GENETICS

Hidden RNA pairings counteract the "first-come, first-served" splicing principle to drive stochastic choice in *Dscam1* splice variants

Haiyang Dong[1]†, Bingbing Xu[1]†, Pengjuan Guo[1]†, Jian Zhang[1]†, Xi Yang[2], Lei Li[1], Ying Fu[1], Jilong Shi[1], Shixin Zhang[1], Yanda Zhu[1], Yang Shi[1], Fengyan Zhou[1], Lina Bian[1], Wendong You[3], Feng Shi[1], Xiaofeng Yang[3], Jianhua Huang[4], Haihuai He[2], Yongfeng Jin[1]*

Drosophila melanogaster Dscam1 **encodes 38,016 isoforms via mutually exclusive splicing; however, the regulatory mechanism behind this is not fully understood. Here, we found a set of hidden RNA secondary structures that balance the stochastic choice of *Dscam1* splice variants (designated balancer RNA secondary structures). In vivo mutational analyses revealed the dual function of these balancer interactions in driving the stochastic choice of splice variants, through enhancement of the inclusion of distal exon 6s by cooperating with docking site–selector pairing to form a stronger multidomain pre-mRNA structure and through simultaneous repression of the inclusion of proximal exon 6s by antagonizing their docking site–selector pairings. Thus, we provide an elegant molecular model based on competition and cooperation between two sets of docking site–selector and balancer pairings, which counteracts the "first-come, first-served" principle. Our findings provide conceptual and mechanistic insight into the dynamics and functions of long-range RNA secondary structures.**

INTRODUCTION

Pre-mRNA alternative splicing is a major source of proteomic and functional diversity (*1*–*3*). In humans, ~95% of multi-exon genes are subjected to alternative splicing (*4*, *5*), and splicing defects are associated with a variety of genetic diseases (*6*, *7*). Common alternative splicing mechanisms include exon skipping, intron retention, alternative 5′ or 3′ splice site usage, and mutually exclusive splicing (*2*). Mutually exclusive splicing occurs when only one exon from a cluster of variable exons is spliced into a specific mRNA product (*8*, *9*). The most astonishing example of this can be found in the *Drosophila melanogaster* Down's syndrome cell adhesion molecule 1 (*Dscam1*) gene, which potentially generates 38,016 different isoforms through mutually exclusive splicing of exon clusters 4, 6, 9, and 17 (*10*). Growing evidence has revealed that the notable diversity of *Dscam1* isoforms is required for both neuronal wiring and immune defense (*11*–*13*).

In an attempt to explain how only one exon variant of *Dscam1* is selected at a time from a cluster of exons, several models based on competing RNA secondary structures have been proposed (*14*–*18*). This mechanism was initially found in the exon 6 cluster of *Dscam1*, where the intronic docking site downstream of exon 5 can pair competitively with selector sequences upstream of each exon 6 variant (*14*, *15*). Similar structural arrangements have been identified in the exon 4 and exon 9 clusters of *Dscam1* (*18*–*20*); however, in contrast to the docking site located upstream of the exon 6 cluster, the docking sites for the exon 4 and exon 9 clusters are located within their respective downstream introns. The heterogeneous nuclear ribonucleoprotein 36 also ensures splicing fidelity in the exon 6 cluster but not in the exon 4 or exon 9 cluster (*21*). Moreover, a locus control region immediately upstream of the docking site is involved in the inclusion of the nearest exon when the docking site is paired with its upstream selector sequence (*22*).

A long-standing question regarding mutually exclusive splicing is how docking site–selector pairing is regulated (*9*, *14*). Proximal selector sequences are only a few hundred nucleotides away from the docking site, whereas distal selector sequences are located over 10,000 nucleotides away from the docking site. If docking site–selector pairing were dictated on a cotranscriptional "first-come, first-served" basis (*23*), there would be a bias toward docking site–proximal exon 6 variants; however, this is not observable in all tissues (*12*, *24*–*26*). These splicing outcomes in terms of exon 6 inclusion cannot be explained by compensation from splice site strength, because distal exon 6s do not display a preference for stronger splice sites compared to proximal exon 6s. The predicted thermodynamic stability of docking site–selector pairing does not correlate with the selection frequency of the associated exons. The frequency of some exon 6 variants actually increased following deletion of the docking site in a previous study examining *D. melanogaster* /*Drosophila virilis* (Dme/Dvi) *Dscam1* mutant constructs (*24*). It is therefore likely that unknown mechanisms regulate selection of the *Dscam1* exon 6 cluster.

In this work, we found a previously unknown RNA secondary structure hidden within the exon 6 cluster that balances the stochastic selection of *Dscam1* splice variants (designated balancer RNA secondary structures). Targeted mutational analyses using a CRISPR-Cas9 system revealed that hidden RNA pairings drive the stochastic choice of splice variants by simultaneously enhancing the inclusion of distal exon 6s by facilitating docking site–selector pairing and by repressing the inclusion of proximal exon 6s. We found that the balancers and docking site–selector base pairings cooperated to form a strong multidomain structure that enhanced distal exon inclusion. From this, we developed a molecular model for the regulation of the stochastic, mutually exclusive splicing of *Dscam1* exon 6 based on competition and cooperation between two sets of docking

[1]MOE Laboratory of Biosystems Homeostasis & Protection and Innovation Center for Cell Signaling Network, College of Life Sciences, Zhejiang University, Hangzhou, China. [2]Department of Neurosurgery and State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu China. [3]Department of Neurosurgery, First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China. [4]Institute of Insect Sciences, Zhejiang University, Hangzhou, China.
*Corresponding author. Email: jinyf@zju.edu.cn
†These authors contributed equally to this work.

site–selector and balancer RNA secondary structures. Our work suggests that *Dscam1* has evolved compensatory mechanisms that balance the distance and strength of docking site–selector base pairing to drive the stochastic selection of *Dscam1* splice variants and provides an explanation for the lack of obvious 5′ to 3′ preference during exon 6 variant selection. Moreover, genetic analysis indicated that disruption of docking site pairing or balancer base pairing led to growth and neuronal anomalies, suggestive of their physiological significance. We provide an additional framework for the regulation of complex, mutually exclusive splicing and new insight into the role of long-range RNA secondary structures in gene regulatory networks.

RESULTS

In vivo mutagenesis confirmed cluster-wide docking site–selector base pairing in the exon 6 cluster

To verify previously predicted docking site–selector base pairings (*14*), we generated mutant flies with both disruptive and double compensatory mutations using CRISPR-Cas9 and examined their effects on exon 6 inclusion (Fig. 1, A and B). We constructed homozygous knock-in flies with docking site mutations (designated *Dscam*^{M1}), mutations in the selector sequence upstream of exon 6.5 (*Dscam*^{M2}), or double compensatory mutations (*Dscam*$^{M1+M2}$, hereafter referred to as *Dscam*^{M12} for simplicity) (Fig. 1C). Reverse transcription polymerase chain reaction (RT-PCR) analyses revealed a small fraction of the transcripts (~10%) lacking exon 6 variants in the head tissues of *Dscam*^{M1} mutant flies compared to the wild type (WT; Fig. 1D and fig. S1A); these results were largely consistent with those of a previous report, in which deletion of the docking site of Dme/Dvi *Dscam1* resulted in a 14-fold increase in exon 6 skipping compared with the control (*24*). Here, high-throughput sequencing of the RT-PCR products revealed that the docking site mutation caused a marked decrease in the inclusion frequency of most exon 6s, including exon 6.5 (Fig. 1E and fig. S1B). Mutation of the selector sequence upstream of exon 6.5 markedly reduced the inclusion frequency of exon 6.5 in *Dscam*^{M2} flies compared to the WT but did not affect the inclusion of other exon 6 variants (Fig. 1E and fig. S1B). *Dscam*^{M12} double mutants exhibited higher exon 6.5 inclusion frequencies compared to the WT (Fig. 1E and fig. S1B); a possible explanation for this is that the mutated docking site may have paired with the mutated selector sequence but not with other selector sequences. Therefore, our in vivo analysis examining disruptive and compensatory docking site/selector sequence mutations suggests that exon 6.5 inclusion is dependent on base pairing interactions between the docking site and its selector sequence.

Further efforts to verify the base pairing between the docking site and distal selector sequences using the disruptive/double compensatory mutation method were hampered by the technical difficulties associated with the large size of interspacing. For example, introducing a double compensatory mutation spanning >11,000 base pairs (bp) between the docking site and selector 6.48 would be extremely challenging (Fig. 1A).

To experimentally verify docking site–selector base pairing in a cluster-wide manner, we introduced a series of point mutations into the endogenous docking site using CRISPR-Cas9 (*Dscam*$^{M3–M10}$; Fig. 1F). This enabled the strength of the base pairing interactions between the docking site and selector sequences to be finely tuned and calculated, as opposed to fully disrupting the interaction. We hypothesized that a strong correlation between the inclusion frequency of a given exon 6 and the base pairing strength among a series of docking site mutation flies would represent bona fide base pairing (i.e., exon 6.5; Fig. 1G). No significant differences in exon 6 inclusion were observed in the head tissue of flies harboring these mutations compared to the WT (fig. S1C). However, these mutant flies exhibited a range of inclusion frequencies among exon 6 variants (fig. S1D). On the basis of previously predicted base pairing (*14*), we calculated the base pairing strength between the selector sequences and each docking site mutation and plotted this against the exon 6 inclusion frequencies for each mutant; for 40 of 47 exon 6 variants (not including 6.11), the inclusion frequency was moderately to strongly positively correlated with the predicted base pairing strength between the docking site and selector sequence (Fig. 1, H and I, and fig. S2). Thus, these data support most of the previously predicted base pairings between the docking site and the selector sequences.

In 7 of the 47 exon 6s, there was no or poor positive correlation between exon 6 inclusion frequency and base pairing strength among flies with the docking site mutation (Fig. 1J). Notably, the inclusion frequency of exon 6.1 did not correlate with base pairing strength, in contrast to the other proximal exon 6s, which exhibited strong correlation (Fig. 1, I and J). To investigate this further, we disrupted the putative selector sequence upstream of exon 6.1 using CRISPR-Cas9, which was predicted to pair with the docking site (*14*). The inclusion frequency of exon 6.1 was higher in homozygous viable *Dscam*$^{Ms6.1}$ flies containing the selector-sequence mutation compared to the WT control (fig. S1, E to G). The data from the correlation and mutation analyses indicated that the predicted selector sequence upstream of exon 6.1 was incorrect, suggesting that docking site–proximal exon 6.1 splicing might not depend on docking site–selector sequence base pairings but might be controlled by a similar mechanism as 14-3-3ξ gene (*18*). Similarly, poor correlation between the inclusion frequency of exon 6.22 and the base pairing strength between its selection sequence and the docking site is likely due to the incorrect prediction of the exon 6.22 selector sequence. On the basis of sequence conservation and correlation analysis, we revised the selector sequences upstream of exons 6.22 and 6.38 (Fig. 1K and fig. S1, H and I). However, poor correlation in docking site–selector base pairings did not necessarily mean that the predictions were incorrect; for example, the correlation between the inclusion frequency of exon 6.21 and base pairing strength was likely poor because the nucleotide mutations examined here scarcely affected the predicted base pairings of exon 6.21 (Fig. 1J).

Notably, we found that in most of the docking site–distal exon 6s (exons 6.35 to 6.48), the correlation between their inclusion frequencies and the docking site–selector sequence base pairing strength was poor (Fig. 1I). This trend is largely consistent with the observation that the docking site mutation almost abrogated the inclusion of docking site–proximal exon 6s (except for exon 6.1) but did not reduce and even increased the inclusion of some docking site–distal exon 6s in *Dscam*^{M1} docking site mutants compared to the WT (fig. S1B). These results indicate that the inclusion of several exon 6s may still be promoted in the absence of a functional docking site, suggesting that an as-yet-unknown mechanism could be involved in exon 6 inclusion.

Exon 6.40 is highly included even in the absence of the docking site

To investigate why some exon 6s were included at high frequencies in the absence of the docking site, we analyzed the effect of the sequences

