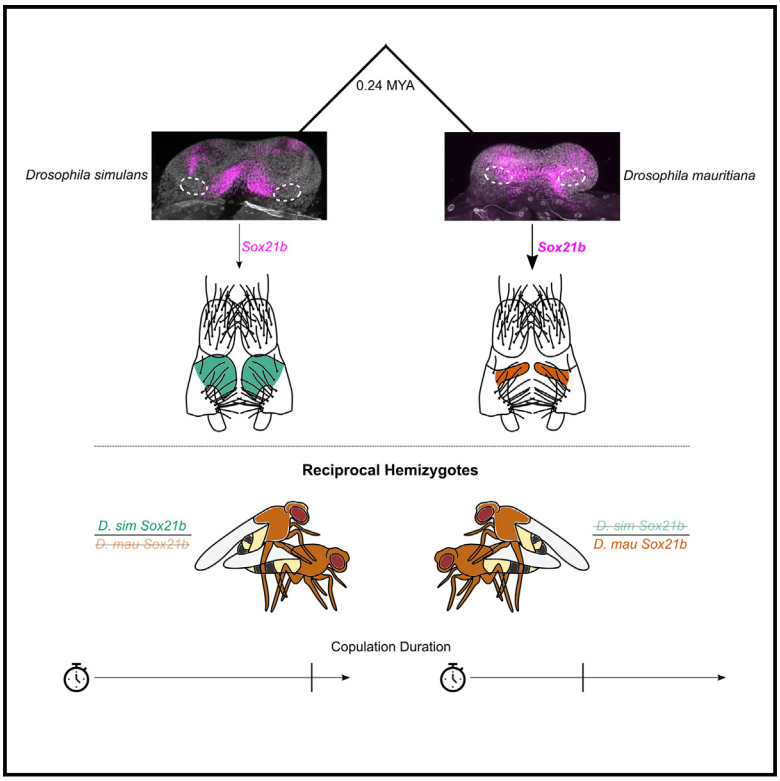
Report

Current Biology

Sox21b underlies the rapid diversification of a novel male genital structure between *Drosophila* species

Graphical abstract



Authors

Amber M. Ridgway, Emily J. Hood, Javier Figueras Jimenez, Maria D.S. Nunes, Alistair P. McGregor

Correspondence

msantos-nunes@brookes.ac.uk (M.D.S.N.), alistair.mcgregor@durham.ac.uk (A.P.M.)

In brief

Ridgway et al. show that the transcription factor Sox21b represses the size of the posterior lobes of male *Drosophila* genitalia. *Sox21b* expression differs during posterior lobe development between *D. mauritiana* and *D. simulans*, and this gene contributes to the evolution of the size of this morphological novelty between these two species.

Highlights

- *Sox21b* regulates development of posterior lobes, novel *Drosophila* genital organs
- Higher *Sox21b* expression in developing genitalia produces smaller posterior lobes
- *Sox21b* underlies posterior lobe divergence between *D. simulans* and *D. mauritiana*
- The species allele of *Sox21b* causes differences in the duration of copulation

Ridgway et al., 2024, Current Biology *34*, 1114–1121
March 11, 2024
https://doi.org/10.1016/j.cub.2024.01.022

CellPress

CellPress
OPEN ACCESS

Current Biology

Report

Sox21b underlies the rapid diversification of a novel male genital structure between *Drosophila* species

Amber M. Ridgway,[1] Emily J. Hood,[1] Javier Figueras Jimenez,[2] Maria D.S. Nunes,[1,*] and Alistair P. McGregor[2,3,4,*]
[1]Department of Biological and Medical Sciences, Oxford Brookes University, Oxford OX3 0BP, UK
[2]Department of Biosciences, Durham University, Durham DH1 3LE, UK
[3]X (formerly Twitter): @McGregorLab
[4]Lead contact
*Correspondence: msantos-nunes@brookes.ac.uk (M.D.S.N.), alistair.mcgregor@durham.ac.uk (A.P.M.)
https://doi.org/10.1016/j.cub.2024.01.022

SUMMARY

The emergence and diversification of morphological novelties is a major feature of animal evolution.[1–9] However, relatively little is known about the genetic basis of the evolution of novel structures and the mechanisms underlying their diversification. The epandrial posterior lobes of male genitalia are a novelty of particular *Drosophila* species.[10–13] The lobes grasp the female ovipositor and insert between her abdominal tergites and, therefore, are important for copulation and species recognition.[10–12,14–17] The posterior lobes likely evolved from co-option of a Hox-regulated gene network from the posterior spiracles[10] and have since diversified in morphology in the *D. simulans* clade, in particular, over the last 240,000 years, driven by sexual selection.[18–21] The genetic basis of this diversification is polygenic but, to the best of our knowledge, none of the causative genes have been identified.[22–30] Identifying the genes underlying the diversification of these secondary sexual structures is essential to understanding the evolutionary impact on copulation and species recognition. Here, we show that *Sox21b* negatively regulates posterior lobe size. This is consistent with expanded *Sox21b* expression in *D. mauritiana*, which develops smaller posterior lobes than *D. simulans*. We tested this by generating reciprocal hemizygotes and confirmed that changes in *Sox21b* underlie posterior lobe evolution between these species. Furthermore, we found that posterior lobe size differences caused by the species-specific allele of *Sox21b* significantly affect copulation duration. Taken together, our study reveals the genetic basis for the sexual-selection-driven diversification of a novel morphological structure and its functional impact on copulatory behavior.

RESULTS AND DISCUSSION

Identification of transcription factors regulating the development of *Drosophila* male external genitalia

To better understand how the male genitalia develop and have evolved between *D. simulans* and *D. mauritiana* (Figure 1A), we used a candidate gene approach to interrogate RNA sequencing (RNA-seq) data and genomic regions identified by introgression mapping.[27] We focused on genes encoding transcription factors (TFs) because in many previous studies they have been shown to occupy key nodes in gene regulatory networks and contribute to morphological evolution.[31–36]

Our previous analysis of RNA-seq data from the developing male genitalia of *D. simulans*^{w501} and *D. mauritiana* D1 revealed 49 differentially expressed TF-encoding genes,[27] 24 on the chromosome arm 3L, where we have previously generated high-resolution introgression maps of regions contributing to divergence in genital structures between these two species.[26,27,37] We tested the function of these 24 genes during genital development using RNAi knockdown in *D. melanogaster* (Data S1A). RNAi against ten of these genes significantly altered the size and/or bristle count of the posterior lobes, surstyli, and/or cerci compared with parental controls (Figures S1A–S1D; Data S1A). Seven out of the ten affected just one structure, while *Mediator complex subunit 24*, *tonalli*, and *Enhancer of split m3, helix-loop-helix* knockdown affected multiple structures (Figures S1A–S1D). For four of these genes, *CKII-α subunit interactor-1, knirps-like, Mediator complex subunit 10* (all cercus area), and *Sox21b* (posterior lobe area), the effect of the knockdown was consistent with the direction of the difference in their expression and phenotype in *D. mauritiana* D1 compared with *D. simulans*^{w501} (Figures S1A and S1C). However, *Sox21b* is the only one of these genes located within a mapped introgressed region.[27] The higher expression of *Sox21b* in *D. mauritiana*, the species with smaller posterior lobes, and enlargement of this structure upon knockdown in *D. melanogaster*, suggests a previous unknown role for *Sox21b* in repressing posterior lobe size during male genital development.[38,39]

Furthermore, although the posterior lobes of the *D. melanogaster* species subgroup have been shown to have evolved from co-option of the Hox-regulated gene network

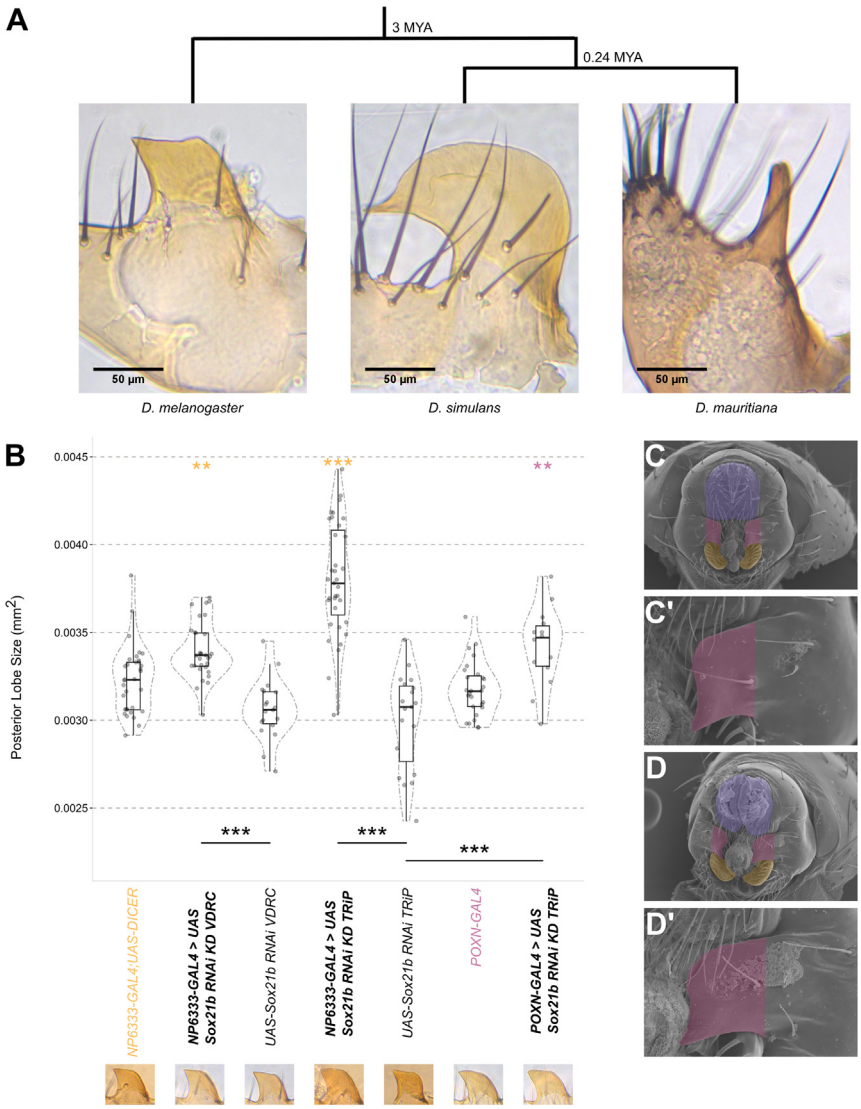


Figure 1. RNAi knockdown of *Sox21b* in *D. melanogaster* increases posterior lobe size

(A) Posterior lobes of *D. melanogaster*, *D. simulans*, and *D. mauritiana*. MYA, million years ago.

