C05_AncT670G-pBPSUw	This paper, GeneScript	N/A
18C05_Anc-SNPall-pBPSUw	This paper, GeneScript	N/A
Abd-B-HD-pGEX-4T-1	S. B. Carroll, [46]	N/A
Software and Algorithms		
Nebuilder Tools	New England Biolabs	https://nebuilder.neb.com/
Jaspar	[43]	http://jaspar.genereg.net
R 3.4	[60]	https://cran.r-project.org/
ImageJ	[37]	https://imagej.nih.gov/ij/download.html
Geneious	Biomatters Ltd.	https://www.geneious.com/download/