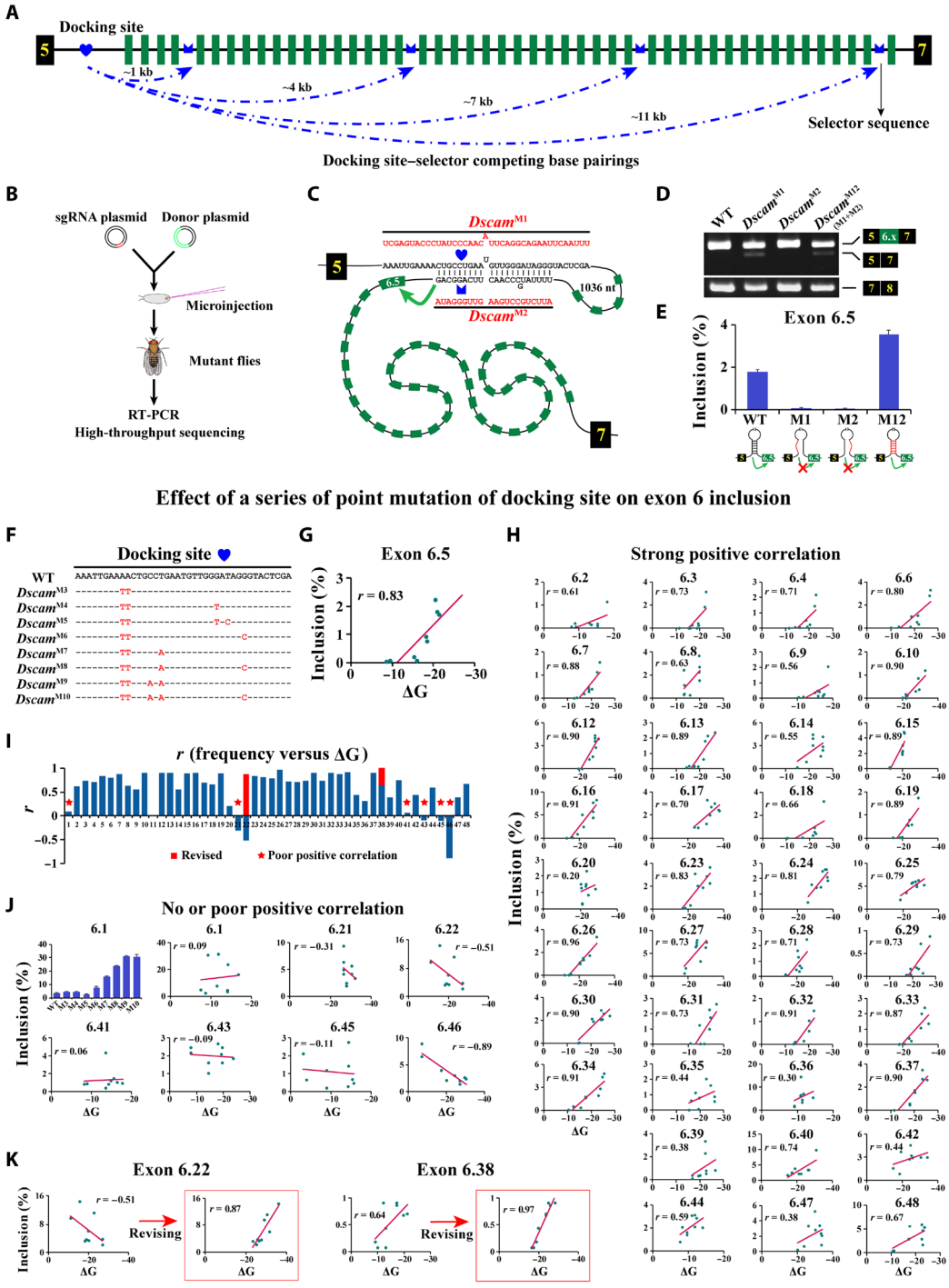

Fig. 1. CRISPR-Cas9 mutagenesis verifies cluster-wide docking site–selector pairing in the exon 6 cluster (see also figs. S1 and S2). (**A**) Schematic diagrams of docking site–selector RNA pairings of *D. melanogaster Dscam1*. Constitutive exons (black boxes), alternative exon 6 (green boxes), docking site (blue heart), and selector sequences (blue crowns) are shown. The dashed arrow represents the RNA pairing interaction. (**B**) Schematic of in vivo mutagenesis in the intact fly using targeted CRISPR-Cas9. (**C**) Schematic of disrupting and double compensatory mutations. Mutations introduced into the predicted RNA secondary structures are indicated in *Dscam*^{M1} and *Dscam*^{M2} flies. A combination of upper and lower sequence mutations (*Dscam*^{M12}) led to the restoration of structure. The green arrow is depicted as activating inclusion of the exon 6.5. (**D**) RT-PCR diagram from the head of WT and mutant flies. (**E**) Effect of mutations on exon 6.5 inclusion is indicated for disruptive mutations (*Dscam*^{M1} and *Dscam*^{M2}) and compensatory double mutations (*Dscam*^{M12}). (**F**) Schematic of docking site mutations at one nucleotide resolution (*Dscam*$^{M3\text{-}M10}$). (**G**) Exon 6.5 inclusion was positively correlated with the predicted pairing strength in WT and mutant flies. (**H**) Exon 6s whose inclusion frequency positively correlated with the predicted pairing strength are shown. (**I**) Correlation coefficients of exon 6 selection with the strength of predicted RNA pairing for each exon 6. Exon 6s with no or poor correlation are marked by red stars. (**J**) Seven exon 6s whose inclusion frequency did not positively correlate with the predicted pairing strength are shown. The inclusion frequency of exon 6.1 in the WT and docking site mutation flies is shown. (**K**) Correlation comparison between previously predicted and revised base pairings in exons 6.22 and 6.38.

downstream of the docking site in exon 6 inclusion. We used targeted CRISPR-Cas9 to generate a series of knockout mutants with varying lengths of deletions in the docking site and its downstream sequence (designated *Dscam*$^{\Delta Dock1-\Delta Dock3}$; Fig. 2A). Similar to *Dscam*^{M1}, a small fraction of the transcripts lacking exon 6 variants was observed in various tissues of *Dscam*$^{\Delta Dock1}$ mutants harboring docking site deletion; frequencies of overall exon 6 inclusion increased as the length of the deleted region increased in *Dscam*$^{\Delta Dock1-\Delta Dock3}$ mutants (Fig. 2B and fig. S3A). High-throughput sequencing of RT-PCR products revealed a marked decrease in the inclusion frequency of all but a

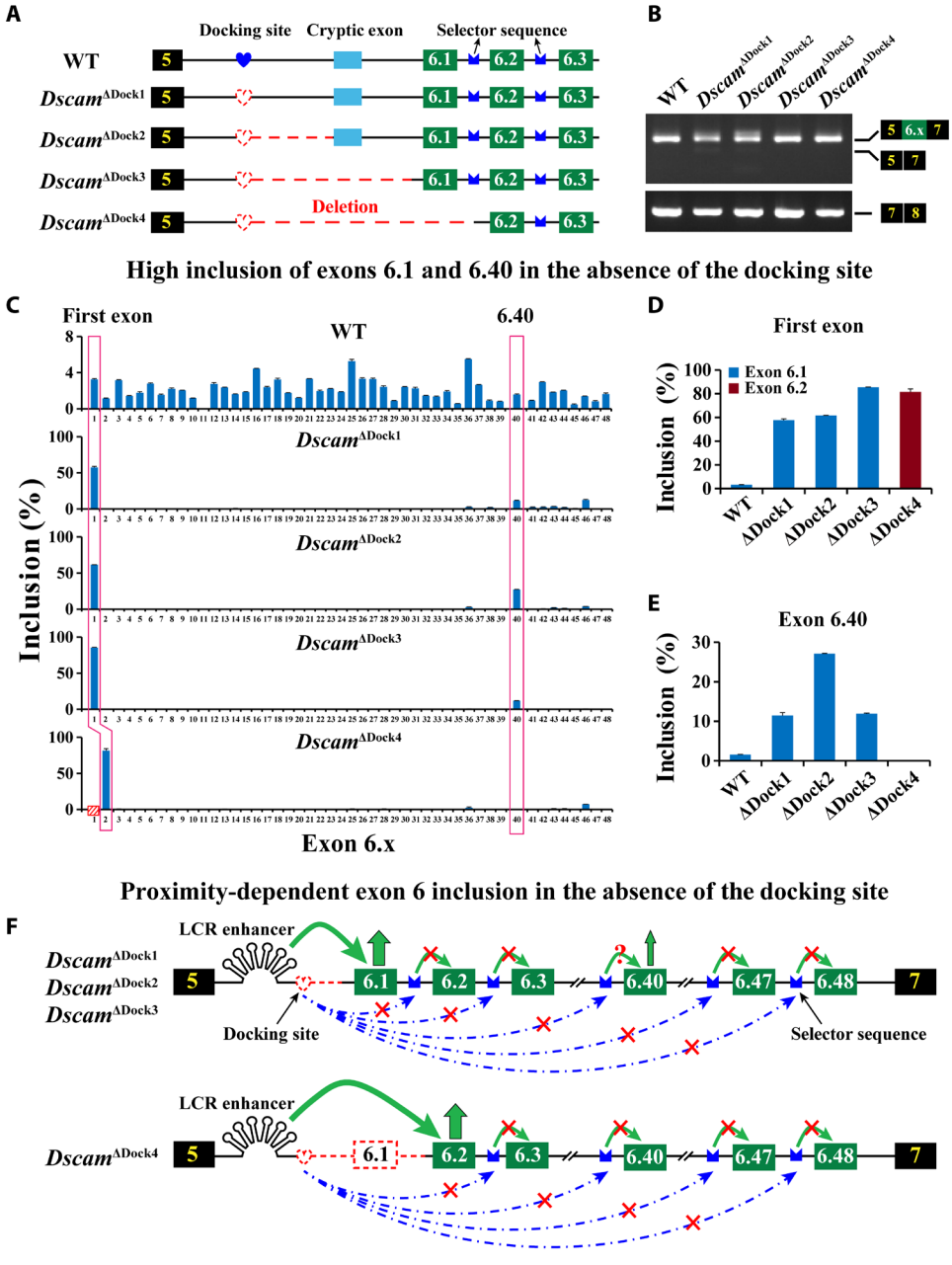

Fig. 2. Exon 6.40 is highly included in absence of the docking site (see also fig. S3). (**A**) Schematic of docking site mutations in *D. melanogaster Dscam1* exon 6 cluster. A series of knockout mutant flies were constructed by targeted CRISPR-Cas9 with varying degrees of deletions of the docking site sequence and its downstream sequence (designated *Dscam*$^{\Delta Dock1-\Delta Dock4}$). Cryptic exons upstream of exon 6.1 are indicated (*43*). (**B**) RT-PCR diagram from the head of WT and mutant flies. *Dscam*$^{\Delta Dock1}$ and *Dscam*$^{\Delta Dock2}$ showed slight exon 6 skipping while *Dscam*$^{\Delta Dock3}$ and *Dscam*$^{\Delta Dock4}$ showed no obvious exon 6 skipping. (**C**) Inclusion frequency of each exon 6 variants in WT and *Dscam*$^{\Delta D1-\Delta D4}$. (**D** and **E**) The effect of RNA pairings on first exon 6 and exon 6.40 inclusion WT and *Dscam*$^{\Delta D1-\Delta D4}$. See also fig. S3. (**F**) Model of proximity-dependent exon 6 inclusion. Exon 6.1 was dominantly included in the deletion of docking sites in *Dscam*$^{\Delta Dock1-\Delta Dock3}$ fly mutants, while inclusion of the other exon 6s might be inhibited except for a few exon 6s (i.e., exon 6.40). Exon 6.2 was dominantly included in the *Dscam*$^{\Delta Dock4}$ fly with deletion of a fragment spanning the docking site and entire exon 6.1. The red dashed line depicts the deletion sequence. The red cross indicates disruption of docking site–selector pairings. The curved green arrow indicates proximity-dependent activation of exon 6 variants. The blue dashed arrow represents the docking site–selector pairing. LCR, locus control region.

few exon 6 variants in the docking site mutants (Fig. 2C and fig. S3B). Exon 6.1 was the most frequently selected exon 6 variant in *Dscam*$^{\Delta Dock1-\Delta Dock3}$ mutants, with inclusion frequencies ranging from ~60 to ~85%. Exon 6.1 accounted for ~85% of the total exons included in *Dscam*$^{\Delta Dock3}$ flies with deletions to approximately 34 bp upstream of the 3′ splice site of exon 6.1 (Fig. 2D and fig. S3B, i). Exon 6.1 inclusion frequencies increased as the length of the deleted downstream region increased, suggesting that exon 6.1 was included in a proximity-dependent manner (Fig. 2F). Furthermore, when the fragment encompassing the docking site and entire exon 6.1 was deleted (*Dscam*$^{\Delta Dock4}$; Fig. 2A), exon 6.2 accounted for ~80% of the total inclusion frequency in *Dscam*$^{\Delta Dock4}$ flies (Fig. 2D). In the Δexon6.1 mutant, exon 6.2 turned to be the closest exon to the docking site and the most frequently included exon. Thus, this result suggests a model of proximity-dependent activation of the proximal exon 6 variant (Fig. 2F).

Unexpectedly, exon 6.40 inclusion was approximately sevenfold higher in the *Dscam*$^{\Delta Dock1}$ mutant compared to the WT (Fig. 2E and fig. S3B, ii). Furthermore, when the deletion was extended to ~171 bp upstream of the 3′ splice site of exon 6.1, the exon 6.40 inclusion frequency was ~17-fold higher in *Dscam*$^{\Delta Dock2}$ flies compared to WT flies (Fig. 2E and fig. S3B, ii). Obviously, these data cannot be explained by the original model (*14*). Moreover, it is unlikely that this was due to off-target effects because a similar phenomenon was observed in three *Dscam*$^{\Delta Dock1-\Delta Dock3}$ mutant lines (Fig. 2C), as well as the *Dscam*^{M1} line containing the docking site mutation (fig. S1B). However, when we deleted a fragment spanning the docking site and exon 6.1 in its entirety, almost no exon 6.40 inclusion was observed in *Dscam*$^{\Delta Dock4}$ flies (Fig. 2E and fig. S3B, ii). The data from the various mutant lines suggest that the exon 6.1 sequence is required for high exon 6.40 inclusion frequencies in the absence of the docking site (Fig. 2F).