(B) RNAi knockdown of *Sox21b* in *D. melanogaster*. Genotypes in bold represent *Sox21b* RNAi knockdown lines. KD, knockdown. UAS lines indicate the parental control lines with only the *UAS-RNAi*. Full genotypes for the lines used can be found in the method details. Asterisks above the knockdown line indicate the levels of statistical significance between the RNAi knockdown and associated parental GAL4 driver: orange for *NP6333-GAL4* and pink for *POXN-GAL4*. The lines below, adjoining the RNAi knockdown to the parental *UAS-RNAi* line, indicate the level of statistical significance between each pair. ***$p < 0.001$ and **$p < 0.01$. Within each violin plot, the following values are represented: the median as the bold horizontal line, the box as the interquartile range, and the range as the vertical line. $n > 13$ for each line.

(C–D′) SEM images of example *NP6333-GAL4* control genitalia (C and C′) and *NP6333-GAL4>UAS-Sox21b RNAi TRiP* genitalia (D and D′). Purple marks the cerci, pink the epandrial posterior lobes, and yellow the surstyli.

See also Figure S1 and Data S1A and S1B.

ancestrally involved in posterior spiracle formation,[10] none of the genes involved in subsequent posterior lobe diversification among these species have yet been identified. Therefore, we further examined the role of *Sox21b* in posterior lobe development and evolution.

Sox21b regulates posterior lobe size

We first carried out additional RNAi knockdowns of *Sox21b* in *D. melanogaster*. We used a second upstream activating sequence (UAS)-RNAi line designed to target a different region of the *Sox21b* mRNA, and an alternative driver line, *POXN-GAL4*, (a posterior-lobe-specific GAL4 driver[40]). All of the RNAi knockdowns of *Sox21b* resulted in an increase in posterior lobe area relative to controls (Figures 1B–1D). In addition, the width of the base of the posterior lobes also significantly increased upon reduction of *Sox21b* expression (Figure S1E). Conversely, *Sox21b* RNAi led to a decrease in the size of the lateral plates, i.e., the structure from which the posterior lobes grow out of (Figure S1F). The reduction of the lateral plate and reciprocal enlargement of the lobe reveals a potential trade-off in the proportion of cells assigned to posterior lobe versus lateral plate fate.

Spatial differences in *Sox21b* expression between species in the posterior lobe primordium

Previous analysis of *Sox21b* in the developing male terminalia of *D. melanogaster* showed expression in the developing posterior lobes and lateral plates during pupal stages.[41] We analyzed the expression of *Sox21b* during larval and pupal stages using *in situ* hybridization chain reaction (HCR). Although we were unable to detect the *Sox21b* expression captured in pupal periphallic structures previously reported,[41] we found that *Sox21b* is expressed earlier in the genital discs of *D. melanogaster*, *D. simulans*, and *D. mauritiana* larvae (Figures 2 and S2).

At 96 h after egg laying (hAEL), *Sox21b* is expressed across the medial and lateral regions of the genital discs of all three species, which encompass the posterior lobe/lateral plate/surstylus primordium (Figures 2A–2C).[42] However, *D. mauritiana* expression was observed across the entirety of this region, whereas *D. simulans* lacks expression in the most lateral parts (Figures 2B and 2C). At 120 hAEL, the medial expression contracts to varying extents among the species (Figures 2A′–2C′). *D. simulans* exhibits the most extreme contraction in expression, with *Sox21b* mostly absent from the posterior lobe primordium (Figure 2B′), whereas in *D. mauritiana*, the broader *Sox21b* expression persists

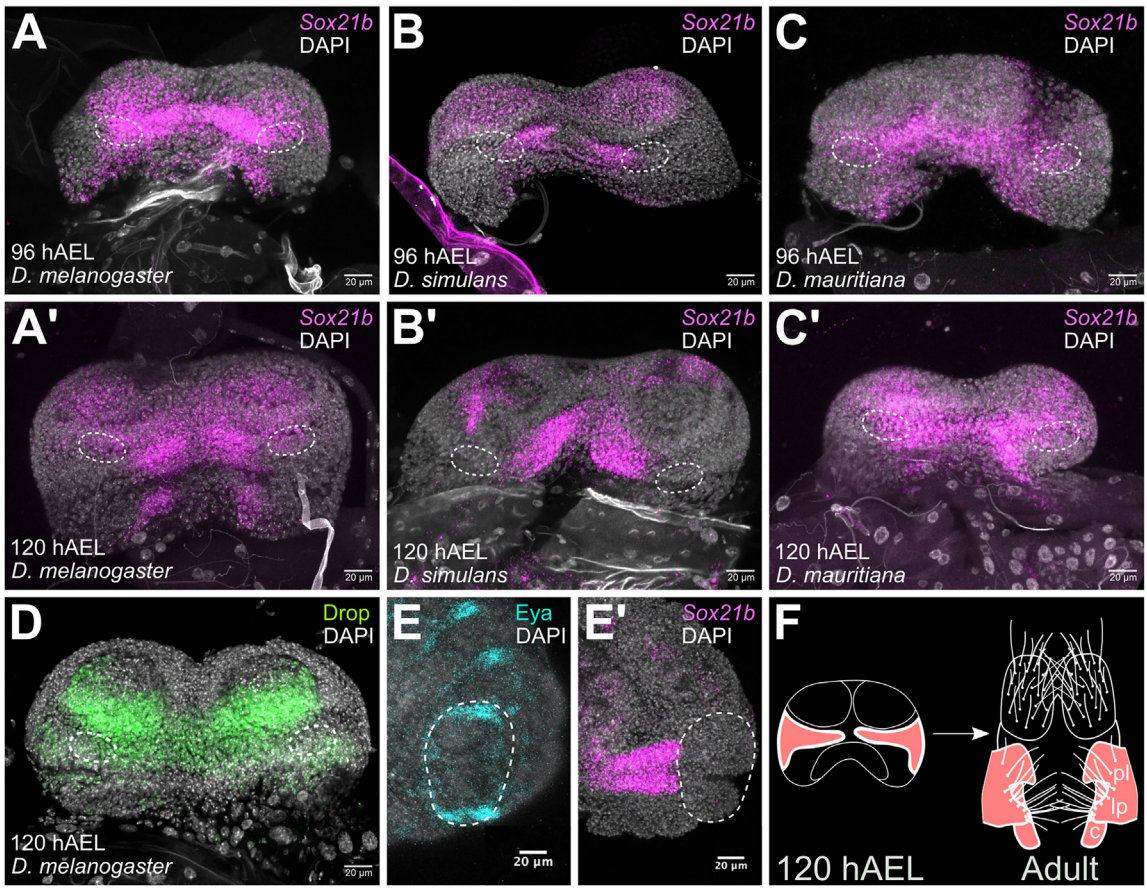


Figure 2. Expression of *Sox21b* in *D. melanogaster*, *D. simulans*, and *D. mauritiana* genital discs

(A–C′) (A–C) *Sox21b in situ* hybridization chain reaction (HCR) of male genital discs at 96 hAEL and (A′–C′) 120 hAEL. (A and A′) *D. melanogaster*^{w1118}. (B and B′) *D. simulans*^{w501}. (C and C′) *D. mauritiana* D1. (A) n = 5, (A′) n = 7, (B) n = 4, (B′) n = 4, (C) n = 4, and (C′) n = 4. Ovals with dashed lines indicate the position of the posterior lobe primordia.

(D) Anti-Drop staining in a *D. melanogaster*^{w1118} male genital disc (n = 8). Positions of the posterior lobe primordia are indicated by ovals with dashed lines.

(E and E′) (E) Anti-Eyes absent (Eya) staining and (E′) *Sox21b* HCR in a *D. melanogaster*^{w1118} embryo at stage 16. Ovals with dashed lines surround the developing posterior spiracles. (E) n = 15 and (E′) n = 8.

(F) Schematic showing lateral plate (lp), posterior lobe (pl), and clasper/surstylus (c) primordia in *D. melanogaster*.

See also Figures S2 and S4.

(Figure 2C′). The expression of *Sox21b* in genital discs is similar to that of Drop (Figure 2D), a known marker of posterior lobe development.[43]

Given our finding that *Sox21b* negatively regulates posterior lobe size, these data suggest that the higher and more expansive expression—detected by RNA-seq and *in situ* hybridization, respectively—in *D. mauritiana* compared with *D. simulans* contributes to the evolutionary difference in posterior lobe size between these two species.

Sox21b is not expressed in the developing posterior spiracles

We next explored whether *Sox21b* could have also been co-opted with the Hox-regulated network during posterior lobe evolution by assaying whether this gene is also expressed in the posterior spiracles, like *eyes absent* (Figures 2E and 2E′).[10] However, we did not observe *Sox21b* expression in the developing posterior spiracles in *D. melanogaster* (Figure 2E′). This suggests that, while *Sox21b* can modulate the expression of this network to regulate posterior lobe size, expression of this gene was likely not co-opted from the posterior spiracles.

Highly conserved coding sequences of *Sox21b* between *D. mauritiana* and *D. simulans*

To further compare *Sox21b* between *D. simulans* and *D. mauritiana*, we analyzed the coding sequence differences in this gene between the two species. We found 24 nucleotide differences in the coding sequence of *Sox21b* between *D. simulans*^{w501} and *D. mauritiana* D1, four of which are non-synonymous, but we found no insertions or deletions. Two of the nonsynonymous changes are located in exon 1 and are derived in *D. simulans* (Data S1F). The other two changes are located in exon 6 and are derived in *D. mauritiana*. None of these changes affect the DNA binding high-mobility group (HMG) box domain. Only the A535V substitution

Table 1. Nonsynonymous changes in the coding sequence of *Sox21b* between *D. simulans* and *D. mauritiana*

		Codon/amino acid			nsSNV frequency[c]			
					ReSeq		Popoolation DB	
Nsyn[a]	*Dsim* coordinate[b]	*Dmel*	*Dsim*^{w501}	*DmauD1*	*Dsim*	*Dmau*	*Dsim*	*Dmau*
A151T^{SP}	13791381	GCG/A	–	GCG/A	5/10	8/11	9/18	11/16
		–	ACG/T	–	5/10	3/11	9/18	5/16
A186S^{S}	NP	GCT/A	–	GCT/A	9/10	11/11	–	
		–	TCT/S	–	1/10	0/11		
E502D^{SP}	13779108	GAG/E	GAG/E	–	9/10	3/10	8/15	9/14
		–	–	GAT/D	1/20	7/10	7/15	5/14
A535V^{F}	NP	GCG/A	GCG/A	–	10/10	0/10	–	
		–	–	GTG/V	0/10	10/10		

NP: variant not present in the Popoolation dataset. nsSNV Frequency obtained from two datasets (1) ReSeq: a dataset made by resequencing ten strains of each species, (2) Popoolation DB[49,50]: Pool-seq data from 107 strains of *D. mauritiana* and 50 strains of sub-Saharan *D. simulans*. Amino acid differences denoted as: SP, shared polymorphism; S, singleton in *Dsim*^{w501}; F, fixed difference. See also Data S1F.

[a]Nonsynonymous.

[b]PopoolationDB.[49]

[c]Nonsynonymous single nucleotide variation.

represents a fixed difference between *D. simulans* and *D. mauritiana* (Table 1). The variants underlying the A151T and E502D change between *D. simulans*^{w501} and *D. mauritiana* D1 are present in populations of both species and the A186S variant is a singleton in *D. simulans*^{w501} (Table 1). All substitutions are fairly conservative[44–46] and unlikely to have a large effect on protein structure and function,[47,48] especially given the pleiotropic roles of Sox21b in these flies.[38,39] Although we cannot completely rule out the contribution of the amino acid substitutions in Sox21b between the two species to their posterior lobe size differences, given the striking spatial differences in the expression of this gene in the genital discs of *D. mauritiana* and *D. simulans*, we hypothesize that regulatory evolution in *Sox21b* is more likely to be responsible.

Sox21b contributes to the evolution of posterior lobe size between *D. simulans* and *D. mauritiana*

Because *Sox21b* regulates posterior lobe development in *D. melanogaster*, is located in a genomic region[27] that contributes to differences in posterior lobe size between *D. mauritiana* and *D. simulans*, and is expressed differently between these two species, we then directly tested whether this gene has contributed to the evolution of their posterior lobe size difference. To do this, we carried out a reciprocal hemizygosity test.[37,51–53] We used CRISPR-Cas9 to direct the insertion of 3XP3-DsRed fluorophore into exon 1 of *Sox21b* in *D. simulans*^{w501} and an introgression line (*IL108*) carrying 2.8 Mb of *D. mauritiana* DNA (3L: 12,277,961–15,075,323), spanning the previously mapped region containing this gene (P5, 1.138 Mb) (3L: 13,393,862–14,532,063)[27] in an otherwise *D. simulans*^{w501} genome (Figure S3A). This successfully disrupted the reading frame in the *Sox21b* locus from both species (Figure S3A). Interestingly, *D. simulans Sox21b* mutants were homozygous viable, and these flies had significantly larger posterior lobes than controls. This result corroborates the *Sox21b* RNAi knockdown results in *D. melanogaster* and confirms that this TF negatively regulates posterior lobe size in *D. simulans* (Figures 1B and S3B; Data S1B and S1C).

Reciprocal hemizygotes were then generated by crossing the two sets of independent *Sox21b* mutant lines to generate flies that were genetically identical except that they had either a working *D. mauritiana Sox21b* allele (*IL108/Dsim*$^{Sox21b1.1/1.2}$) or a working *D. simulans Sox21b* allele (*Dsim/IL108*$^{Sox21b1.1/1.2}$) (Figure 3A). Remarkably, posterior lobe size was significantly different, depending upon the species-origin of the working *Sox21b* allele and consistent with the direction of the species difference (Figure 3A). Male flies with the *D. mauritiana Sox21b* allele had significantly smaller lobes with narrower bases compared with flies with a working *D. simulans Sox21b* allele (Cohen's effect size = −1.172; Figures 3A and S3C). Surstylus bristle count, cercus bristle count, and cercus area did not differ significantly between the reciprocal hemizygotes (Data S1B and S1C). This difference of 5.5% in the size of the posterior lobes between the reciprocal hemizygotes is consistent with the 9% effect of the introgression containing *Sox21b*.[27] These results show that variation in *Sox21b* has contributed to the evolution of posterior lobe size between *D. mauritiana* and *D. simulans*.

The shape of the posterior lobe is altered according to the species origin of *Sox21b*

To test whether variation in *Sox21b* contributes to the evolution of posterior lobe shape between *D. simulans* and *D. mauritiana*, we used elliptical Fourier analysis (EFA) to summarize the morphometric changes in shape between the reciprocal hemizygotes to identify variation originating from species-specific alleles of *Sox21b* (Figures 3B, S3D, and S3E; Data S1D). The two principal components (PCs) summarizing the highest proportion of shape variation, PC1 and PC2, contributing 27.1% and 12.5%, respectively, were compared to assess the relative distribution in shape differences captured by species-specific alleles of *Sox21b* (Figures S3E and S3E′). PC1 captured shape alterations from the artificial baseline to the beginning of the beak extension. The *D. mauritiana Sox21b* allele reduced the height of the "neck" leading up to the "beak" extension, in comparison to the *D. simulans* working allele (Figure 3B). When

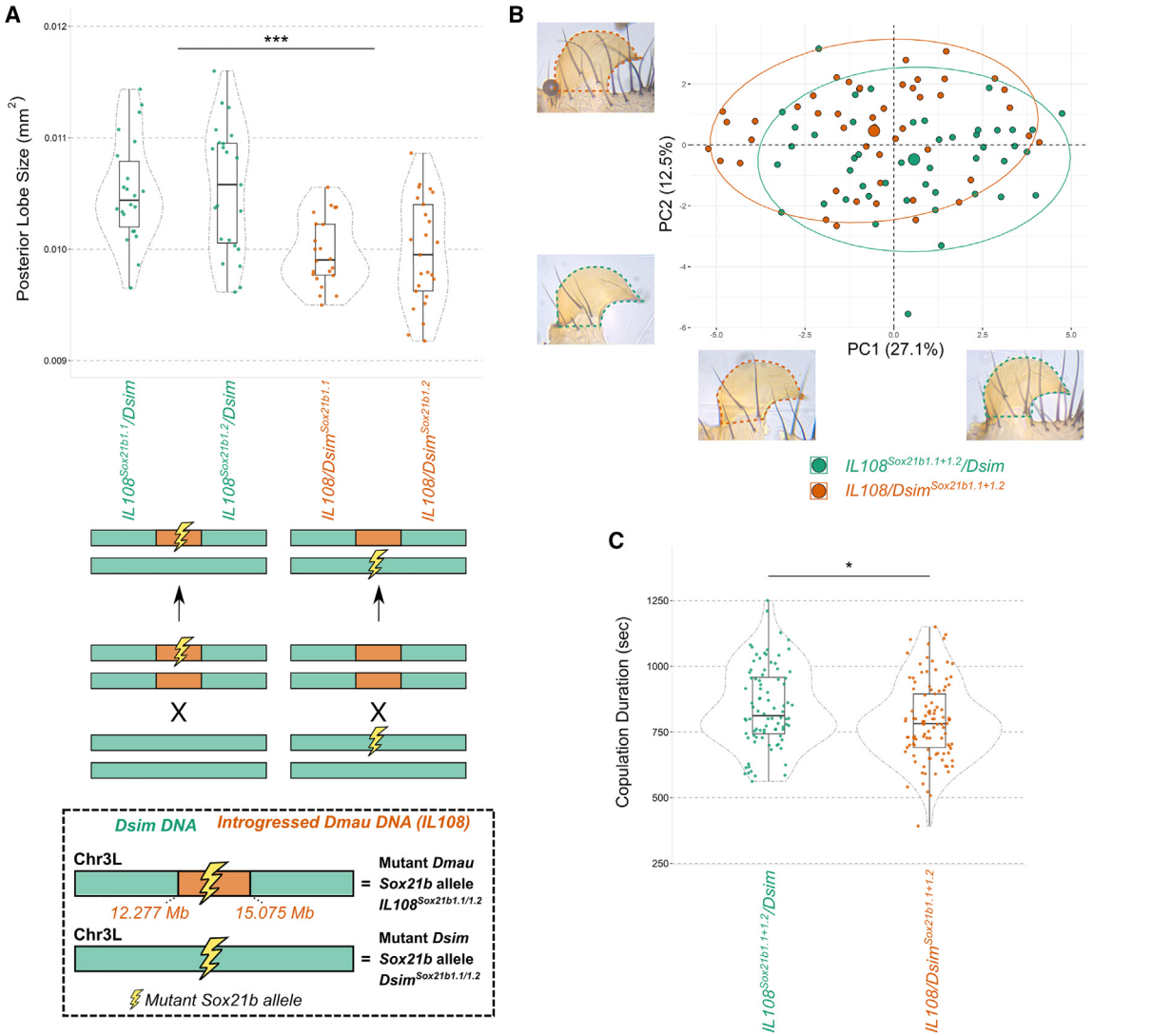


Figure 3. Effects of species-specific *Sox21b* alleles on posterior lobe size and shape

(A) Posterior lobes are smaller in size when only the *D. mauritiana* allele of *Sox21b* is working (*IL108/Dsim*$^{Sox21b1.1}$ and *IL108/Dsim*$^{Sox21b1.2}$) (n > 22). The lower schematic illustrates the generation of reciprocal hemizygotes, containing either a *D. simulans* working allele or *D. mauritiana* working allele of *Sox21b*. The lightning bolt represents the disrupted *Sox21b* locus, the shaded rectangle represents the *D. mauritiana* introgressed region (3L: 12,277,961–15,075,323).