Balancer RNA secondary structures in the exon 6 cluster of *Dscam1* in flies

Exon 6.40 inclusion is dependent on the exon 6.1 sequence in *Dscam*$^{\Delta Dock1-\Delta Dock3}$ mutants; therefore, we speculated that a sequence upstream of exon 6.40 may pair with exon 6.1 to form a long-distance RNA secondary structure, thereby activating exon 6.40 inclusion via an approximation-activation mechanism (*16*). Sequence alignment revealed a conserved intronic element (designated Dbs6.40) in the intron upstream of exon 6.40 across Schizophora species, spanning ~100 million years (Fig. 3 and fig. S4A). By probing the exon 6 cluster in silico using the Dbs6.40 sequence, we identified a highly complementary sequence (Ubs6.1) in the intron-exon 6.1 boundary region. The predicted base pairings were highly conserved, with a common core in the exon 6 cluster of *Dscam1* in all 20 Schizophora species (fig. S4B). Clear evidence of compensatory structural covariations

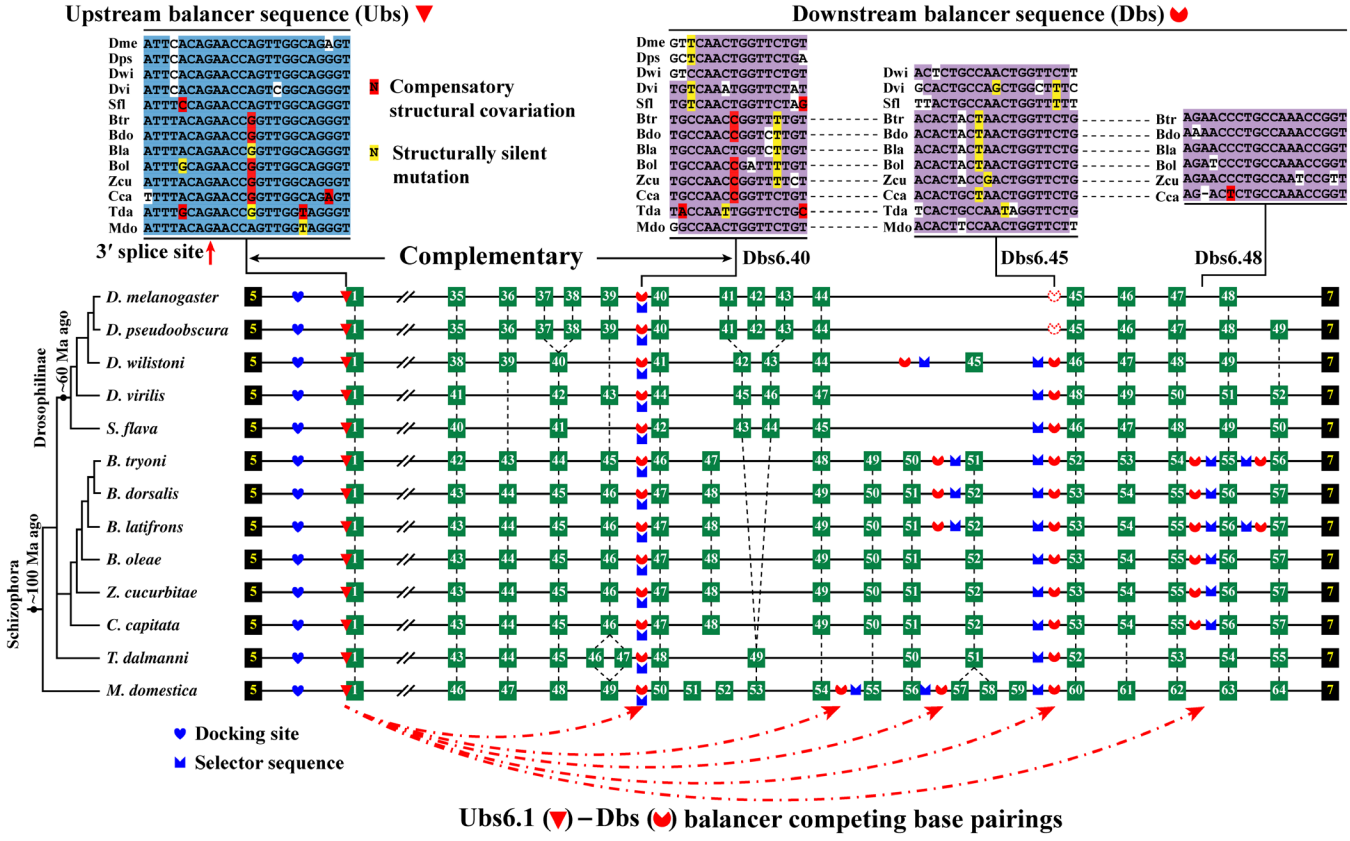

Fig. 3. Balancer RNA secondary structures in the exon 6 cluster of fly *Dscam1* (see also figs. S4 and S5). The arrangement of conserved and lineage-specific balancer sequences across fly species. A phylogenetic tree depicting the relationships between each species is shown on the left. Above are conserved sequences for different clades (see table S1 for abbreviations). Nucleotide structural covariations that maintain the structural integrity of the balancer base pairing are shaded in red, and their covariation intermediates (U-G and G-U) are shaded in yellow. Each upstream balancer sequence (marked by red triangles) was supposed to pair with downstream balancer sequences (marked by red sectors). The introns are not drawn to scale. The red dashed arrow represents the balancer base pairings.

and evolutionary intermediates exist within the core region (Fig. 3 and figs. S4B and S5B).

Further comparative genomic analyses revealed species-specific and conserved sequences within the exon 6 cluster body, which potentially pair with the upstream intron-exon 6.1 boundary region (Fig. 3, figs. S4 and S5, and Supplementary Text). Dbs6.40 homolog sequences might be simultaneously duplicated and purifying-selected with duplication-mediated variable exon expansion (*27*, *28*). This type of RNA secondary structure may be involved in a compensatory mechanism to balance stochastic exon 6 variant selection (see below); therefore, depending on the position of these elements (upstream or downstream), they were designated upstream balancer sequences (Ubs6.x) or downstream balancer sequences (Dbs6.x) (Fig. 3A). For example, the Ubs located in exon 6.1 was denoted Ubs6.1, and the Dbs upstream of exon 6.40 was denoted Dbs6.40. Moreover, the balancers tended to be located toward the 3′ region of the exon 6 cluster, and their frequency increased with the size of the exon 6 clusters. The distribution of the balancer sequences may be linked to their specific functions.

In vivo verification of balancer RNA secondary structures

To verify this balancer RNA secondary structure, double compensatory mutations in the core of the predicted stem to test the effects on exon 6.40 inclusion would have been ideal. However, efforts to perform these experiments were hampered due to technical challenges associated with the very large interspacing of base pairing (~10,000 nt). To confirm the balancer RNA secondary structures described above, we mutated Ubs6.1 in *Dscam*$^{\Delta Dock3}$ using CRISPR-Cas9 to investigate how changes in pairing strength could influence exon 6.1 or 6.40 inclusion frequency (Fig. 4A). We hypothesized that if Ubs6.1 and Dbs6.40 paired to form an RNA secondary structure, then exon 6.40 inclusion frequency would be strongly correlated with Ubs6.1-Dbs6.40 base pairing strength. Strengthening the base pairing significantly enhanced exon 6.40 inclusion in *Dscam*$^{\Delta Dock3M1}$ flies (Fig. 4, B and C). By contrast, weakening the Ubs6.1-Dbs6.40 pairing strength significantly reduced the exon 6.40 inclusion frequency in *Dscam*$^{\Delta Dock3M2}$ and *Dscam*$^{\Delta Dock3M3}$ flies. Our results indicated that exon 6.40 inclusion was positively correlated with base pairing strength (Fig. 4C), suggesting that Ubs6.1 and Dbs6.40 work together to regulate exon 6 inclusion by base pairing. In addition, exon 6.1 inclusion was negatively correlated with Ubs6.1-Dbs6.40 pairing strength (Fig. 4D), consistent with the notion that RNA secondary structure can inhibit splicing by masking acceptor splice sites (*16*). Thus, the two cis elements Ubs6.1 and Dbs6.40 likely regulate exon 6 inclusion by forming long-distance RNA secondary structures at least in the absence of a docking site.

To further verify the balancer RNA secondary structure in the presence of docking site, we mutated Dbs6.40 to test the effects on the inclusion of exon 6.1 and exon 6.40 (*Dscam*$^{\Delta Dbs6.40}$; Fig. 4E). Dbs6.40 is located in a region similar to that of the selector sequence upstream of exon 6.40; therefore, it was unclear whether the reduced exon 6.40 inclusion frequencies observed in Δ*Dbs6.40* mutants were caused by the disruption of the balancer RNA secondary structure, by docking siteselector base pairing, or by both. This prompted us to compare the exon 6.1 inclusion frequencies between WT and Δ*Dbs6.40* mutant flies. Δ*Dbs6.40* flies exhibited higher exon 6.1 inclusion frequencies than WT flies (Fig. 4E, right). These data further supported our hypothesis that Ubs6.1 pairs with Dbs6.40 to regulate exon 6 inclusion. To rule out the possibility that the increase in exon 6.1 inclusion frequency in Δ*Dbs6.40* flies was a result of the decreased inclusion of other distal exon 6s, we constructed two mutant lines in which exons 6.40 to 6.48 were deleted both with and without Dbs6.40 (*Dscam*$^{\Delta Ex40-48}$ and *Dscam*$^{\Delta Dbs40-Ex48}$; Fig. 4F). *Dscam*$^{\Delta Dbs40-Ex48}$ flies lacking Dbs6.40 displayed significantly higher exon 6.1 inclusion frequencies compared to *Dscam*$^{\Delta Ex40-48}$ mutants with Dbs6.40, which is consistent with disruption of the Ubs6.1-Dbs6.40 balancer RNA secondary structure in *ΔDbs6.40* flies. We concluded that the balancer elements Ubs6.1 and Dbs6.40 likely form an RNA secondary structure.

Dual functions of balancer pairing in choice of exon 6 variants

Next, we investigated the potentially evolutionary and regulatory benefits of the balancer RNA structures. In addition to the repression of exon 6.1 inclusion by masking the 3′ splice site, we propose two functions for the RNA base pairing interactions. First, balancer RNA secondary structures may cooperate with docking site–selector secondary structures, potentially generating stronger, multidomain RNA structures (Fig. 5A, right). For example, when Ubs6.1 pairs with the balancer sequence upstream of exon 6.40 (Dbs6.40), this interaction could bring the docking site within close proximity of the selector sequence upstream of exon 6.41 or its downstream exon 6s. In this scenario, the inclusion of exon 6.41 would be enhanced. The second possibility is that when Ubs6.1 pairs with the balancer sequence upstream of exon 6.40 (Dbs6.40), this could position the docking site away from the selector sequence upstream of the proximal exon 6s (i.e., exons 6.2, 6.3, and 6.4), thereby antagonizing RNA pairing between the docking site and selector sequence upstream of the proximal exon 6s. In this scenario, balancer base pairings might inhibit the selection of certain proximal exon 6 variants (Fig. 5A, left).

Balancer base pairing enhances distal exon 6 inclusion

We then investigated how balancer RNA secondary structures could affect the inclusion of exon 6 variants by altering the Ubs6.1 or Dbs6.40 sequence. We found that deleting exon 6.1 led to decreased exon 6.41 inclusion in *Dscam*^{BM1} flies (Fig. 5B, i, and fig. S6A), which could be explained by failed Ubs6.1-Dbs6.40 pairing. To preclude the potential effect of distance on exon 6 inclusion, we generated fly mutants by introducing single mutations to destroy the 5′ splice site in combination with Ubs6.1 mutations (*Dscam*^{BM2} and *Dscam*^{BM3}; Fig. 5B). A similar trend was displayed by both the *Dscam*^{BM2} and *Dscam*^{BM1} mutant lines (Fig. 5B, ii). Likewise, weakening the Ubs6.1-Dbs6.40 pairing led to a marked decrease in exon 6.41 inclusion frequency (*Dscam*^{BM3}; Fig. 5B, iii). Notably, the inclusion of exons 6.42 and 6.43 was also reduced in *Dscam*^{BM1}, *Dscam*^{BM2}, and *Dscam*^{BM3}, with the degree of inclusion depending on exon order (Fig. 5B, i to iii, and fig. S6, A and B). Our analysis revealed that the reducing effect caused by three independent Ubs6.1 mutations strongly correlated with distance from Dbs6.40 to downstream selector sequences (Fig. 5B, I to III). By contrast, strengthening the Ubs6.1-Dbs6.40 pairing markedly increased the inclusion frequency of exon 6.41 and downstream exon 6s, exhibiting the greatest effect on exon 6.41 inclusion and the least notable effect on distal exon inclusion (*Dscam*^{BM4}; Fig. 5B, iv and IV). Similar changes in exon 6 usage were observed at different developmental stages of mutants, although the alterations were subject to development- and tissue-specific regulation (fig. S8). Collectively, these data indicate that balancer RNA secondary structures can enhance the inclusion of