(B) PC1 versus PC2, whereby both statistical identifiers summarize 37.6% of total shape variation between the two reciprocal hemizygotes. Outlines of lobes show the minimal and maximal data points for each principal component.

(C) The *D. mauritiana* working *Sox21b* allele significantly decreased copulation duration compared with the working *D. simulans Sox21b* allele when reciprocal hemizygote males were paired with *D. simulans*^{w501} females (n > 87 for each reciprocal hemizygote).

See also Figures S3 and S4 and Data S1B–S1E.

assessing the variation captured by PC2, a notable shape alteration in the beak region of the lobe (extending from the main body of the lobe, parallel to the artificial baseline) was identified (Figure 3B). Therefore, the EFA of the posterior lobes between the *Sox21b* reciprocal hemizygotes revealed shape variation affecting the beak extension and the neck, consistent with the direction of the species difference.

The behavioral consequences of *Sox21b* species-specific alleles

We next sought to understand whether the species-specific alleles of *Sox21b* that contribute to evolutionary differences in posterior lobe morphology also cause detectable differences in copulation, by crossing reciprocal hemizygote males to *D. simulans*^{w501} females. In contrast to a previous study that used laser microdissection to alter the morphology of the posterior lobes in *D. simulans*,[16] we found no difference in mating frequency between reciprocal hemizygote males (Data S1C and S1E). We also found no difference in copulation latency between reciprocal hemizygote males (Figure S4A; Data S1C and S1E) but, indeed, to the best of our knowledge, copulation latency differences do not segregate in these two species, although a previous study using introgression lines between *D. mauritiana* and *D. sechellia* found that posterior lobe morphology did affect copulation latency.[14] Interestingly,

however, we did observe that males carrying a working *D. mauritiana Sox21b* allele engaged in significantly shorter copulations than those carrying a working *D. simulans Sox21b* allele, which may result in a difference in sperm transfer time (Figure 3C; Data S1C). This result is consistent with the previous study of *D. mauritiana* and *D. sechellia* and, more importantly, a species-specific difference in copulation duration, with shorter copulations in *D. mauritiana* than in *D. simulans*,[54–56] although we cannot exclude that this could be a general effect of smaller lobes rather than representing the true species difference.

Conclusions

We have shown that inter-specific allelic variation in *Sox21b* contributes to the diversification of an evolutionary novelty, the posterior lobes, and that this affects copulatory behavior, suggesting that this gene has been subject to sexual selection to help shape lobes with different morphologies. Genes have been identified that contribute to the evolution of other sexual traits,[37,57,58] but *Sox21b* is the first for the diversification of the posterior lobes, to the best of our knowledge. As it appears that *Sox21b* regulates the early stages of posterior lobe development, one possible mechanism for this occurrence is that this gene may restrict the allocation of lateral plate cells to posterior lobe fate and thereby influence the size and shape of these structures. Our results strongly suggest that a difference in *Sox21b* expression between *D. mauritiana* and *D. simulans* underlies the role of this gene in the evolution of posterior lobe morphology between these two species. However, we cannot rule out the possibility that coding changes may also be involved because we did not detect a consistent difference in *Sox21b* expression between the reciprocal hemizygotes (Figures S4B and S4C). This is likely because the *Sox21b* expression difference in the reciprocal hemizygotes is too subtle to detect, and the full difference in the expression of this gene, and difference in posterior lobe morphology between *D. mauritiana* and *D. simulans*, probably involves the evolution of the expression of TFs that regulate *Sox21b*. Therefore, it will be important to identify the other genes that contribute to posterior lobe diversity between *D. mauritiana* and *D. simulans*, and other species, to understand more fully how the gene regulatory network for posterior lobes has evolved. Several other genes have already been found that regulate posterior lobe development,[10,28,41] and a crucial role has been identified for the apical extracellular matrix (ECM) in cell extension from the lateral plate.[59] Therefore, it will be interesting to determine how *Sox21b* is integrated into this network.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - RNAi knockdown of differentially expressed transcription factors in *D. melanogaster*
  - Phenotyping of RNAi knockdown flies
  - Generation of Sox21b-DsRed
  - Generation of reciprocal hemizygotes
  - Scanning electron microscopy
  - Posterior lobe shape analysis
  - *In situ* HCR of *Sox21b* in genital discs, embryos and pupal terminalia
  - Immunohistochemistry in genital discs and embryos
  - *Sox21b* coding sequence analysis
  - Behaviour assays using *Sox21b* reciprocal hemizygotes
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2024.01.022.

ACKNOWLEDGMENTS

This study was funded in part by BBSRC (BB/X006689/1) and NERC grants (NE/M001040/1) to A.P.M. and M.D.S.N., a BBSRC DTP studentship (BB/M011224/1) to A.M.R., and an Oxford Brookes University Nigel Groome studentship to E.J.H. We thank the Bloomington *Drosophila* Stock Center and Vienna *Drosophila* RNAi Center for fly lines, as well as Molecular Instruments for the custom-made *Sox21b* HCR probe. We also thank members of the Rebeiz lab for their thoughtful discussions.

AUTHOR CONTRIBUTIONS

The project was conceived by A.P.M. and M.D.S.N. Experiments were carried out by A.M.R. and E.J.H. Data were analyzed by all authors. A.M.R., A.P.M., and M.D.S.N. wrote the paper, assisted by E.J.H. and J.F.J.

DECLARATION OF INTERESTS

The authors declare they have no conflicting interests.

Received: August 21, 2023
Revised: December 2, 2023
Accepted: January 8, 2024
Published: February 2, 2024

REFERENCES

1. York, J.R., and McCauley, D.W. (2020). The origin and evolution of vertebrate neural crest cells. Open Biol. *10*, 190285.
2. Tomoyasu, Y. (2021). What crustaceans can tell us about the evolution of insect wings and other morphologically novel structures. Curr. Opin. Genet. Dev. *69*, 48–55.
3. Moczek, A.P. (2009). On the origins of novelty and diversity in development and evolution: a case study on beetle horns. Cold Spring Harbor Symp. Quant. Biol. *74*, 289–296.
4. Kijimoto, T., Pespeni, M., Beckers, O., and Moczek, A.P. (2013). Beetle horns and horned beetles: emerging models in developmental evolution and ecology. Wiley Interdiscip. Rev. Dev. Biol. *2*, 405–418.
5. Rebeiz, M., and Tsiantis, M. (2017). Enhancer evolution and the origins of morphological novelty. Curr. Opin. Genet. Dev. *45*, 115–123.
6. Shubin, N., Tabin, C., and Carroll, S. (2009). Deep homology and the origins of evolutionary novelty. Nature *457*, 818–823.

7. Bruce, H.S., and Patel, N.H. (2022). The Daphnia carapace and other novel structures evolved via the cryptic persistence of serial homologs. Curr. Biol. *32*, 3792–3799.e3.
8. Colizzi, E.S., Hogeweg, P., and Vroomans, R.M.A. (2022). Modelling the evolution of novelty: a review. Essays Biochem. *66*, 727–735.
9. Muller, G.B., and Wagner, G.P. (1991). Novelty in Evolution: Restructuring the Concept. Annu. Rev. Ecol. Syst. *22*, 229–256.
10. Glassford, W.J., Johnson, W.C., Dall, N.R., Smith, S.J., Liu, Y., Boll, W., Noll, M., and Rebeiz, M. (2015). Co-option of an Ancestral Hox-Regulated Network Underlies a Recently Evolved Morphological Novelty. Dev. Cell *34*, 520–531.
11. Jagadeeshan, S., and Singh, R.S. (2006). A time-sequence functional analysis of mating behavior and genital coupling in *Drosophila*: role of cryptic female choice and male sex-drive in the evolution of male genitalia. J. Evol. Biol. *19*, 1058–1070.
12. Yassin, A., and Orgogozo, V. (2013). Coevolution between Male and Female Genitalia in the *Drosophila melanogaster* Species Subgroup. PLoS One *8*, e57158.
13. Kopp, A., and True, J.R. (2002). Evolution of male sexual characters in the Oriental *Drosophila melanogaster* species group. Evol. Dev. *4*, 278–291.
14. Frazee, S.R., Harper, A.R., Afkhami, M., Wood, M.L., McCrory, J.C., and Masly, J.P. (2021). Interspecific introgression reveals a role of male genital morphology during the evolution of reproductive isolation in *Drosophila*. Evolution *75*, 989–1002.
15. Frazee, S.R., and Masly, J.P. (2015). Multiple sexual selection pressures drive the rapid evolution of complex morphology in a male secondary genital structure. Ecol. Evol. *5*, 4437–4450.
16. LeVasseur-Viens, H., Polak, M., and Moehring, A.J. (2015). No evidence for external genital morphology affecting cryptic female choice and reproductive isolation in *Drosophila*. Evolution *69*, 1797–1807.
17. Robertson, H.M. (1988). Mating asymmetries and phylogeny in the *Drosophila melanogaster* species complex. Pac. Sci. *42*, 72–80.
18. McDermott, S.R., and Kliman, R.M. (2008). Estimation of Isolation Times of the Island Species in the *Drosophila simulans* Complex from Multilocus DNA Sequence Data. PLOS One *3*, e2442.
19. Coyne, J.A. (1983). Genetic Basis of Differences in Genital Morphology Among Three Sibling Species of *Drosophila*. Evolution (NY) *37*, 1101–1118.
20. House, C.M., Lewis, Z., Hodgson, D.J., Wedell, N., Sharma, M.D., Hunt, J., and Hosken, D.J. (2013). Sexual and Natural Selection Both Influence Male Genital Evolution. PLoS One *8*, e63807.
21. House, C.M., Lewis, Z., Sharma, M.D., Hodgson, D.J., Hunt, J., Wedell, N., and Hosken, D.J. (2021). Sexual selection on the genital lobes of male *Drosophila* simulans. Evolution *75*, 501–514.
22. Liu, J., Mercer, J.M., Stam, L.F., Gibson, G.C., Zeng, Z.-B., and Laurie, C.C. (1996). Genetic Analysis of a Morphological Shape Difference in the Male Genitalia of *Drosophila simulans* and *D. mauritiana*. Genetics *142*, 1129–1145.
23. Zeng, Z.B., Liu, J., Stam, L.F., Kao, C.H., Mercer, J.M., and Laurie, C.C. (2000). Genetic architecture of a morphological shape difference between two *Drosophila* species. Genetics *154*, 299–310.
24. LeVasseur-Viens, H., and Moehring, A.J. (2014). Individual Genetic Contributions to Genital Shape Variation between *Drosophila simulans* and *D. mauritiana*. Int. J. Evol. Biol. *2014*, 808247.
25. True, J.R., Liu, J., Stam, L.F., Zeng, Z.-B., and Laurie, C.C. (1997). Quantitative Genetic Analysis of Divergence in Male Secondary Sexual Traits Between *Drosophila simulans* and *Drosophila mauritiana*. Evolution *51*, 816–832.
26. Tanaka, K.M., Hopfen, C., Herbert, M.R., Schlötterer, C., Stern, D.L., Masly, J.P., McGregor, A.P., and Nunes, M.D.S. (2015). Genetic Architecture and Functional Characterization of Genes Underlying the Rapid Diversification of Male External Genitalia Between *Drosophila simulans* and *Drosophila mauritiana*. Genetics *200*, 357–369.
27. Hagen, J.F.D., Mendes, C.C., Booth, S.R., Figueras Jimenez, J., Tanaka, K.M., Franke, F.A., Baudouin-Gonzalez, L., Ridgway, A.M., Arif, S., Nunes, M.D.S., et al. (2021). Unraveling the Genetic Basis for the Rapid Diversification of Male Genitalia between *Drosophila* Species. Mol. Biol. Evol. *38*, 437–448.
28. Hackett, J.L., Wang, X., Smith, B.R., and Macdonald, S.J. (2016). Mapping QTL Contributing to Variation in Posterior Lobe Morphology between Strains of *Drosophila melanogaster*. PLoS One *11*, e0162573.
29. Laurie, C.C., True, J.R., Liu, J., and Mercer, J.M. (1997). An Introgression Analysis of Quantitative Trait Loci That Contribute to a Morphological Difference Between *Drosophila simulans* and *D. mauritiana*. Genetics *145*, 339–348.
30. Masly, J.P., Dalton, J.E., Srivastava, S., Chen, L., and Arbeitman, M.N. (2011). The Genetic Basis of Rapidly Evolving Male Genital Morphology in *Drosophila*. Genetics *189*, 357–374.
31. Zhang, L., Mazo-Vargas, A., and Reed, R.D. (2017). Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. Proc. Natl. Acad. Sci. USA *114*, 10707–10712.
32. Williams, T.M., Selegue, J.E., Werner, T., Gompel, N., Kopp, A., and Carroll, S.B. (2008). The Regulation and Evolution of a Genetic Switch Controlling Sexually Dimorphic Traits in *Drosophila*. Cell *134*, 610–623.
33. Stern, D.L. (2011). Evolution, Development, and the Predictable Genome (Wiley-Blackwell).
34. Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., et al. (2010). Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a Pitx1 Enhancer. Science *327*, 302–305.
35. Stern, D.L., and Frankel, N. (2013). The structure and evolution of cis-regulatory regions: the *shavenbaby* story. Philos. Trans. R. Soc. Lond. B Biol. Sci. *368*, 20130028.
36. Carroll, S.B., Grenier, J.K., and Weatherbee, S.D. (2009). From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design, Second Edition (Wiley-Blackwell).
37. Hagen, J.F.D., Mendes, C.C., Blogg, A., Payne, A., Tanaka, K.M., Gaspar, P., Figueras Jimenez, J., Kittelmann, M., McGregor, A.P., and Nunes, M.D.S. (2019). *tartan* underlies the evolution of *Drosophila* male genital morphology. Proc. Natl. Acad. Sci. USA *116*, 19025–19030.
38. Akhund-Zade, J., Lall, S., Gajda, E., Yoon, D., Ayroles, J.F., and de Bivort, B.L. (2021). Genetic basis of offspring number-body weight tradeoff in *Drosophila melanogaster*. G3 (Bethesda) *11*, jkab129.
39. McKimmie, C., Woerfel, G., and Russell, S. (2005). Conserved genomic organisation of Group B Sox genes in insects. BMC Genet. *6*, 26.
40. Boll, W., and Noll, M. (2002). The *Drosophila* Pox neuro gene: control of male courtship behavior and fertility as revealed by a complete dissection of all enhancers. Development *129*, 5667–5681.
41. Vincent, B.J., Rice, G.R., Wong, G.M., Glassford, W.J., Downs, K.I., Shastay, J.L., Charles-Obi, K., Natarajan, M., Gogol, M., Zeitlinger, J., et al. (2019). An Atlas of Transcription Factors Expressed in Male Pupal Terminalia of *Drosophila melanogaster*. G3 (Bethesda) *9*, 3961–3972.
42. Keisman, E.L., and Baker, B.S. (2001). The *Drosophila* sex determination hierarchy modulates wingless and decapentaplegic signaling to deploy dachshund sex-specifically in the genital imaginal disc. Development *128*, 1643–1656.
43. 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.
44. Müller, T., and Vingron, M. (2000). Modeling amino acid replacement. J. Comput. Biol. *7*, 761–776.
45. Henikoff, S., and Henikoff, J.G. (1992). Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA *89*, 10915–10919.
46. Mount, D.W. (2008). Using BLOSUM in Sequence Alignments. Cold Spring Harb. Protoc. *2008*, pdb.top.39.
47. Barnes, M.R., and Gray, I.C. (2003). Bioinformatics for Geneticists (Wiley).

48. Jonson, P.H., and Petersen, S.B. (2001). A critical view on conservative mutations. Protein Eng. *14*, 397–402.

49. Pandey, R.V., Kofler, R., Orozco-terWengel, P., Nolte, V., and Schlötterer, C. (2011). PoPoolation DB: a user-friendly web-based database for the retrieval of natural polymorphisms in *Drosophila*. BMC Genet. *12*, 27.

50. Nolte, V., Pandey, R.V., Kofler, R., and Schlötterer, C. (2013). Genome-wide patterns of natural variation reveal strong selective sweeps and ongoing genomic conflict in *Drosophila mauritiana*. Genome Res. *23*, 99–110.

51. Stern, D.L. (2014). Identification of loci that cause phenotypic variation in diverse species with the reciprocal hemizygosity test. Trends Genet. *30*, 547–554.

52. Lamb, A.M., Wang, Z., Simmer, P., Chung, H., and Wittkopp, P.J. (2020). ebony Affects Pigmentation Divergence and Cuticular Hydrocarbons in *Drosophila americana* and *D. novamexicana*. Front. Ecol. Evol. *8*, 184.

53. Ding, Y., Berrocal, A., Morita, T., Longden, K.D., and Stern, D.L. (2016). Natural courtship song variation caused by an intronic retroelement in an ion channel gene. Nature *536*, 329–332.

54. Price, C.S.C., Kim, C.H., Gronlund, C.J., and Coyne, J.A. (2001). Cryptic reproductive isolation in the *Drosophila simulans* species complex. Evolution *55*, 81–92.

55. Cobb, M., Burnet, B., and Connolly, K. (1988). Sexual isolation and courtship behavior in *Drosophila simulans*, *D. mauritiana*, and their interspecific hybrids. Behav. Genet. *18*, 211–225.

56. Coyne, J.A. (1993). The genetics of an isolating mechanism between two sibling species of *Drosophila*. Evolution *47*, 778–788.

57. Nagy, O., Nuez, I., Savisaar, R., Peluffo, A.E., Yassin, A., Lang, M., Stern, D.L., Matute, D.R., David, J.R., and Courtier-Orgogozo, V. (2018). Correlated Evolution of Two Copulatory Organs via a Single cis-Regulatory Nucleotide Change. Curr. Biol. *28*, 3450–3457.e13.

58. Gao, J.J., Barmina, O., Thompson, A., Kim, B.Y., Suvorov, A., Tanaka, K., Watabe, H., Toda, M.J., Chen, J.-M., Katoh, T.K., et al. (2022). Secondary reversion to sexual monomorphism associated with tissue-specific loss of doublesex expression. Evolution *76*, 2089–2104.

59. Smith, S.J., Davidson, L.A., and Rebeiz, M. (2020). Evolutionary expansion of apical extracellular matrix is required for the elongation of cells in a novel structure. eLife *9*, e55965.

60. Gratz, S.J., Ukken, F.P., Rubinstein, C.D., Thiede, G., Donohue, L.K., Cummings, A.M., and O'Connor-Giles, K.M. (2014). Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in *Drosophila*. Genetics *196*, 961–971.

61. R Core Team (2020). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

62. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

63. Chan, I.Z.W., Stevens, M., and Todd, P.A. (2019). PAT-GEOM: A Software Package for the Analysis of Animal Patterns. Methods Ecol. Evol. *10*, 591–600.

64. Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous Inference in General Parametric Models. Biom. J. *50*, 346–363.

65. Kassambara, A., and Mundt, F. (2020). Extract and visualize the results of multivariate data analyses. R Package Version 1.0.7. https://scirp.org/reference/referencespapers?referenceid=3067217.

66. 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.

67. Stieper, B.C., Kupershtok, M., Driscoll, M.V., and Shingleton, A.W. (2008). Imaginal discs regulate developmental timing in *Drosophila melanogaster*. Dev. Biol. *321*, 18–26.