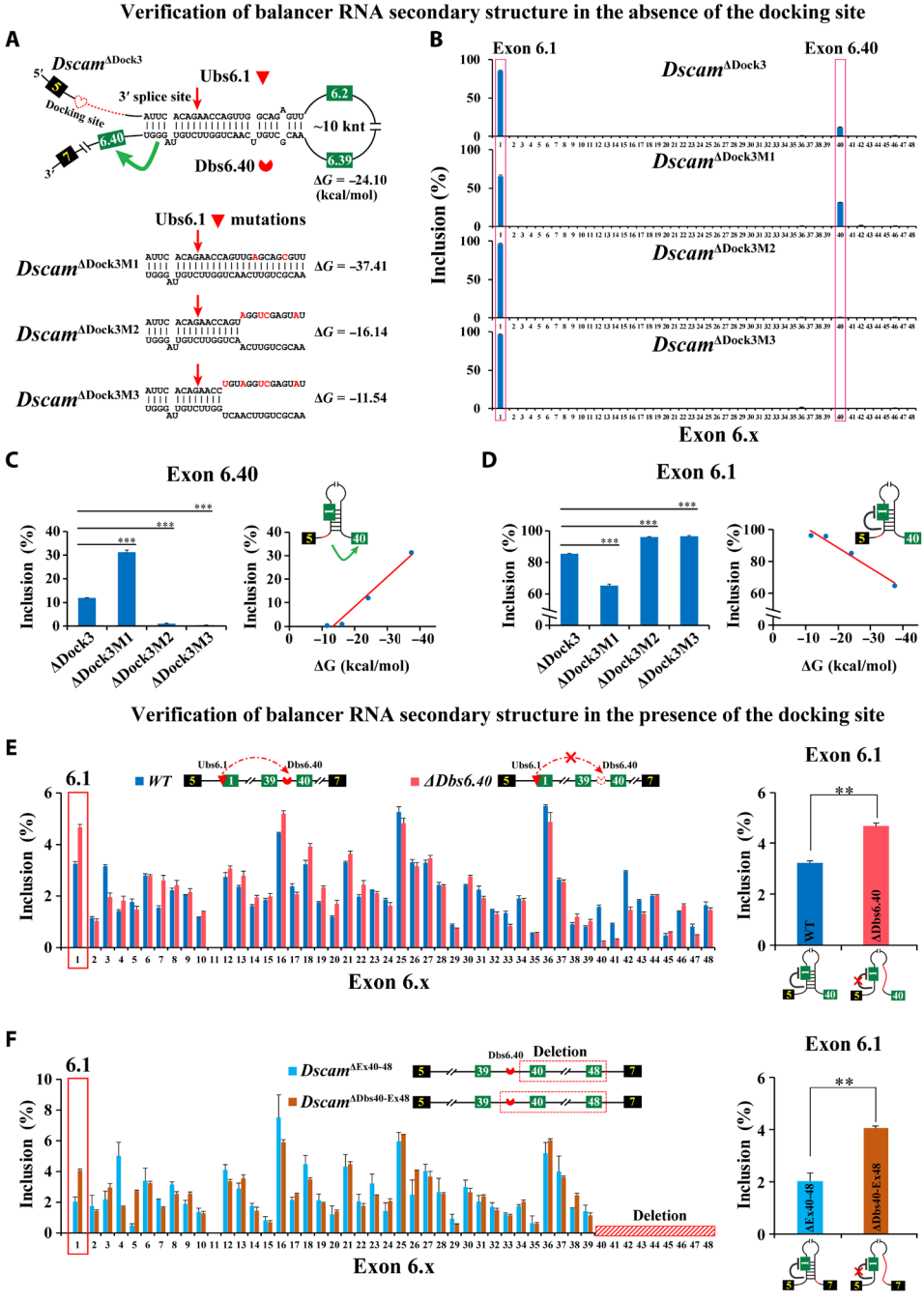

Fig. 4. In vivo verification of balancer RNA secondary structures. (**A**) Schematic of the Ubs6.1 mutation to verify balancer RNA secondary structure in the absence of the docking site. We mutated Ubs6.1 from *Dscam*$^{\Delta Dock3}$ to change the predicted pairing strength to test the effects on the inclusion of exons 6.1 and 6.40. The red dashed line depicts the deletion sequence. (**B**) Effect of Ubs6.1 mutation on exon 6 inclusion in the absence of the docking site. (**C**) Effect of Ubs6.1 mutation on exon 6.40 inclusion. Exon 6.40 inclusion was positively correlated with base pairing strength, indicating that Ubs6.1 and Dbs6.40 regulate exon 6 inclusion through base pairing. The proposed model is shown in the inset. (**D**) Effect of Ubs6.1 mutations on exon 6.1 inclusion. Exon 6.1 inclusion was negatively correlated with Ubs6.1-Dbs6.40 pairing strength, consistent with the notion that RNA secondary structure could inhibit splicing by masking splice sites. The proposed model is shown in the inset. (**E**) Dbs6.40 deletion (*ΔDbs6.40*) led to a remarkable increase of exon 6.1 inclusion in the presence of the docking site. (**F**) Exon 6.1 inclusion in the *Dscam*$^{\Delta Dbs40\text{-}Ex48}$ fly lacking Dbs6.40 is significantly more than in *Dscam*$^{\Delta Ex40\text{–}48}$ with Dbs6.40. These data support the formation of the Ubs6.1-Dbs6.40 balancer RNA secondary structure. $^{**}P < 0.01$; $^{***}P < 0.001$ (Student's *t*-test, two-tailed).

multiple distal exon 6 variants downstream of exon 6.40 in a proximity-correlated manner (Fig. 5C).

We next introduced mutations to Dbs6.40 to change the strength of the balancer RNA secondary structure. We found that flies carrying Dbs6.40 deletions displayed decreased exon 6.41 inclusion frequencies compared to the WT (*Dscam*^{BM5}; Fig. 5B, v), consistent with the disruption of Ubs6.1-Dbs6.40 balancer pairing interactions. Moreover, the inclusion frequencies of exons 6.42 and 6.43 were lower in

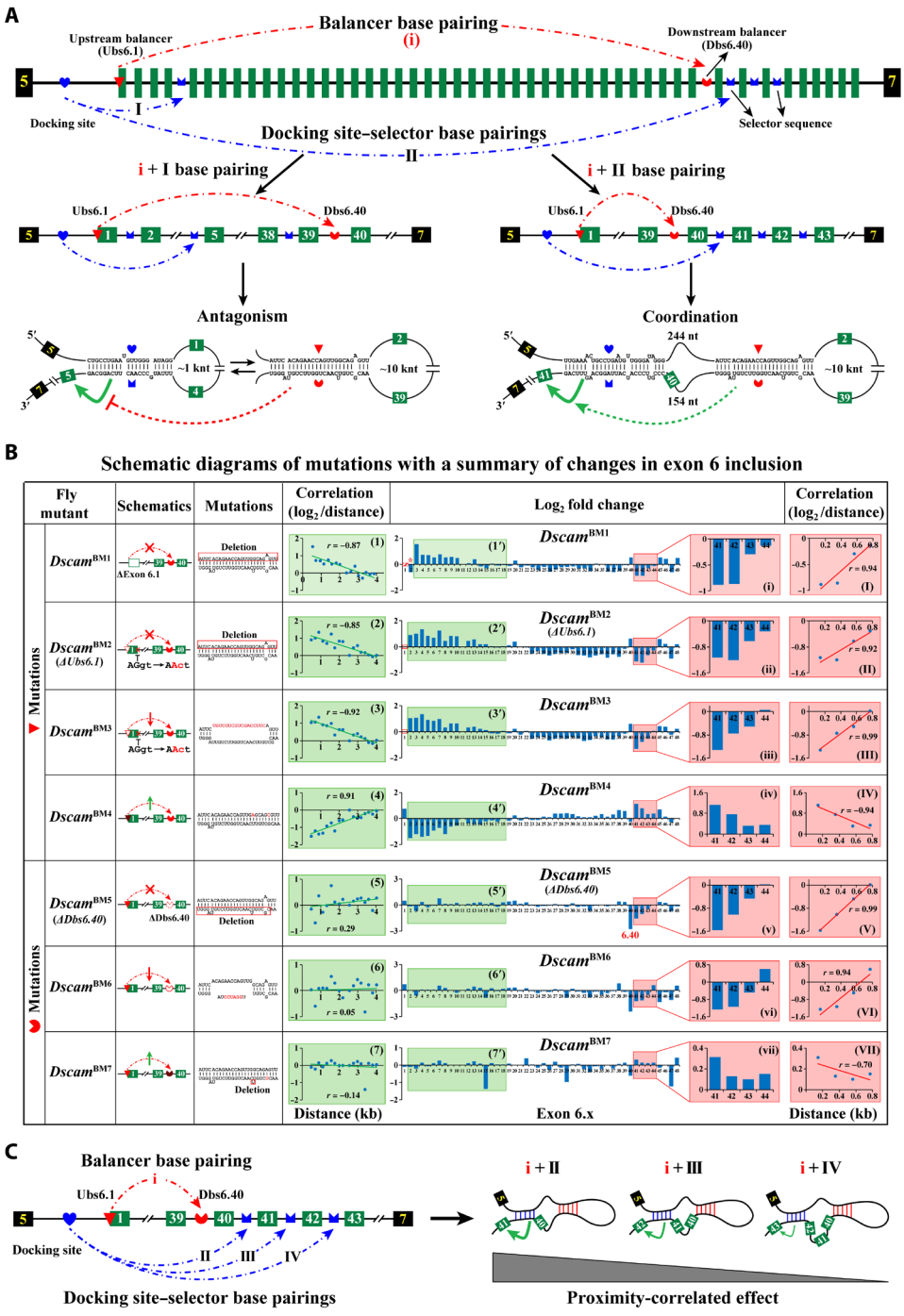

Fig. 5. Dual roles of balancer RNA secondary structure in regulating exon 6 inclusion (see also figs. S6 to S10). (**A**) Possible model of roles of balancer RNA secondary structures in mediating exon 6 inclusion. The balancer base pairing may cooperate with docking site–selector base pairing to potentially form multidomain RNA secondary structures, thereby enhancing the inclusion of distal alternative exon 6.41. Alternatively, balancer base pairing might act to inhibit the choice of proximal exon 6s by antagonizing docking site–selector base pairing. (**B**) Schematic diagrams of *Dscam1* mutant flies with a summary of the frequency of exon 6 inclusion. Red and green arrows mark the decrease and increase in pairing strength, respectively; a red cross marks complete disruption of base pairing. Predicted RNA secondary structure for the WT and a series of mutants (*Dscam*$^{BM1–BM7}$; mutated nucleotides are shown in red). The $\log_2$ fold change of the frequency of variable exon inclusion in *Dscam*$^{BM1–BM7}$ flies compared to WT is shown. The boxed area is magnified in the inset showing the effect of mutations on exon 6 inclusion. Data are from three independent experiments. * in *Dscam*^{BM1} represents abnormal exon 6.2 reduction due to the deletion of a partial 6.2 selector sequence. These data indicate that balancer RNA secondary structure inhibited the choice of proximal exon 6s but enhanced the inclusion of distal exon 6s in a proximity-dependent manner. The distance ($\log_2$/distance; left) refers to the nucleotide number from the docking site to 5′ of the target exon variant, and the distance ($\log_2$/distance; right) refers to the nucleotide number from Dbs6.40 to the downstream selector sequence. (**C**) Model of balancer base pairing could enhance the inclusion of exon 6.41 and its downstream exon 6s in a proximity-correlated manner. The thickness of the green arrows shows the strength of activation (right).

Dscam^{BM5} flies compared to the WT, suggesting that the effects of the mutations were correlated with proximity (Fig. 5B, v and V). Likewise, we observed a reduction in exon 6.41 inclusion frequency in *Dscam*^{BM6} flies that carried partial deletions in the Dbs6.40 element, albeit to a lesser degree than in *Dscam*^{BM5} flies (Fig. 5B, vi and VI). By contrast, strengthening the Ubs6.1-Dbs6.40 pairing (*Dscam*^{BM7}) increased the inclusion frequency of exons 6.41, 6.42, and 6.43 in a proximity-correlated manner (Fig. 5B, vii and VII). These data further support our proposal that balancer RNA secondary structures enhance the inclusion of multiple distal exon 6 variants in a proximity-correlated manner (Fig. 5C).

To further elucidate how balancer RNA secondary structures mediate exon 6 inclusion, we mimicked balancer base pairing by artificially inserting a Ubs-like element into the intron upstream of *Dscam1* exon 6.1 (fig. S9A). We hypothesized that the resulting mutant would potentially form a stronger two-domain RNA secondary structure (*Dscam*^{In1}; fig. S9B). We found that the frequencies of exon 6.46 and exon 6.47 inclusion were markedly increased in *Dscam*^{In1} flies (fig. S9C), which is consistent with RNA secondary structure enhancement by balancer base pairing. We obtained similar results by mimicking another balancer base pairing interaction via knock-in mutagenesis (*Dscam*^{In2}; fig. S9, A to C). The results indicate that balancer RNA secondary structures may be part of the mechanism enhancing the inclusion of distal exon 6 variants. In addition, we observed a progressive increase in the inclusion of exon 6.28 onward in *Dscam*^{In1} flies, but this trend was not obvious in *Dscam*^{In2} flies. These data show that other factors (i.e., the potential existence of other RNA structures) may contribute to the inclusion of exon 6 variants.

Balancer base pairing inhibits the selection of proximal exon 6s

We next investigated how Ubs6.1-Dbs6.40 balancer base pairings affect the choice of proximal exon 6 variants. Ubs6.1 deletion led to an overwhelming bias toward the inclusion of 5′-proximal exon 6s in *Dscam*$^{BM1–BM3}$ mutants, which carry a mutation that disrupts the Ubs6.1 balancer (Fig. 5B, 1′ to 3′, and fig. S6A). The degree of exon preference increased in a proximity-correlated manner, with the highest preference exhibited for docking site–proximal exon 6s (except for exon 6.1). To examine the relationship between inclusion frequency and proximity, we plotted the $\log_2$ fold change against the distance to the targeted exon variant, which revealed a negative linear relationship between the two variables in *Dscam*^{BM1} ($r = -0.87$; $P = 1.5 \times 10^{-5}$; Fig. 5B, 1). Likewise, significant negative correlation between the $\log_2$ fold change value and distance was also observed in *Dscam*^{BM2} and *Dscam*^{BM3} flies, in which Ubs6.1 was deleted and mutated, respectively (Fig. 5B, 2 and 3). By contrast, strengthening the Ubs6.1-Dbs6.40 pairing (*Dscam*^{BM4}) through mutations caused a notable decrease in the inclusion frequency of proximal exon 6s in a proximity-correlated manner ($r = 0.91$; Fig. 5B, 4). Together, these data indicate that Ubs6.1-mediated base pairing inhibits the selection of proximal exon 6 variants in a proximity-correlated manner.

We then performed a series of deletions in the exon 6 region (fig. S10). Similar to Δ*Exon6.1* (*Dscam*^{BM1}), we observed an obvious preference for the selection of 5′-proximal exon 6s in Δ*Exon6.1-6.6* and Δ*Exon6.1-6.16* flies (fig. S10A), which was consistent with an absence of Ubs6.1-mediated pairings. In these deletion mutants, the inclusion frequencies of proximal exon 6 variants were notably higher compared to distal exon 6s in a proximity-correlated manner. By contrast, we found no notable bias toward 5′-proximal exon 6s in mutants containing deletions in regions other than exon 6.1, such as Δ*Exon6.2-6.10*, Δ*Exon6.10-6.20*, and Δ*Exon6.21-6.30* (fig. S10B), which is consistent with the presence of Ubs6.1-mediated pairings. Together, our results indicate that Ubs6.1-mediated pairings likely inhibit the inclusion of proximal exon 6 variants.

We found a ~50% increase in the inclusion frequency of exon 6.1 in *Dscam*^{BM5} and *Dscam*^{BM6} (Fig. 5B, 5′ and 6′), consistent with disruption of the Ubs6.1-Dbs6.40 pairing interaction. However, in contrast to the three Ubs6.1 mutant lines, the Dbs6.40 mutants did not display a pronounced increase in the other 5′-proximal exon 6s (Fig. 5B, 5′ and 6′). Conversely, mutations in Dbs6.40 that strengthened the Ubs6.1-Dbs6.40 pairing (*Dscam*^{BM7}) only led to a slight reduction in the inclusion of three 5′-proximal exon 6s (Fig. 5B, 7′). This imbalance, resulting from independent Ubs6.1 and Dbs6.40 mutations, suggests that Ubs6.1 pairs competitively with multiple downstream balancer sequences, similarly to the pairings in other closely related fly species (Fig. 3A). Although mutated Dbs6.40 is unable to pair with Ubs6.1, other downstream balancer sequences might pair with Ubs6.1 to compensate for the lack of Dbs6.40. In addition, these data indicate that exon 6.40 inclusion was mediated by competition between docking site–selector 6.40 pairing and Dbs6.40-Ubs6.1 balancer pairing (fig. S7 and Supplementary Text).

In summary, given the combined data from the mutation studies and the observation that these downstream balancer sequences are mostly located in the back region of the exon 6 cluster body, we conclude that balancer RNA secondary structures function to enhance the inclusion of docking site–distal exon 6s. In contrast to the docking site–selector base pairing, which specifically activated the inclusion of the most proximal exon 6 variants outside the loop, balancer base pairing could instead enhance the inclusion of exon 6.41 and its downstream exon 6s in a proximity-correlated manner (Fig. 5C). Balancer RNA secondary structures may also repress the inclusion of proximal exon 6s, perhaps by antagonizing their docking site–selector base pairings. Thus, our data suggest that balancer RNA secondary structures orchestrate base pairing between the docking site and proximal and distal selectors to balance the stochastic choice of splice variants.

Multiple balancer base pairings confer long-range enhancement of distal exon 6 inclusion

Comparative genome analyses revealed that fly *Dscam1* evolved multiple lineage-specific and conserved balancer sequences. Unlike Dbs6.40, Dbs6.45 is conserved in most fly species whereas had degenerated in some species, such as *D. melanogaster* (Fig. 3A). We next manipulated *D. virilis* balancer sequences to examine their effects on exon 6 inclusion frequencies. Because of technical limitations in generating knock-in *D. virilis* mutants, we constructed a Dme/Dvi chimera by replacing the distal region of the exon 6 cluster of *D. melanogaster Dscam1* with the corresponding sequence from *D. virilis* (Fig. 6A).

Because both the docking site and Ubs6.1 are highly conserved between *D. melanogaster* and *D. virilis*, the predicted balancer RNA secondary structures were almost identical between the Dme/Dvi chimera and *D. virilis* (Fig. 6B and fig. S11B). High-throughput sequencing revealed that the exon 6 inclusion frequencies were largely similar between the Dme/Dvi chimera and *D. melanogaster* (fig. S11D). The Dme/Dvi chimera was therefore deemed a largely suitable tool for examining the functions of multiple balancer RNA secondary structures formed during exon 6 alternative splicing.

To investigate how various balancer RNA secondary structures influence the inclusion of distal exon 6s, we generated mutant lines containing altered Dbs6.44 or Dbs6.48 elements to change balancer

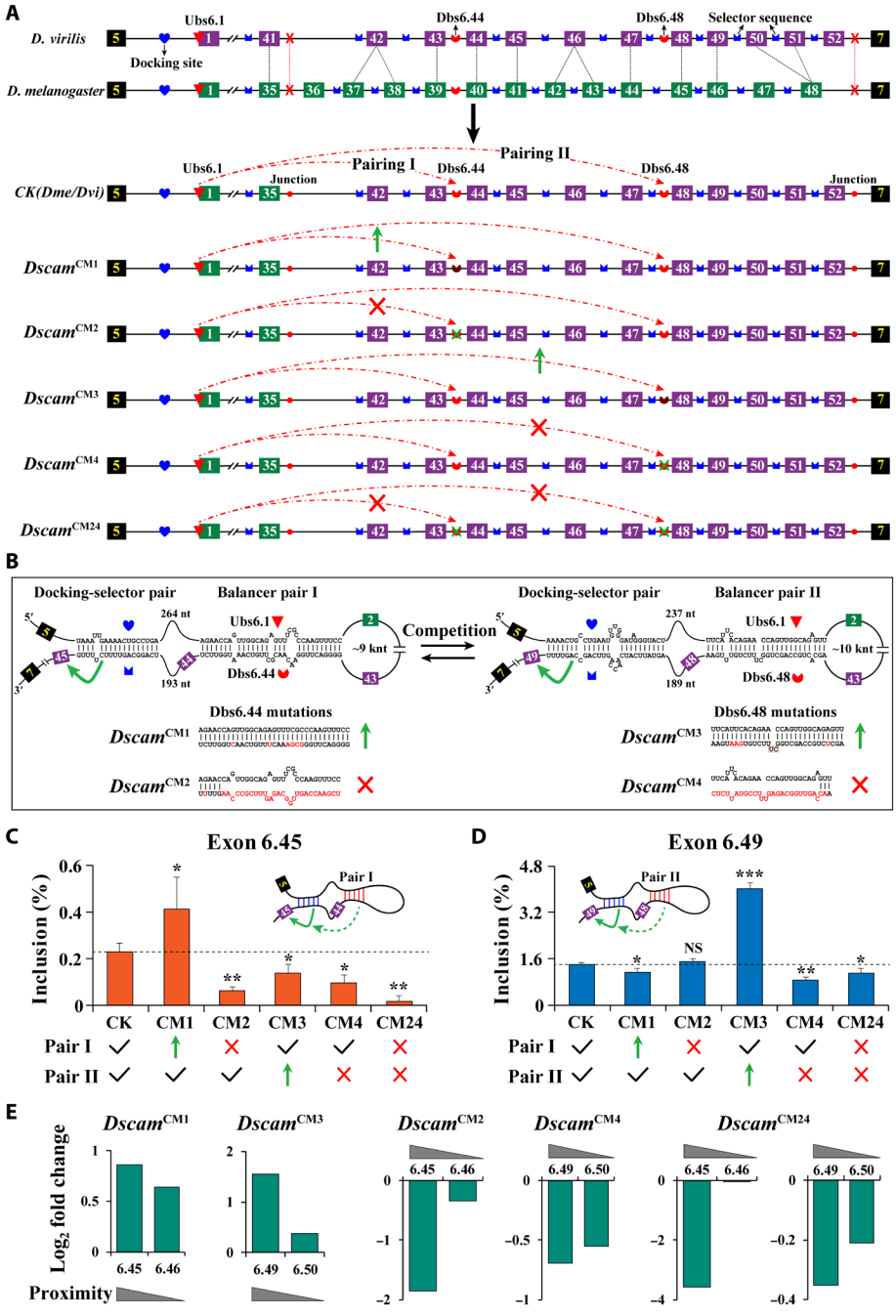

Fig. 6. Multiple balancer RNA secondary structures enhance long-range back exon 6 inclusion (see also fig. S11). (**A**) Schematic of *D. melanogaster*/*D. virilis* chimera and mutant flies. The *D. melanogaster*/*D. virilis* (Dme/Dvi) chimera was constructed by replacing the back region of exon 6 of *D. melanogaster Dscam1* with the corresponding sequence from *D. virilis*. A green arrow marks the increase in pairing strength, and a red cross marks complete disruption of base pairing. (**B**) The predicted balancer RNA secondary structures are shown in the *D. melanogaster*/*D. virilis* chimera. Mutations are introduced into double-stranded RNA to strengthen or disrupt balancer RNA pairing (*Dscam*$^{CM1–CM4}$) and double disruptive mutations of balance RNA pairing (*Dscam*^{CM24}). (**C** and **D**) Effect of balancer RNA secondary structure mutation on the inclusion of exon 6.45 (C) and exon 6.49 (D). Data are expressed as means ± SD from three independent experiments. (**E**) Balancer RNA secondary structures affect the inclusion of downstream exons in a proximity-dependent manner. The $\log_2$ fold change of the frequency of variable exon inclusion in fly mutants (*Dscam*$^{CM1–CM4}$, *Dscam*^{CM24}) compared to WT is shown. NS, not significant; **P* < 0.05; ***P* < 0.01; ****P* < 0.001 (Student's *t*-test, two-tailed).

base pairings (*Dscam*$^{CM1–CM4}$) (Fig. 6, A and B). In *Dscam*^{CM1} lines harboring mutations that strengthened Ubs6.1-Dbs6.44 pairing, the inclusion frequency of exon 6.45 markedly increased, whereas the inclusion frequency of exon 6.49 decreased (Fig. 6, C and D), which was suggestive of competitive base pairing. By contrast, *Dscam*^{CM2} flies harboring mutations that weakened the Ubs6.1-Dbs6.44 pairing exhibited markedly reduced exon 6.45 inclusion frequencies compared to the WT and slight (8.9%) but insignificant increased

exon 6.49 inclusion (Fig. 6, C and D). *Dscam*^{CM3} flies harboring strengthened Ubs6.1-Dbs6.48 pairing mutations exhibited two-fold higher exon 6.49 inclusion and significantly lower exon 6.45 inclusion (Fig. 6, C and D). Similarly, when we weakened the Ubs6.1-Dbs6.48 pairing, exon 6.49 inclusion was significantly decreased (*Dscam*^{CM4}; Fig. 6D). However, exon 6.45 inclusion was not correspondingly increased but decreased in the *Dscam*^{CM4} mutant. It is likely that the competitive effect might be overridden by other factors in this mutant. In flies containing mutations in both Dbs6.44 and Dbs6.48 (*Dscam*^{CM24}: *Dscam*$^{CM2+CM4}$), exon 6.45 inclusion frequencies were >8-fold lower compared to the WT, whereas exon 6.49 inclusion frequencies were lower but to a lesser extent. These observations provided further evidence for our hypothesis that balancer base pairings promote exon 6.45 and exon 6.49 inclusion. Moreover, the exon 6s downstream of exon 6.45 and exon 6.49 were affected in a proximity-correlated manner in all mutant flies (Fig. 6E). Overall, the data obtained by introducing independent and combined mutations demonstrated that multiple balancer RNA secondary structures act to enhance the inclusion of distal exon 6s over a long-range distance.

RNA secondary structure disruptions caused growth and neuronal defects

Why did organisms evolve these specific types of RNA secondary structures? Examining the phenotypic defects in fly mutants with disrupted RNA secondary structures may provide clues to their physiological role. To explore the phenotypic consequences of the disruption of RNA secondary structures, we performed a detailed characterization of growth and neuronal morphology of flies lacking the docking site or balancer elements. Because the vast majority (57 to 85%) of transcripts used the first exon (6.1) in the docking site deletion mutants (*Dscam*$^{\Delta Dock1\text{-}\Delta Dock3}$), we constructed a fly mutant expressing only exon 6.1 (designated *Dscam*$^{SingleEx6.1}$) as the negative control. *Dscam*$^{\Delta Dock1\text{-}\Delta Dock3}$ mutants showed a significantly decreased survival rate from embryos to adults. Moreover, disruption of balancer base pairing (*Dscam*$^{\Delta Ubs6.1}$ and *Dscam*$^{\Delta Dbs6.40}$) resulted in a modest decrease in the survival rate (Fig. 7A). Overall, the mutants with docking site deletion had a largely lower survival rate than mutants with disruption of the balancers but higher than *Dscam*$^{SingleEx6.1}$. In addition, egg production was significantly reduced in all these mutants (Fig. 7B). Because the upstream balancer element was located in the intron-exon 6.1 boundary region, it remains unclear whether these phenotypic defects observed in *Dscam1* mutants are caused by changes in the upstream balancer or exon 6.1. However, considering that deletion of a single variable exon had little effect on the phenotype, we speculate that these phenotypic defects were attributed to disruption of the balancer base pairings. These data indicate that these RNA secondary structures play an important role in normal growth and fecundity of the fly by mediating exon 6 inclusion.

Given the essential role of Dscam1 isoforms in neural development (*10*, *29*–*34*), we next focused on the morphology of three classes of neurons in fly mutants. We first examined the patterning of dendritic arborization (da) neurons in the larval epidermis of the mutants (Fig. 7C). Expression of a single Dscam1 isoform in da neurons is sufficient to ensure avoidance between self-branches, whereas thousands of isoforms are essential to distinguish between self and non–self-branches (*30*, *32*). In all mutant flies, self-dendrites of class I neurons seldom overlapped, as in the WT control (Fig. 7, C and D, and fig. S12A). However, dendrites of class I and class III neurons, which normally overlap in the WT, exhibited obvious avoidance in the *Dscam*$^{\Delta Dock3}$ mutant, although the repulsion was weaker than in the *Dscam*$^{SingleEx6.1}$ mutant only expressing exon 6.1 (Fig. 7, C and E). Dendrite overlaps of class I and class III neurons decreased as exon 6.1 inclusion increased (Fig. 7E and fig. S12C), because the dominant inclusion in exon 6.1 greatly limited access to the full diversity in these docking site deletion mutants. By contrast, the disruption of balancer base pairing did not significantly affect the coexistence between class I and class III, consistent with the splicing pattern being similar to the WT control (Fig. 7E).

We next examined the morphology of mushroom bodies (MBs) in mutant flies. The MB is the olfactory learning and memory center of *Drosophila* and is composed of ~2500 neurons (*35*). Most MB axons form two branches, extending into two orthogonal paths, thus forming two distinct lobes (*35*, *36*) (Fig. 8A); reducing Dscam1 isoform diversity resulted in only one lobe present or other defects (*30*, *37*). We observed a weak phenotype in the docking site deletion mutants (*Dscam*$^{\Delta Dock1\text{-}\Delta Dock3}$). *Dscam*$^{\Delta Dock3}$, with 85% exon 6.1 frequency, exhibited less than 10% of the phenotypic defects. The main defective phenotype included one lobe absent, one lobe thinner, or misprojection of one lobe. By contrast, almost all MBs exhibited defective morphology in the *Dscam*$^{SingleEx6.1}$ mutant only expressing exon 6.1 (Fig. 8, A and B). In addition, no obvious defective phenotype was observed in *Dscam*$^{\Delta Ubs6.1}$ and *Dscam*$^{\Delta Dbs6.40}$ mutants that disrupted the balancer base pairing (Fig. 8B). Overall, these data indicate that the MB phenotype is not sensitive to alteration of exon 6 inclusion caused by disruption of these base pairings.