68. Blake, A.J., Finger, D.S., Hardy, V.L., and Ables, E.T. (2017). RNAi-Based Techniques for the Analysis of Gene Function in *Drosophila* Germline Stem Cells. In RNAi and Small Regulatory RNAs in Stem Cells: Methods and Protocols, B. Zhang, ed. (Springer), pp. 161–184.

69. 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.

70. Gratz, S.J., Rubinstein, C.D., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2015). CRISPR-Cas9 Genome Editing in *Drosophila*. Curr. Protoc. Mol. Biol. *111*, 31.2.1–31.2.20.

71. Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. *16*, 1215.

72. Caple, J., Byrd, J., and Stephan, C.N. (2017). Elliptical Fourier analysis: fundamentals, applications, and value for forensic anthropology. Int. J. Legal Med. *131*, 1675–1690.

73. Younger, M.A., Herre, M., Goldman, O.V., Lu, T.-C., Caballero-Vidal, G., Qi, Y., Gilbert, Z.N., Gong, Z., Morita, T., Rahiel, S., et al. (2022). Non-Canonical Odor Coding in the Mosquito. https://doi.org/10.1101/2020.11.07.368720.

74. Choi, H.M.T., Schwarzkopf, M., Fornace, M.E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., and Pierce, N.A. (2018). Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. Development *145*, dev165753.

75. Schwarzkopf, M., Liu, M.C., Schulte, S.J., Ives, R., Husain, N., Choi, H.M.T., and Pierce, N.A. (2021). Hybridization chain reaction enables a unified approach to multiplexed, quantitative, high-resolution immunohistochemistry and in situ hybridization. Development *148*, dev199847.

76. Stern, D.L., Crocker, J., Ding, Y., Frankel, N., Kappes, G., Kim, E., Kuzmickas, R., Lemire, A., Mast, J.D., and Picard, S. (2017). Genetic and transgenic reagents for *Drosophila simulans*, *D. mauritiana*, *D. yakuba*, *D. santomea*, and *D. virilis*. G3 (Bethesda) *7*, 1339–1347.

77. Krüger, A.P., Vieira, J.G.A., Scheunemann, T., Nava, D.E., and Garcia, F.R.M. (2021). Effects of temperature and relative humidity on mating and survival of sterile *Drosophila suzukii*. J. Appl. Entomol. *145*, 789–799.

78. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. B Methodol. *57*, 289–300.

79. Rosenthal, R., Cooper, H., and Hedges, L.V. (1994). Parametric measures of effect size. In The Hand-book of Research Synthesis (Russell Sage Foundation), pp. 231–244.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Drop (rabbit)	C. Doe	N/A
Anti-rabbit AlexaFluor 488 (goat)	ThermoFisher Scientific	CAT# A-11008; RRID: AB_143165
Anti-Eya (mouse)	DSHB	CAT# eya10h6; RRID: AB_528232
Anti-mouse 647 (donkey)	Thermo Fisher Scientific	CAT# A32787; RRID: AB_2762830
Bacterial and virus strains		
DH5-alpha competent *E. coli*	NEB	CAT# C2987H
Chemicals, peptides, and recombinant proteins		
Hoyer's Liquid	entomopraxis	CAT# A901B
Q5® High-Fidelity DNA Polymerase	NEB	CAT# M0491
Normal Goat Serum	Life Technologies	CAT# PCN5000
Formaldehyde	Merck	CAT# F8775-4X25ML
EcoRI-HF	NEB	CAT# R3101
BbsI-HF®	NEB	CAT# R3539
SapI	NEB	CAT# R0569
AarI	ThermoFisher Scientific	CAT# ER1581
Hydromount Histology Mounting Media	Scientific Laboratory Supplies	CAT# NAT1324
Triton X-100	Sigma Aldrich	CAT# X100
TWEEN® 20	Sigma Aldrich	CAT# P9416
Dextran Sulfate	Sigma Aldrich	CAT# D6001
Formamide	ThermoFisher Scientific	CAT# AM9342
Heparin	Sigma Aldrich	CAT# H3393
Critical commercial assays		
DAPI	Merck	CAT# 10236276001
QIAGEN Plasmid Mini/Midi Kit	QIAGEN	CAT# 12123
Ampicillin Sodium Salt (Crystalline Powder)	ThermoFisher Scientific	CAT# BP1760-25
HCR Custom *Sox21b* Synthetic DNA Oligonucleotide and B3 647 Hairpins	Molecular Instruments	N/A
Experimental models: Organisms/strains		
D. melanogaster yw; NP6333-GAL4, UAS-DICER	(Without UAS-DICER) Kyoto DGGR	RRID: DGGR_113920
D. melanogaster w; Poxn-GAL4$^{4.14}$/TM6B	BDSC	RRID: BDSC_66685
D. simulans nos>cas9;; IL108 Sox21b-DsRed (IL108$^{Sox21b1.1/1.2}$)	This Paper	N/A
D. simulans nos>cas9;; Sox21b-DsRed (D.sim$^{Sox21b1.1/1.2}$)	This Paper	N/A
D. simulans^{w501}	National Drosophila Species Stock Center	CAT# 14021-0251.195
D. mauritiana: D1	True et al.[25]; Tanaka et al.[26]; Hagen et al.[27]	N/A
D. melanogaster^{w1118}	BDSC	RRID: BDSC_3605
D. simulans^{w501};; IL108 (164.2.16.4 D. mauritiana^{w-12} introgression)	Hagen et al.[27]	N/A
D. simulans^{w501} nos>cas9	Hagen et al.[27]	N/A
D. simulans^{w501} nos>cas9;; IL108	This Paper	N/A
D. melanogaster; UAS-Sox21b RNAi ^{60120}	BDSC	RRID: BDSC_60120