Last, we assessed the axonal branching pattern of mechanosensory (MS) neurons resulting from the disruption of RNA secondary structures. MS neurons exhibited a stereotypic axon branching pattern in the ventral nerve cord (*38*, *39*); previous studies have shown that Dscam1 isoform diversity is essential for controlling axonal branching (*30*, *38*, *39*). We found that *Dscam*$^{\Delta Dock1\text{-}\Delta Dock3}$ flies with deletion of the docking site displayed marked defects in the axonal branching patterns of posterior scutellar neurons (Fig. 8C). The axons branched substantially less in *Dscam*$^{\Delta Dock1\text{-}\Delta Dock3}$ mutants than in the WT, and their lateroanterior branches were much shorter than those of the WT (Fig. 8D). These phenotypic defects were exacerbated with increased exon 6.1 inclusion. In addition, disruption of balancer RNA secondary structures significantly reduced axon branches in *Dscam*$^{\Delta Ubs6.1}$ and modestly but insignificantly in *Dscam*$^{\Delta Dbs6.40}$. However, the length of the lateroanterior branch was significantly reduced in both *Dscam*$^{\Delta Ubs6.1}$ and *Dscam*$^{\Delta Dbs6.40}$ mutants (Fig. 8D). These results demonstrate a role for these RNA secondary structures in axonal branching of posterior scutellar neurons.

In summary, these data demonstrate that these RNA secondary structures, which mediate exon 6 inclusion, are required for normal fly growth and development in a context-dependent manner. Overall, deletion of the docking site caused more severe defective phenotypes than disruption of balancer base pairing, consistent with its stronger alteration of exon 6 inclusion.

DISCUSSION

In this study, we uncovered a set of hidden RNA secondary structures that counteract the first-come, first-served splicing principle to balance the stochastic choice of *Dscam1* splice variants. We found evolutionarily conserved RNA secondary structures imbedded within exon 6 clusters that were distinct from docking site–selector RNA

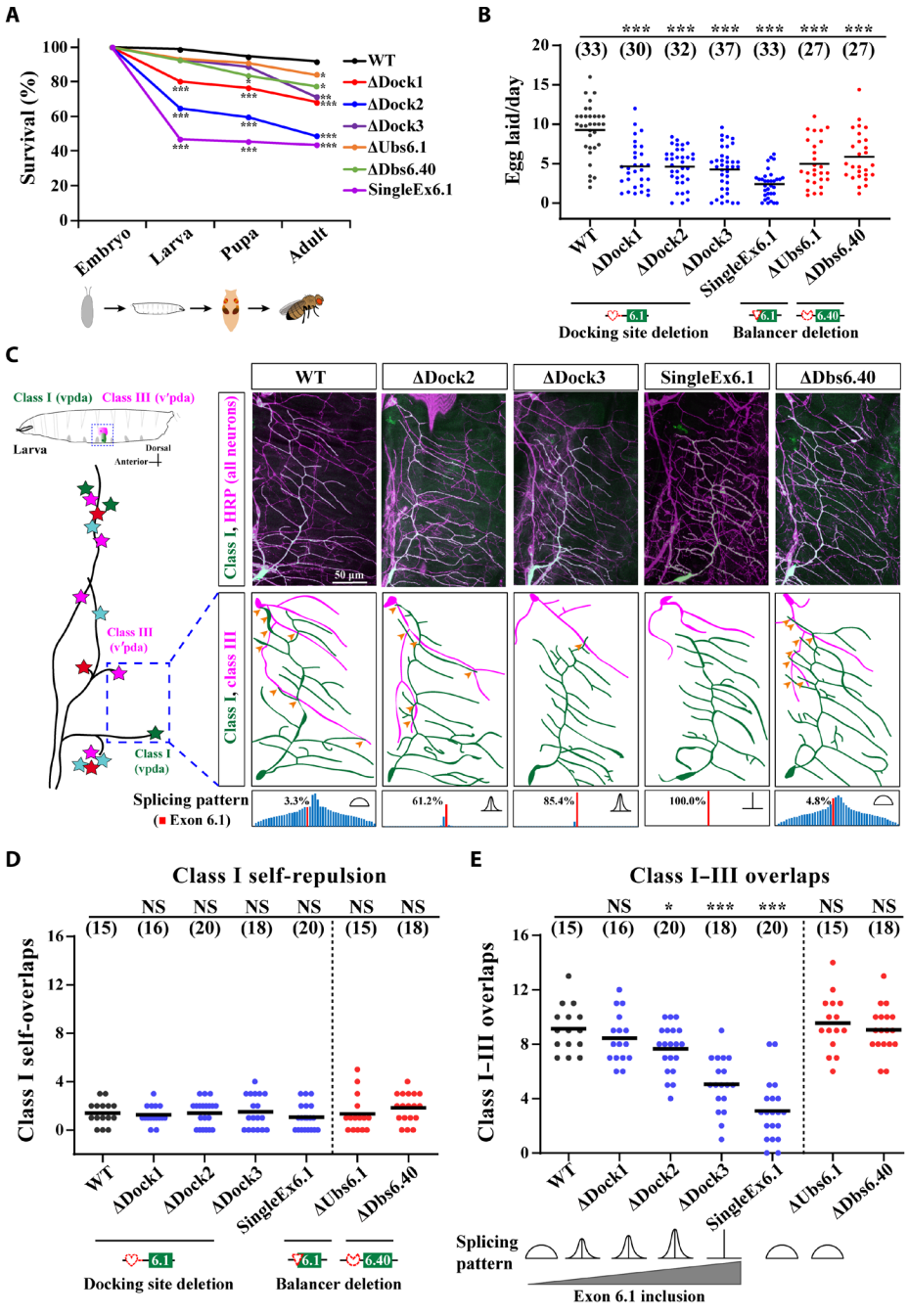

Fig. 7. RNA secondary structures are required for *Drosophila* development (see also fig. S12). (**A**) Reduced survival rates of different *Dscam1* mutant flies at different developmental stages compared with the WT control. (**B**) Disruption of RNA secondary structure caused reduced fecundity in different mutant flies. (**C**) Representative images of dendrites of dendritic arborization (da) neurons in WT and *Dscam1* mutants. Additional images are shown in fig. S12. Schematic diagram of the distribution of four types of da neurons (colored stars) is shown on the left. All neurons were visualized with anti–horseradish peroxidase (HRP) antibody (magenta), and class I (vpda) neurons were labeled with green fluorescent protein (GFP) (green; appears white because they overlap with magenta). Yellow arrowheads indicate crossing between class I and class III dendrites. Scale bar, 50 μm. The splicing pattern of *Dscam1* variable exon 6 of WT and mutants is shown at the bottom, and the inclusion of exon 6.1 is marked in red. (**D**) Self-repulsion was intact in all mutants similar to the WT control. Numbers in parentheses refer to the analyzed class1 neurons of each genotype. (**E**) Overlaps between class I and class III dendrites of different *Dscam1* mutant flies. NS, not significant; **P* < 0.05; ***P* < 0.01; ****P* < 0.001 (Student's *t* test, two-tailed).

secondary structures. At first glance, these previously unknown balancer RNA secondary structures appeared to be similar to the docking site and selector sequence base pairing structures. Both structures were composed of two types of conserved elements: an upstream element, which was located in the intron downstream of the constitutive exon 5, and multiple downstream elements, which were located in the exon 6 cluster body. The upstream element could pair competitively with various downstream elements.

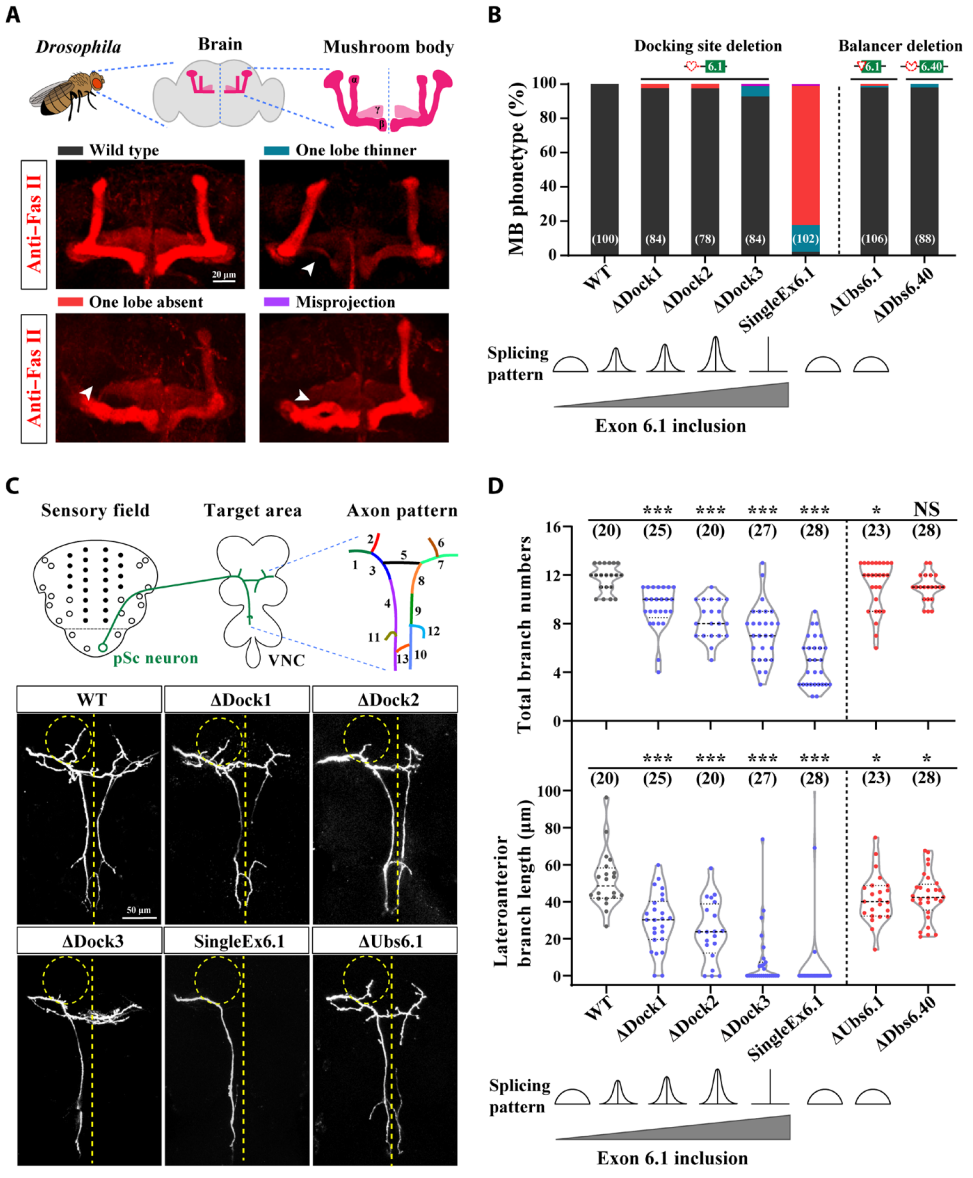

Fig. 8. Disruption of RNA secondary structure caused MB and posterior scutellar neuron defects (see also fig. S12). (**A**) Schematics of MB morphology in the fly brain and representative images of MB lobe morphological defects. Scale bar, 20 μm. (**B**) Quantification of MB defective phenotypes in mutant flies. (**C**) Schematic of axon branching pattern of the WT posterior scutellar (pSc) neuron in ventral nerve cord (VNC) and representative dye tracing images of WT and mutant flies. Dashed circles indicate the location of the lateroanterior branch. (**D**) Quantitative analysis of total branch number and length of the lateroanterior branch in WT and mutant flies. The docking site deletion mutants ($Dscam^{\Delta Dock1\text{-}\Delta Dock3}$) showed a significant decrease in the number of axon branches and the length of the lateroanterior branch. In addition, disruption of balancer RNA secondary structures ($Dscam^{\Delta Ubs6.1}$ and $Dscam^{\Delta Dbs6.40}$) reduced the number of axon branches and the length of the lateroanterior branch, albeit to a lesser degree than docking site deletion. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ (Student's *t* test, two-tailed).

However, the two types of competing RNA secondary structures were fundamentally different in both their function and mechanism. In contrast to the base pairing between the docking site and various exon 6 upstream selectors, the balancer RNA secondary structures tended to be located in the central and 3′ regions of the exon 6 cluster. Moreover, the balancer RNA secondary structures themselves did not promote the inclusion of exon 6 variants but rather enhanced the inclusion of distal exon 6s by strengthening docking site–selector base pairing or inhibited the inclusion of proximal exon 6 s by antagonizing docking site–selector base pairing. Our data suggest that the regulatory function of balancer RNA secondary structures was dependent on the base pairing between the docking site and selector sequence. Furthermore, the formation of balancer RNA secondary structures enhanced the inclusion of multiple exon 6s in a proximity-dependent manner. For example, the Ubs6.1-Dbs6.40 RNA secondary structure may act as a fulcrum that guides the docking site to pair with the selector upstream of exon 6.41 or 6.42 (Fig. 5C). Thus, base pairing between the docking site and a selector sequence may activate selection of the most proximal exon variant 6, whereas the balancer sites

are thought to be effective over a long range. We concluded that while docking site–selector base pairing interactions determined which exon 6 variant was selected, the balancer base pairings drove the stochastic choice of exon 6 variants.