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
D. melanogaster UAS-Sox21b RNAi [41098]	Vienna Biocentre	CAT# 41098
D. melanogaster; UAS-Ssb-c31a RNAi [65060]	BDSC	RRID: BDSC_65060
D. melanogaster; UAS-Ssb-c31a RNAi [47383]	Vienna Biocentre	CAT# 47383
D. melanogaster;; UAS-E(spl)m3-HLH RNAi [55302]	BDSC	RRID: BDSC_55302
D. melanogaster;; UAS-Myb RNAi [35053]	BDSC	RRID: BDSC_35053
D. melanogaster;; UAS-Myb RNAi [37710]	Vienna Biocentre	CAT# 37710
D. melanogaster; UAS-nom RNAi [67295]	BDSC	RRID: BDSC_67295
D. melanogaster; UAS-odj RNAi [62184]	BDSC	RRID: BDSC_62184
D. melanogaster;; UAS-pnr RNAi [33697]	BDSC	RRID: BDSC_33697
D. melanogaster;; UAS-pnr RNAi [6224]	Vienna Biocentre	CAT# 6224
D. melanogaster; UAS-CG6276 RNAi [60485]	BDSC	RRID: BDSC_60485
D. melanogaster; UAS-CG6276 RNAi [30142]	Vienna Biocentre	CAT# 30142
D. melanogaster;; UAS-ich RNAi [44046]	BDSC	RRID: BDSC_44046
D. melanogaster; UAS-CG11966 RNAi [20127]	Vienna Biocentre	CAT# 20127
D. melanogaster;; UAS-E(var)3-9 RNAi [31948]	BDSC	RRID: BDSC_31948
D. melanogaster;; UAS-CKIIalpha-il RNAi [10714]	Vienna Biocentre	CAT# 24722
D. melanogaster; UAS-CKIIalpha-il RNAi [60102]	BDSC	RRID: BDSC_60102
D. melanogaster;; UAS-CG10147 RNAi [31183]	Vienna Biocentre	CAT# 31183
D. melanogaster;; UAS-CG10147 RNAi [31943]	BDSC	RRID: BDSC_31943
D. melanogaster UAS-Asciz RNAi [39856]	Vienna Biocentre	CAT# 39856
D. melanogaster;; UAS-MED24 RNAi [33755]	BDSC	RRID: BDSC_33755
D. melanogaster; UAS-Asciz RNAi [51002]	BDSC	RRID: BDSC_51002
D. melanogaster; UAS-TAF RNAi [37549]	Vienna Biocentre	CAT# 37549
D. melanogaster; UAS-CG13894 RNAi [32078]	Vienna Biocentre	CAT# 32078
D. melanogaster;; UAS-CG13894 RNAi [27243]	BDSC	RRID: BDSC_27243
D. melanogaster;; UAS-CG17359 RNAi [25256]	Vienna Biocentre	CAT# 25256
D. melanogaster;; UAS-CG17359 RNAi [26776]	BDSC	RRID: BDSC_26776
D. melanogaster;; UAS-Mirror RNAi [50134]	Vienna Biocentre	CAT# 50134
D. melanogaster; UAS-Meics RNAi [51284]	Vienna Biocentre	CAT# 51284
D. melanogaster;; UAS-Meics RNAi [50636]	BDSC	RRID: BDSC_50636
D. melanogaster;; UAS-Su(z)12 RNAi [42423]	Vienna Biocentre	CAT# 42423
D. melanogaster; UAS-knrl RNAi [47217]	Vienna Biocentre	CAT# 47217
D. melanogaster;; UAS-knrl RNAi [36664]	BDSC	RRID: BDSC_36664
D. melanogaster; UAS-tna RNAi [28071]	Vienna Biocentre	CAT# 28071
D. melanogaster;; UAS-tna RNAi [29372]	BDSC	RRID: BDSC_29372
D. melanogaster; UAS-mu2 RNAi [28343]	Vienna Biocentre	CAT# 28343
D. melanogaster;; UAS-MED24 RNAi [15878]	Vienna Biocentre	CAT# 15878
D. melanogaster; UAS-CG6843 RNAi [61168]	BDSC	RRID: BDSC_61168
D. melanogaster; UAS-CG6843 RNAi [109411]	Vienna Biocentre	CAT# 109411
D. melanogaster UAS-MED10 RNAi [12755]	Vienna Biocentre	CAT# 12755
D. melanogaster;; UAS-MED10 RNAi [34031]	BDSC	RRID: BDSC_34031 (No longer available)
Oligonucleotides		
Sox21b sense gRNA 5' GCAGCAGCAACAATCCGACCA 3'	This Paper	N/A
Sox21b antisense gRNA 5' TGGTCGGATTGTTGCTGCTGC 3'	This Paper	N/A
w501 Sox21b Left HA Forward Primer 5' CTTTCCATTATGCGACGGGG 3'	This Paper	N/A
w501/IL108 Sox21b Left HA Reverse Primer 5' TCGGATTGTTGCTGCTGCTG 3'	This Paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
w501/IL108 Sox21b Right HA Forward Primer 5' CCAGGGACTCGGCCACTCG 3'	This Paper	N/A
w501 Sox21b Right HA Reverse Primer 5' TTTCGTGGCTTGCGTTACAC 3'	This Paper	N/A
IL108 Sox21b Left HA Forward Primer 5' GCTTGTTTTGGACACGCTGG 3'	This Paper	N/A
IL108 Sox21b Right HA Reverse Primer 5' CGAAAGAAACGTTGCCACCA 3'	This Paper	N/A
12.1 Forward Primer (IL108 Introgression) 5' CCAGGGTCGTTCACTT 3'	Hagen et al.[27]	N/A
12.1 Reverse Primer (IL108 Introgression) 5' CCCAGCTTTGTTTCGAATGT 3'	Hagen et al.[27]	N/A
11.9 Forward Primer (IL108 Introgression) 5' CGGACTTGAGCGACCTTCTA 3'	Hagen et al.[27]	N/A
11.9 Reverse Primer (Il108 Introgression) 5' AAAACGAGCGGACTGCTTC 3'	Hagen et al.[27]	N/A
14.2 Forward Primer (IL108 Introgression) 5' TGAGGACATGAGCTTTTCTT 3'	Hagen et al.[27]	N/A
14.2 Reverse Primer (IL108 Introgression) 5' CTTGGCCAACTTATGTGAAC 3'	Hagen et al.[27]	N/A
Dmau_Sox21b_Exon1-2 Forward Primer 5' CTGAAACCGTGCTAAAGGCG 3' AGGCAAACACAATTCAACAGG	This Paper	N/A
Dmau_Sox21b_Exon1-2 Reverse Primer 5' AGGCAAACACAATTCAACAGG 3'	This Paper	N/A
Dsim_Sox21b_Exon1-2 Forward Primer 5' TTGAAGGGGCAATTGAGGCA 3'	This Paper	N/A
Dsim_Sox21b_Exon1-2 Reverse Primer 5' ACGTATCTGATCATTTCCTTTAAGC 3'	This Paper	N/A
Dmau_Sox21b_Exon3-4 Forward Primer 5' TGTGCTTCAGCCGCTAGTTT 3' AGGCAAACACAATTCAACAGG	This Paper	N/A
Dmau_Sox21b_Exon3-4 Reverse Primer 5' GGTGACCCCGACCAAACAT 3'	This Paper	N/A
Dsim_Sox21b_Exon3-4 Forward Primer 5' TTCAGGGGCTCTTAATGCGG 3'	This Paper	N/A
Dsim_Sox21b_Exon3-4 Reverse Primer 5' GGTGACCCCGACCAAACATT 3'	This Paper	N/A
Dmau_Sox21b_Exon5 Forward Primer 5' GGCGTCTTCCGTAGGAGTTT 3' AGGCAAACACAATTCAACAGG	This Paper	N/A
Dmau_Sox21b_Exon5 Reverse Primer 5' GGTTGCATCGCGGTGATAAT 3'	This Paper	N/A
Dsim_Sox21b_Exon5 Forward Primer 5' ATTGGTCTCCGCTACCGTTC 3'	This Paper	N/A
Dsim_Sox21b_Exon5 Reverse Primer 5' TCACCTTTAATAATTGCCTATGCC 3'	This Paper	N/A
Dmau_Sox21b_Exon6 Forward Primer 5' TTGCGAACGGAAAAGAAGCG 3' AGGCAAACACAATTCAACAGG	This Paper	N/A
Dmau_Sox21b_Exon6 Reverse Primer 5' GGCACCCTTAGATGTTATTCAA 3'	This Paper	N/A
Dsim_Sox21b_Exon6 Forward Primer 5' GGTTTCGCGTAGCAAATTCTGA 3'	This Paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Dsim_Sox21b_Exon6 Reverse Primer 5' TATCCATGTCCTTGCCCCTC 3	This Paper	N/A
Recombinant DNA		
pCFD3-dU6:3gRNA	Addgene	RRID: Addgene_49410
pHD-DsRed-attp	DGRC	RRID: DGRC_1361
Software and algorithms		
CRISPR Optimal Target Finder	Gratz et al.[60]	http://targetfinder.flycrispr.neuro.brown.edu/
R Studio 4.2.0	R Core Team[61]	https://rstudio.com
ImageJ/Fiji	Schindelin et al.[62]	https://fiji.sc
PAT-GEOM	Chan et al.[63]	http://ianzwchan.com/my-research/pat-geom/
Multcomp	Hothorn et al.[64]	https://cran.r-project.org/package=multcomp
FactoExtra	Kassambara et al.[65]	https://CRAN.R-project.org/package=factoextra
prcomp	Kassambara et al.[65]	https://stat.ethz.ch/R-manual/R-devel/library/stats/html/prcomp.html
FactoMineR	Kassambara et al.[65]	http://factominer.free.fr
Other		
Standard wall borosilicate glass with filament	World Precision Instruments	CAT# BF100-50-10
0.14mm stainless steel pins (A3)	Watdon	CAT# E6873
Hitachi S-3400N SEM	Hitachi	N/A
Zeiss Axioplan Light Microscope	Zeiss	N/A
Jenoptik ProgRes C3 Camera	Jenoptik	N/A
Zeiss AxioZoom V16	Zeiss	N/A
Femtojet 4i	Eppendorf	N/A
Leica light microscope	Leica	N/A
Zeiss LSM800 Confocal	Zeiss	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Alistair McGregor (alistair.mcgregor@durham.ac.uk)

Materials availability

All unique/stable reagents generated in this study are available from the lead contact without restriction.

Data and code availability

- RNA-seq data[27] underlying this study are deposited in the ArrayExpress database at EMBL-EBI under accession number E-MTAB-9465 (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9465). Sequence data used for allele frequency analysis is available in Data S1F.
- This paper does not report original code.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All *Drosophila* strains and species used in this study can be found in the key resources table. Flies were maintained in a 12 hour light/dark cycle incubator on standard cornmeal food and transferred every two days. All crosses were carried out at 25°C, unless otherwise stated. For this study, all phenotypic analyses were performed on adult male genitalia, aged for at least 3 days prior to storage in 70% ethanol.

METHOD DETAILS

RNAi knockdown of differentially expressed transcription factors in *D. melanogaster*

Differentially expressed TFs were identified from RNA-Seq data generated from the developing male genitalia of *D. simulans* and *D. mauritiana* at 30 – 36 hours after puparium formation as previously reported.[27] Those selected for further analysis in this study were filtered based on chromosome 3 genomic location with respect to previous QTL and introgression mapping studies.[26,27,30,37] Selected differentially expressed TFs were assessed for roles in the development of the male periphallic genitalia using the GAL4-UAS system to drive RNAi in *D. melanogaster*. *UAS-RNAi* lines for each gene were provided by the Vienna Drosophila Resource Centre (VDRC)[66] and from Bloomington Drosophila Stock Centre (BDSC) (NIH P40OD018537). *UAS-RNAi* males were crossed to *NP6333-GAL4 (P(GawB)PenNP6333)* virgin females, (which drives GAL4 expression in all imaginal discs from larval stages and during metamorphosis) also carrying *UAS-Dicer-2* (*P[UAS*-Dcr-*2.D]*.[67] *Sox21b* RNAi was repeated using the *POXN-GAL4 (14.1.1)* driver, which specifically drives in the posterior lobe primordium from larval imaginal disc stage[40] (Data S1A). All crosses were carried out at 25°C[68] (Data S1A). All crosses were performed using a 1:2 male to female ratio. The same conditions were used for each control line, as well as reciprocal hemizygote crosses. Three biological replicates, with a total sample size of n > 13, were phenotyped for each cross. Crosses were transferred to standard cornmeal food every two days and maintained in a 12-hour light/dark cycle incubators. Males aged at least 3 days were collected and then stored in 70% EtOH at -20°C for phenotyping.

Phenotyping of RNAi knockdown flies

The cerci, epandrial posterior lobes and surstyli of the adult male genitalia were dissected in Hoyer's Solution using 0.14 mm diameter stainless steel pins and then mounted in Hoyer's solution. This was done using slides containing eight individual 6 mm diameter wells. To account for body size, the T2 legs of each fly were also dissected and mounted. A Zeiss Axioplan light microscope with a Jenoptik ProgRes C3 camera was used to image each dissected structure. 250X magnification was used for the genital structures, and 160X magnification for the T2 legs. The area of posterior lobes and cerci, and length of T2 tibias were measured manually using ImageJ[69] (Data S1B). The bristles were counted using the light microscope and a tap counter. When drawing the outline of the posterior lobe, an artificial baseline was used as previously described.[28] The area and bristle count were recorded for both pairs of structures per individual and the average was then used for the latter statistical analysis.

Generation of Sox21b-DsRed

To disrupt the reading frame of *Sox21b*, 3XP3-DsRed was inserted 152 bp into exon 1 using CRISPR/Cas9[70] (Figure 2C). This exon was chosen because it did not include restriction sites required in the cloning procedure 1 kb either side from the gRNA cut site, as well as bypassing the conserved HMG-Box domain that may have resulted in off-target effects. The gRNAs were designed using FlyCRISPR[60] and inserted into the pCFD3 plasmid.[70] The homology arms (HA) were amplified by PCR from salt extracted genomic DNA of the focal strains (adapted from Miller et al.[71]), and inserted into the plasmid pHD-DsRed-attp.[60] Plasmids were sequenced prior to injections to verify homology arms and gRNA incorporation. 200 ng/μl of gRNA-pCFD3 and 500 ng/μl of HA-pHD-DsRed-attp were injected (using a Eppendorf FemtoJet 4i and Leica light microscope) into *D. simulans*^{w501} and *IL108* (*D. mauritiana* introgression region spanning 3L: 12,277,961-15,075,323 in an otherwise *D. simulans*^{w501} genetic background) embryos,[27] both carrying *nanos-Cas9* on the X chromosome. 48 hours prior to injection, cages were set up containing apple juice plates and yeast paste, which were changed twice per day prior to injections. Surviving adults from injected embryos were then backcrossed to non-injected adults of the same strain. Progeny were screened for the DsRed marker in their eyes using a Zeiss Axiozoom microscope and those positive were amplified and sequenced to verify genome editing (Figure 2C).