Our study provides a reasonable explanation for the long-standing puzzle in the choice of exon 6 variants of *Dscam1*. By combining various techniques including informatics, evolutionary analyses, and in vivo mutagenesis experiments, we generated a new model depicting the regulation of exon 6 variant inclusion. According to our model (Fig. 9A), for a 5′-distal exon 6 (i.e., exon 6.z) to be included in mature *Dscam1* mRNA, the docking site–selector interaction may cooperate with multiple balancer base pairings to form a stronger multidomain RNA secondary structure (Fig. 9A, right); balancer base pairings stabilize the docking site–selector interaction, thereby increasing the frequency of exon 6.z inclusion. For a 5′-proximal exon 6 (i.e., exon 6.x) to be included in mature mRNA, the docking site may pair with the selector sequence to form short RNA secondary structures containing few domains. Therefore, 5′-proximal exon 6 inclusion relies upon docking site–selector base pairing strength. Simultaneously, balancer secondary structures also inhibit the inclusion of 5′-proximal exon 6s by antagonizing their docking site–selector base pairing interactions (Fig. 9A, left). For a middle exon 6 to be included in the mature mRNA, strong intermediate RNA secondary structures may be formed (Fig. 9A, middle). Although docking site–selector pairings did not exhibit an obvious preference for docking site–distal exon 6s, the combined overall base pairings exhibited a low-to-high strength gradient from docking site–proximal to docking site–distal exon 6s (Fig. 9B). In this study, we found that docking site deletion led to a bias toward the inclusion of docking site–distal exon 6s (except for exon 6.1), whereas deleting upstream balancers led to increased docking site–proximal exon 6 inclusion (fig. S13). Therefore, we speculate that the gradient of pairing strength

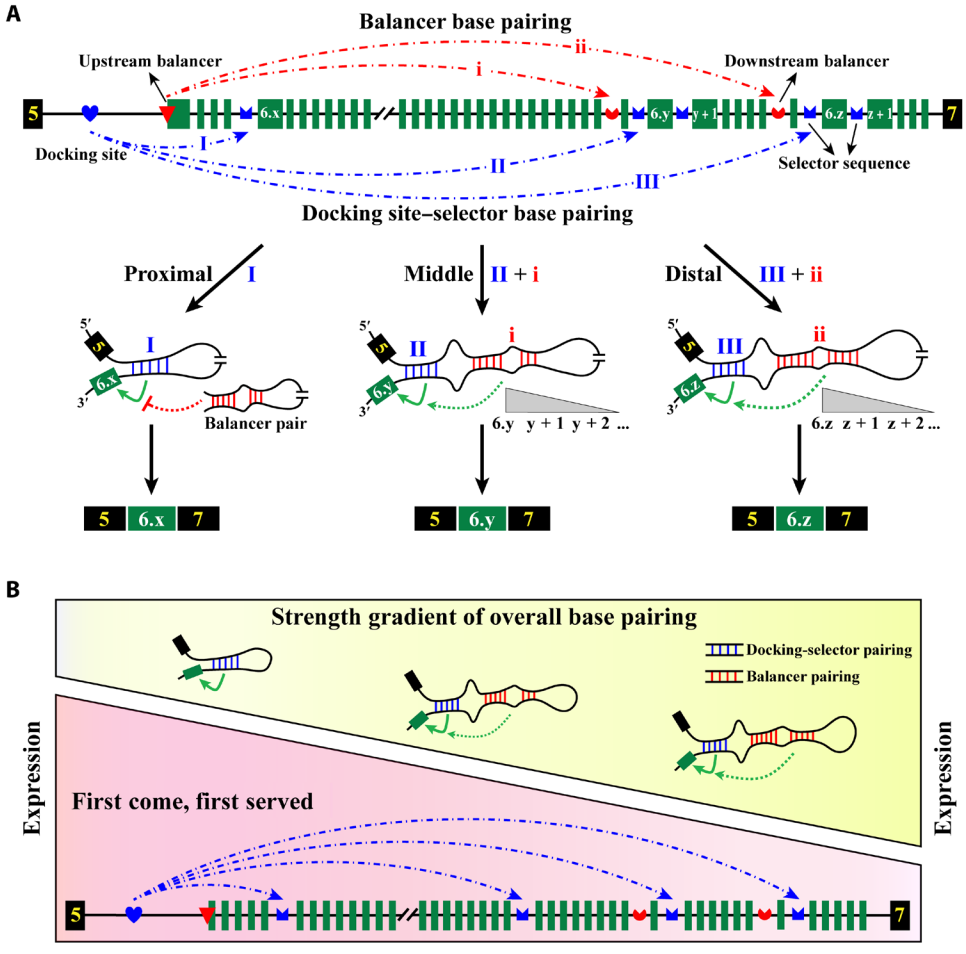

Fig. 9. Model for balancer RNA secondary structures in balancing the stochastic choice of *Dscam1* splice variants (see also fig. S13). (**A**) Model of balancer RNA secondary structures in regulating exon 6 inclusion. Symbols used are the same as those in Fig. 1. For a proximal exon 6 (i.e., exon 6.x) to be included in the mature mRNA, the docking site would pair with the selector sequence to form the short RNA secondary structures (left). Simultaneously, balancer secondary structures also inhibit the inclusion of proximal exon 6s by antagonizing docking site–selector base pairings. For a distal exon 6 (i.e., exon 6.z) to be included in the mature mRNA, the docking site–selector interaction cooperates with the balancer base pairings to form a multidomain RNA secondary structure, thereby activating exon 6.z inclusion at higher efficiency (right). For a middle exon 6 to be included in the mature mRNA, strong intermediate RNA secondary structures may be formed (middle). Balancer RNA secondary structures enhance the inclusion of exon 6.z and its downstream exons in a proximity-correlated manner. (**B**) Balancer RNA secondary structure compensation counteracts the first-come, first-served splicing principle to balance stochastic choice of *Dscam1* splice variants. Although docking site–selector pairings did not exhibit an obvious preference for docking site–distal exon 6s, the combined overall base pairings exhibit a low-to-high strength gradient from proximal to distal exons.

compensates for the docking site–proximal to docking site–distal preference of the exon 6 variants caused by the first-come, first-served principle. Thus, the final splicing outcome did not exhibit an obvious docking site–proximal preference in exon 6 variants. This finding suggests that flies have evolved an intricate compensatory pairing mechanism that drives the stochastic choice of *Dscam1* splice variants.

The conservation of balancer RNA secondary structures and genetic analyses highlight their regulatory benefits and physiological roles. First, it would be advantageous to increase balancer base pairing interactions alongside the duplication of exon clusters. Moreover, balancer pairing compensation could provide specific advantages to large exon clusters (i.e., *Dscam1*) through a variety of mechanisms. An intriguing possibility is that the balancer base pairing interaction induces the juxtaposition of the docking site and its downstream selector sequences, thereby enhancing their pairing interaction. For example, Ubs6.1-Dbs6.40 base pairing acts as a fulcrum that guides the docking site to pair with multiple selector sequences (Fig. 5C). Balancer base pairing interactions not only reduce the base pairing distance across large RNA molecules but also help with avoiding the formation of nonspecific secondary structures and heterogeneous nuclear ribonucleoprotein complexes that might affect the accuracy or efficiency of specific docking site–selector base pairings. Thus, balancer base pairing strategies may have been evolutionarily favorable in complex alternative splicing units such as *Dscam1*, in which long-distance base pairing is required. Coordination and competition between diverse balancer and docking site–selector base pairing interactions have increased the flexibility and efficiency of folding-mediated splicing. Our phenotypic analyses revealed obvious defects in these mutants bearing disruption of two types of RNA secondary structures (Figs. 7 and 8), suggesting that two types of RNA secondary structures coordinate to ensure the proper frequency of exon 6 variants, which is required for normal development.

Is this type of functional RNA secondary structures common in vivo? Competing RNA secondary structures were initially identified in the exon 6 cluster of *Dscam1* (*14*), which at that time was believed to be unique. Similar structural codes have recently been revealed in several exon clusters, such as *Drosophila* 14-3-3ξ, exon 4 and 9 clusters in *Drosophila Dscam1*, *srp*, *RIC-3*, *Branchiostoma MRP*, human *dynamin 1*, and *CD55* (*18*, *20*, *40*–*43*). In addition, this docking site–mediated model governs the regulation of 3′-end alternative splicing in *Drosophila PGRP-LC*, *CG42235*, and *Pip* genes (*44*). Therefore, competing RNA secondary structures may be a broadly applicable mechanism to regulate mutually exclusive splicing. In addition to mutually exclusive exons, competing RNA secondary structures may regulate variable exon skipping of human *SF1* (*45*) and alternative backsplicing of human *POLR2A* (*46*). Moreover, we observed similar balancer RNA secondary structures in the *Dscam1* exon 9 cluster and *MRP1* (figs. S14 and S15). Analogous nested RNA secondary structures have been found in the catalog of pairs of conserved complementary regions (PCCRs) in human protein-coding genes (*47*). Likewise, two groups of RNA structural modules have been shown to operate together to dynamically regulate alternative splicing in the human *Ate1* gene (*48*). It is therefore likely that the framework of balancer RNA secondary structures developed here is widely applicable to the regulation of mutually exclusive splicing and other RNA processing events.

Moreover, some questions related to the functional regulation of RNA structures should be further addressed. First, why are these exons in the stem-loop not spliced compared to flanking exons? Previous studies have shown that RNA pairing can repress the splicing of exons within a loop (*18*, *49*). Moreover, *Dscam1* exon 6 is kept silent through weak splice sites (*22*) in combination with splicing repressors, such as the heterogeneous nuclear ribonucleoprotein hrp36 (*21*). When the docking site pairs with certain selector sequence, the distant enhancers specifically activate the proximal alternative exon by promoting recognition of the splicing site and/or dissociation of the repressors, while other exons in a stem-loop and downstream exons away from the RNA loop are repressed (*8*, *14*, *17*). Second, can the balancer and docking site–selector base pairings form a long-range pseudoknot? It is likely that two sets of RNA secondary structures form a long-range pseudoknot. However, the mutational evidence in the present study supports the notion that balancer secondary structures act to antagonize docking site–selector base pairing of proximal exon 6s (Fig. 5). Therefore, we propose that these elements function by forming RNA secondary structures rather than pseudoknot structures.

We have provided sufficient evidence for the splicing regulation of docking site–selector and balancer RNA secondary structures in the *Dscam1* gene. Future challenges include directly demonstrating the existence of the purported and/or unknown base pairs and mapping RNA-protein interactions and the effects on the choice of different exons. It will also be important to test whether these variable exons are spliced cotranscriptionally and, if so, how far the RNA polymerase progresses before splicing is detected.

MATERIALS AND METHODS

Materials: Fly stocks

The *{nos-Cas9}attP2* fly was used for embryo microinjection in which Cas9 can be specifically expressed in germ cells (*50*). To avoid expression differences between fly strains, we used *{nos-Cas9}attP2* fly as the WT. *if/CyO* was used as the balancer stock for mutant fly screening. *D. virilis* line was a gift from Q. Zhou's laboratory (Zhejiang University, China). The 221-GFP Gal4 line was used to drive green fluorescent protein (GFP) expression in class I da neurons as previously described (*51*). All flies were cultured on standard cornmeal medium at 25°C.

Generation of *Dscam1* mutant fly lines

To generate precise point mutations or other mutations in the fly *Dscam1* locus, we used the CRISPR-Cas9–mediated homology-directed repair system (*50*, *52*) to perform site-targeted mutagenesis. In this system, two single guide RNA (sgRNA) plasmids and one donor plasmid were coinjected into *{nos-Cas9}attP2* embryos (Fig. 1B). The donor plasmid included an upstream homologous arm (approximately 1 kb), mutant sequences, and a downstream homologous arm (approximately 1 kb). Injection was performed at UniHuaii Co. Ltd., China. After injection, each mosaic fly (G0 generation) was individually crossed with the balancer stock. The female offspring (G1 generation) were prescreened by PCR or PCR product restriction digestion analysis after the genome was extracted. Next, individual male flies (G1 generation) in tubes where female flies have mutations would be crossed with the balancers again. Similarly, single male flies (G_1 generation) were screened by PCR or PCR product restriction digestion analysis after the genome was extracted. Last, homozygous viable mutant flies were obtained through self-crossing of heterozygous mutants. Primers used for sgRNA, mutant screening, and RT-PCR are listed in table S2. All mutation sequences are summarized in table S3.

Identification of *Dscam1* exon 6 orthologs

Genomic sequences of the *Dscam1* gene in insects were obtained by comparative genome analysis from FlyBase and NCBI (http://flybase.org/blast/ and https://blast.ncbi.nlm.nih.gov/Blast.cgi) (table S1). Sequences of the *Dscam1* exon 6 cluster were determined by PCR and sequencing after genomic DNA of *{nos-Cas9)attP2* fly was extracted.

Sequence alignments and RNA pairing predictions

We aligned the selector or balancer sequences among species using ClustalW2 (www.ebi.ac.uk/clustalw/index.html) (*53*). The consensus sequences of the balancer sequences were obtained using WebLogo (http://weblogo.berkeley.edu/logo.cgi) (*54*). To decipher the evolutionary relationship among the variable exons of *Dscam1*, nucleotide sequences of the exon variants were aligned using the MUSCLE program according to the system default parameters. A phylogenetic tree was constructed using MEGA X (www.megasoftware.net/). The tree was reconstructed using a nearest-neighbor interchange method based on the maximum composite likelihood model. The RNA secondary structures were predicted using the Mfold Web Server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form) (*55*). The scores of 5′ and 3′ splice sites were calculated using the splice site predictor (www.fruitfly.org/seq_tools/splice.html) (*56*).

Reverse transcription polymerase chain reaction

Total RNA from fly heads or other tissues was harvested using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Total RNA was reverse-transcribed using the SuperScript III System (Invitrogen) with specific primers for *Dscam1* constitutive exon 7. The reverse transcription products were amplified with PrimeSTAR DNA Polymerase (TaKaRa) using specific primers for *Dscam1* constitutive exon 5 and exon 7, and the final products were detected via 1.5% agarose gel electrophoresis. PCR was performed as follows: denaturing at 98°C for 5 min, 35 cycles of denaturing at 98°C for 30 s, annealing at 58°C for 30 s, extension at 72°C for 30 s, and a final elongation step of 10 min at 72°C. Electrophoresis of the RT-PCR product was photographed using the GIS 1D Gel Image System (version 3.73; Tanon, Shanghai, China), and ImageJ software was used for semi-quantitative densitometric analysis of each band.