Generation of reciprocal hemizygotes

To generate the reciprocal hemizygote males, we crossed *D. simulans*^{w501} male flies with a mutation in *Sox21b* (1.1 or 1.2) (*D.sim*$^{Sox21b1.1/1.2}$) to *IL108* virgin females to generate male progeny with the genotype *IL108/D.sim*$^{Sox21b1.1/1.2}$ (i.e. flies with a working copy of only the *D. mauritiana Sox21b* allele). Or *IL108* male flies with a mutation in *Sox21b* (1.1 or 1.2) (*IL108*$^{Sox21b1.1/1.2}$) to *D. simulans*^{w501} virgin females to generate male progeny with the genotype *IL108*$^{Sox21b1.1/1.2}$*/D.sim* (i.e. flies with a working copy of only the *D. simulans Sox21b* allele).

Scanning electron microscopy

Flies stored in 70% EtOH were moved to 100% EtOH at least 24 hours prior to imaging. The posterior of the fly was dissected in EtOH. Samples were processed in a critical point dryer and mounted on SEM stubs, then gold coated for 30 seconds. The genitalia were imaged using SE mode at 5 kV in a Hitachi S-3400N SEM, with a working distance of 13 to 14 mm. Whole genitalia was imaged at a magnification of 250x, and individual periphallic structures at 900x.

Posterior lobe shape analysis

Posterior lobes were manually traced as described above. To quantitatively assess shape variation, PAT-GEOM[63] was used to perform Elliptical Fourier Analysis (EFA) on each region of interest (ROI).[72] This software benefitted from being trace-start point, scale, rotation, and translation insensitive. One posterior lobe at random was assessed per fly (n = 45), and twenty descriptors were assigned to each ROI for EFA (Data S1D). Principal component analysis (PCA) was performed using prcomp and factoextra[65] in R,

to evaluate variation between lines. This package standardised the data to have a mean of zero and variance of one prior to computing the PCA. The eigenvalues for each principal component were also computed to identify those with a value above 1, to which these principal components were retained for analysis (Figure S3D). Outlines of lobes corresponding to the extremities of the minimum and maximum values of the principal components labelled on each axis were extracted from the ROI data.

In situ HCR of *Sox21b* in genital discs, embryos and pupal terminalia

To capture *Sox21b* expression, we carried out *in situ* HCR on larval genital discs (at 96 hAEL and 120 hAEL) to complement previous analysis of *Sox21b* in developing male terminalia that showed expression in the posterior lobes and lateral plates.[41] *D. melanogaster*^{w1118}, *D. simulans*^{w501}, *D. mauritiana* D1 and *Sox21b* reciprocal hemizygote male larvae were dissected in ice cold 1XPBS. Each larva was cut in half and the posterior half inverted then placed into 4% formaldehyde in 0.3% PBT (TRITON X-100) for 20 minutes. For *D. melanogaster*^{w1118} stage 16 embryos were collected and fixed in 4% formaldehyde and an equal volume of heptane for 20 minutes, before being washed in methanol. For pupal terminalia, prepupa were collected and denoted as 0 hAPF. Pupae were then cut in half and fixed at the required developmental timepoints using 4% formaldehyde in 1X PBS for 30 minutes. Samples were washed in 1X PBS, and fixed tissue was removed from the pupal casing. The HCR procedure was based on an established protocol.[73,74] The probe from Molecular Instruments to target *Sox21b* was designed as 20 individual hairpins spanning the entirety of the gene, ensuring that all isoforms were captured.[75] A 16 nM probe solution was used, and the sample was incubated for 24 hours at 37 °C on an orbital shaker at 60 RPM. Note, we used half the volumes of each solution compared to the protocol.[73,74] DAPI was diluted in the probe wash solution for nuclear staining of the samples. Samples were then dissected using forceps, mounted in 1X PBS and imaged. Images were obtained on the Zeiss LSM800 upright Confocal Laser Scanning Microscope with a 20X objective for larval discs and embryos, and a 40X objective for pupal samples.

Immunohistochemistry in genital discs and embryos

Genital discs of L3 larvae from *D. melanogaster* were dissected and fixed as described for HCR. Samples were incubated in 10% NGS in 0.1% PBST (Tween-20) for 1 hour prior to the addition of the primary antibody. 1:200 dilution of Rabbit Anti-Drop was used, and samples were incubated overnight at 4°C. 1:600 Anti-Rabbit 488 secondary was incubated with samples the following day in 10% NGS at 4°C overnight. Samples were then dissected and mounted in 1X PBS, and imaged that same day with the same imaging parameters as described above. As Drop is a male-specifically expressed factor, female genital discs could be readily sorted from male discs when imaging.[43] Embryos were fixed as described for HCR and blocked as described for larval samples. 1:100 dilution of Mouse Anti-Eya was used, and samples were incubated overnight at 4 °C. 1:400 Donkey Anti-Mouse 647 secondary was incubated with samples the following day in 10% NGS at room temperature for 2 hours. Samples were then mounted and imaged as described above.

Sox21b coding sequence analysis

Genomic DNA was isolated for ten strains of each species, including the mapped strains *D. simulans*^{w501} and *D. mauritiana* D1 using the high salt extraction method.[71] The full coding sequence of *Sox21b* was amplified with OneTaq® 2X Master Mix with Standard Buffer (New England BioLabs M0482) following the manufacturer recommendations, in four overlapping fragments using the primers listed in the key resources table, all with annealing temperature of 55°C and 30 cycles. PCR products were purified following the GeneJET PCR Purification kit (ThermoFisher Scientific) protocol and sequenced in both directions using Sanger sequencing technology via the Eurofins Sequencing service. As we only found non-synonymous differences between *D. simulans*^{w501} and *D. mauritiana* D1 in exon 1 and 6, we only sequenced the remaining strains of each species for these two fragments. To further evaluate the population frequency of the non-synonymous differences found between *D. simulans*^{w501} and *D. mauritiana* D1, we used polymorphism data from Pool-seq data from 107 strains of *D. mauritiana* and from 50 strains of sub-Saharan *D. simulans*[76] available at http://www.popoolation.at/pgt/.

Behaviour assays using *Sox21b* reciprocal hemizygotes

All mating assays were carried out at 25°C and 70% humidity. Flies were kept in these conditions 24 hours prior to their respective mating assay, allowing acclimatisation. Flies used in the mating assays were reared at 25°C in a 12-hour light/dark cycle. Mating experiments were carried out within the first hour of lights as previously described.[77] Reciprocal hemizygote males were collected as pupae (identified by the presence of sex combs) and aged individually for 4-5 days in separate vials, ensuring they were socially naive. *D. simulans*^{w501} females were similarly collected as pupae (identified by the absence of sex combs) and aged for 4-5 days in vials in groups of 5-10, ensuring virgin status. Reciprocal hemizygote males were screened for the DsRed marker at least 48 hours prior to the mating to ensure full recovery from the brief CO_2 exposure. Single males were paired with single females in a standard food vial with the stopper pushed down to 1 cm above the food, creating a restricted 'mating chamber'. Each pairing was observed for a total of 90 minutes. Mating frequency refers to whether there was evidence of mating in the 90-minute observation period (Y/N), characterised as mounting of the female by the male for a minimum of 7 minutes because this is considered the minimal time for sperm transfer to occur.[19,54] Copulation latency was measured as the time between pairing and copulation onset. Copulation duration was quantified as the time elapsed from initial male mounting and dismounting the female.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were carried out using R version 4.2.0.[61] Measurements of each structure described above were first assessed for normality using the Shapiro Wilk test. Normally distributed data were analysed using Dunnett's test and ANOVA, whereas the Kruskal Wallis test was used for non-normally distributed data.[64] All comparisons included the RNAi knockdown compared to both parental controls. If this test was evaluated as statistically significant, the Tukey's Test / Wilcoxon Rank Sum Test (BH p-adjusted method[78]) was used to identify if the RNAi knockdown was significantly different to both parental controls (Data S1A). Results were concluded as non-significant when p =/> 0.05 or the effect detected in the RNAi knockdown was an intermediate of the two parental controls.

Where the T2 tibia length was significantly different to both parental controls following the tests described above, Pearson/Spearman analysis was performed on the cercus and posterior lobe area of these crosses dependent on Shapiro Wilk test results. If statistically significant, the area of the structure was divided by the tibia length squared, and statistical tests were carried out on the normalised version of the measurements (Data S1A). Effect sizes between each parental control to the RNAi knockdown were calculated using Cohens' d, where the coefficient value represents the number of standard deviations different between the two population means under investigation.[79] To calculate the Cohens' d coefficient, the difference of the means from the two populations of samples were divided by the cumulative pooled standard deviation of them both.

The phenotypic measurements for the reciprocal hemizygotes and null mutants were analysed using an independent t-test or Wilcoxon Rank Sum Test dependent on the normality of the data. Details of morphometric analysis of posterior lobe shape can be found above. For PCA, a MANOVA test was used, followed by univariate analysis to assess the significance of each individual principal component between the two reciprocal hemizygotes.

A Fisher's exact test was conducted to analyse copulation frequency. The Shapiro Wilk test was conducted to assess the distribution of the datasets. As the copulation latency dataset was not normally distributed, a Kruskal Wallis test was performed. Copulation duration was analysed using an independent t-test. Violin plots indicate the mean of the data, with the first quartile and the third quartile value shown. The width of the violin plots represents the frequency of data points at assigned values.