Utilization assay of exon 6 variants

Multiplex high-throughput sequencing was used to identify the relative usage of the variable exon 6s as previously described (May 2011). Following reverse transcription, high-throughput sequencing primers (table S2) were designed to amplify the exon 6 cluster using PrimeSTAR DNA Polymerase (TaKaRa). Amplification was performed as follows: denaturing at 98°C for 5 min, 25 cycles of denaturing at 98°C for 30 s, annealing at 58°C for 30 s, extension at 72°C for 30 s, and a final elongation step of 10 min at 72°C. The RT-PCR products were examined via 1.5% agarose gel electrophoresis. The products were excised and gel-purified using the Biospin Gel Extraction Kit (BioFlux). Purified PCR products were pooled on the Illumina MiSeq platform (Illumina, San Diego, CA, USA) according to the standard protocol by G-BIO. The data from the high throughput sequencing analysis were shown in table S4.

Growth and development detection

Approximately 120 virgin female and male fruit flies were collected. All flies were mixed for 48 hours in a bottle containing a solidified juice tray and yeast extract. The juice tray was replaced with a new one every 4 hours three times. A total of 200 embryos were collected from each replacement juice tray and placed in a new, larger juice tray. After 48 hours, the number of hatched embryos was counted to obtain the hatching rate.

Each group of 90 larvae was collected from the second-instar larvae, hatched, and distributed into three new food tubes, representing three groups in total. After 3 to 4 days, the number of pupae on the tube wall was counted, and the pupation rate was calculated. After another 4 to 5 days, the number of adult flies in the food tube was counted, and the eclosion rate was calculated.

Fecundity detection

Approximately 40 virgin female and male flies were collected. Then, a female and male fly were placed in the same fresh food tube. The fruit flies were transferred to fresh food tubes every day, and the number of eggs laid each day by each pair was counted. Fecundity is expressed as the number of eggs laid per tube per day. The experimental group in which the fruit flies died during the analysis was excluded.

Immunostaining

For coexistence between class I and class III da neurons, the third-instar larvae of GFP-labeled class I da neurons were dissected, and the larval epidermis was fixed with 4% paraformaldehyde for 25 min at room temperature followed by three washes with phosphate-buffered saline containing 0.1% Triton X-100 (PBST) for 20 min each. Then, the larval epidermis was blocked with 5% bovine serum albumin (BSA) (diluted in PBST) and incubated with Cy3-conjugated Affinipure goat anti–horseradish peroxidase (1:200) at 4°C overnight. After three washes with PBST for 20 min each, samples were mounted in prolong gold antifade reagent (Invitrogen). Immunofluorescence staining was imaged using a confocal microscope LSM800 (Carl Zeiss). The class I (vpda) neuron self-branch images were exported in a single channel using software.

For MB immunofluorescence staining, adult flies at 3 to 5 days were dissected in PBS, and the brains were fixed in 4% paraformaldehyde at room temperature for 45 min. Following three washes with PBST for 20 min each, the brains were blocked with 5% BSA for 1 hour at room temperature. After three washes with PBST for 20 min each, the brains were incubated for 48 hours at 4°C with the primary antibody [anti–Fas II, Developmental Studies Hybridoma Bank; diluted 1:2 in PBST]. After standard washing, the brains were incubated for 4 hours with the appropriate secondary antibody (Alexa 594–goat anti-mouse immunoglobulin G, Earthox; diluted 1:400 in PBST) at room temperature. After standard washing, the brains were mounted in prolong gold antifade reagent (Invitrogen). Last, MB immunofluorescence staining was imaged with a laser scanning confocal microscope (LSM800, Carl Zeiss). The branching patterns of pSc neurons were visualized by 1,1′-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate or 1,1'-Dioctadecyl-3,3,3',3'-Tetramethylindodicarbocyanine,4-Chlorobenzenesulfonate Salt tracing essentially as previously described (*38*, *39*, *57*).

Statistical analysis

Quantitative analysis was performed from three biological replicates. Error bars were calculated as the average of three independent experiments. The significance of the difference was determined by the two-tailed Student's *t* test, and $P < 0.05$, $P < 0.01$, and $P < 0.001$ were used to indicate different levels of statistical significance. The correlation between the percent inclusion frequency and the strength or

distance of the long-range RNA secondary structure was fitted with a two-variable linear regression, with the $\log_2$ fold change or exon variant inclusion frequency as the dependent variable and the distance or strength of the base pairing between the docking site and selector sequence as the independent variable. Statistical analysis was performed using Origin 2018 software (OriginLab Corporation, Northampton, MA, USA).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abm1763

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. Y. Lee, D. C. Rio, Mechanisms and regulation of alternative pre-mRNA splicing. *Annu. Rev. Biochem.* **84**, 291–323 (2015).
2. T. W. Nilsen, B. R. Graveley, Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**, 457–463 (2010).
3. J. Ule, B. J. Blencowe, Alternative splicing regulatory networks: Functions, mechanisms, and evolution. *Mol. Cell* **76**, 329–345 (2019).
4. Q. Pan, O. Shai, L. J. Lee, B. J. Frey, B. J. Blencowe, Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
5. M. Sultan, M. H. Schulz, H. Richard, A. Magen, A. Klingenhoff, M. Scherf, M. Seifert, T. Borodina, A. Soldatov, D. Parkhomchuk, D. Schmidt, S. O'Keeffe, S. Haas, M. Vingron, H. Lehrach, M. L. Yaspo, A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321**, 956–960 (2008).
6. S. C. Bonnal, I. Lopez-Oreja, J. Valcarcel, Roles and mechanisms of alternative splicing in cancer - implications for care. *Nat. Rev. Clin. Oncol.* **17**, 457–474 (2020).
7. Y. Wang, Y. Bao, S. Zhang, Z. Wang, Splicing dysregulation in cancer: From mechanistic understanding to a new class of therapeutic targets. *Sci. China Life Sci.* **63**, 469–484 (2020).
8. Y. Jin, H. Dong, Y. Shi, L. Bian, Mutually exclusive alternative splicing of pre-mRNAs. *Wiley Interdiscip. Rev. RNA* **9**, e1468 (2018).
9. C. W. Smith, Alternative splicing—When two's a crowd. *Cell* **123**, 1–3 (2005).
10. D. Schmucker, J. C. Clemens, H. Shu, C. A. Worby, J. Xiao, M. Muda, J. E. Dixon, S. L. Zipursky, Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* **101**, 671–684 (2000).
11. Y. Jin, H. Li, Revisiting Dscam diversity: Lessons from clustered protocadherins. *Cell. Mol. Life Sci.* **76**, 667–680 (2019).
12. F. L. Watson, R. Puttmann-Holgado, F. Thomas, D. L. Lamar, M. Hughes, M. Kondo, V. I. Rebel, D. Schmucker, Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* **309**, 1874–1878 (2005).
13. S. L. Zipursky, J. R. Sanes, Chemoaffinity revisited: Dscams, protocadherins, and neural circuit assembly. *Cell* **143**, 343–353 (2010).
14. B. R. Graveley, Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. *Cell* **123**, 65–73 (2005).
15. D. Anastassiou, H. Liu, V. Varadan, Variable window binding for mutually exclusive alternative splicing. *Genome Biol.* **7**, R2 (2006).
16. B. Xu, Y. Meng, Y. Jin, RNA structures in alternative splicing and back-splicing. *WIREs RNA* **12**, e1626 (2020).
17. B. Xu, Y. Shi, Y. Wu, Y. Meng, Y. Jin, Role of RNA secondary structures in regulating Dscam alternative splicing. *Biochim. Biophys. Acta Gene Regul. Mech.* **1862**, 194381 (2019).
18. Y. Yang, L. Zhan, W. Zhang, F. Sun, W. Wang, N. Tian, J. Bi, H. Wang, D. Shi, Y. Jiang, Y. Zhang, Y. Jin, RNA secondary structure in mutually exclusive splicing. *Nat. Struct. Mol. Biol.* **18**, 159–168 (2011).
19. W. Hong, Y. Shi, B. Xu, Y. Jin, RNA secondary structures in Dscam1 mutually exclusive splicing: Unique evolutionary signature from the midge. *RNA* **26**, 1086–1093 (2020).
20. Y. Yue, Y. Yang, L. Dai, G. Cao, R. Chen, W. Hong, B. Liu, Y. Shi, Y. Meng, F. Shi, M. Xiao, Y. Jin, Long-range RNA pairings contribute to mutually exclusive splicing. *RNA* **22**, 96–110 (2016).
21. S. Olson, M. Blanchette, J. Park, Y. Savva, G. W. Yeo, J. M. Yeakley, D. C. Rio, B. R. Graveley, A regulator of Dscam mutually exclusive splicing fidelity. *Nat. Struct. Mol. Biol.* **14**, 1134–1140 (2007).
22. X. Wang, G. Li, Y. Yang, W. Wang, W. Zhang, H. Pan, P. Zhang, Y. Yue, H. Lin, B. Liu, J. Bi, F. Shi, J. Mao, Y. Meng, L. Zhan, Y. Jin, An RNA architectural locus control region involved in Dscam mutually exclusive splicing. *Nat. Commun.* **3**, 1255 (2012).
23. R. Zeller, J. Deschamps, first come, first served. *Nature* **420**, 138–139 (2002).
24. G. E. May, S. Olson, C. J. McManus, B. R. Graveley, Competing RNA secondary structures are required for mutually exclusive splicing of the Dscam exon 6 cluster. *RNA* **17**, 222–229 (2011).
25. G. Neves, J. Zucker, M. Daly, A. Chess, Stochastic yet biased expression of multiple Dscam splice variants by individual cells. *Nat. Genet.* **36**, 240–246 (2004).
26. W. Sun, X. You, A. Gogol-Doring, H. He, Y. Kise, M. Sohn, T. Chen, A. Klebes, D. Schmucker, W. Chen, Ultra-deep profiling of alternatively spliced Drosophila Dscam isoforms by circularization-assisted multi-segment sequencing. *EMBO J.* **32**, 2029–2038 (2013).
27. C. Lee, N. Kim, M. Roy, B. R. Graveley, Massive expansions of Dscam splicing diversity via staggered homologous recombination during arthropod evolution. *RNA* **16**, 91–105 (2010).
28. T. M. Ivanov, D. D. Pervouchine, An evolutionary mechanism for the generation of competing RNA structures associated with mutually exclusive exons. *Genes (Basel)* **9**, 356 (2018).
29. X. L. Zhan, J. C. Clemens, G. Neves, D. Hattori, J. J. Flanagan, T. Hummel, M. L. Vasconcelos, A. Chess, S. L. Zipursky, Analysis of Dscam diversity in regulating axon guidance in Drosophila mushroom bodies. *Neuron* **43**, 673–686 (2004).
30. D. Hattori, Y. Chen, B. J. Matthews, L. Salwinski, C. Sabatti, W. B. Grueber, S. L. Zipursky, Robust discrimination between self and non-self neurites requires thousands of Dscam1 isoforms. *Nature* **461**, 644–648 (2009).
31. M. E. Hughes, R. Bortnick, A. Tsubouchi, P. Baumer, M. Kondo, T. Uemura, D. Schmucker, Homophilic Dscam interactions control complex dendrite morphogenesis. *Neuron* **54**, 417–427 (2007).
32. B. J. Matthews, M. E. Kim, J. J. Flanagan, D. Hattori, J. C. Clemens, S. L. Zipursky, W. B. Grueber, Dendrite self-avoidance is controlled by Dscam. *Cell* **129**, 593–604 (2007).
33. J. Wang, C. T. Zugates, I. H. Liang, C.-H. Lee, T. Lee, Drosophila Dscam is required for divergent segregation of sister branches and suppresses ectopic bifurcation of axons. *Neuron* **33**, 559–571 (2002).
34. P. Soba, S. Zhu, K. Emoto, S. Younger, S. J. Yang, H. H. Yu, T. Lee, L. Y. Jan, Y. N. Jan, Drosophila sensory neurons require Dscam for dendritic self-avoidance and proper dendritic field organization. *Neuron* **54**, 403–416 (2007).
35. K. Ito, W. Awano, K. Suzuki, Y. Hiromi, D. Yamamoto, The Drosophila mushroom body is a quadruple structure of clonal units each of which contains a virtually identical set of neurones and glial cells. *Development* **124**, 761–771 (1997).
36. T. Lee, A. Lee, L. Luo, Development of the Drosophila mushroom bodies: Sequential generation of three distinct types of neurons from a neuroblast. *Development* **126**, 4065–4076 (1999).
37. D. Hattori, E. Demir, H. W. Kim, E. Viragh, S. L. Zipursky, B. J. Dickson, Dscam diversity is essential for neuronal wiring and self-recognition. *Nature* **449**, 223–227 (2007).
38. B. E. Chen, M. Kondo, A. Garnier, F. L. Watson, R. Puettmann-Holgado, D. R. Lamar, D. Schmucker, The molecular diversity of Dscam is functionally required for neuronal wiring specificity in Drosophila. *Cell* **125**, 607–620 (2006).
39. H. He, Y. Kise, A. Izadifar, O. Urwyler, D. Ayaz, A. Parthasarthy, B. Yan, M. L. Erfurth, D. Dascenco, D. Schmucker, Cell-intrinsic requirement of Dscam1 isoform diversity for axon collateral formation. *Science* **344**, 1182–1186 (2014).
40. Y. Yue, S. Hou, X. Wang, L. Zhan, G. Cao, G. Li, Y. Shi, P. Zhang, W. Hong, H. Lin, B. Liu, F. Shi, Y. Yang, Y. Jin, Role and convergent evolution of competing RNA secondary structures in mutually exclusive splicing. *RNA Biol.* **14**, 1399–1410 (2017).
41. M. Suyama, Mechanistic insights into mutually exclusive splicing in dynamin 1. *Bioinformatics* **29**, 2084–2087 (2013).
42. K. Hatje, R. U. Rahman, R. O. Vidal, D. Simm, B. Hammesfahr, V. Bansal, A. Rajput, M. E. Mickael, T. Sun, S. Bonn, M. Kollmar, The landscape of human mutually exclusive splicing. *Mol. Syst. Biol.* **13**, 959 (2017).
43. W. Hong, J. Zhang, H. Dong, Y. Shi, H. Ma, F. Zhou, B. Xu, Y. Fu, S. Zhang, S. Hou, G. Li, Y. Wu, S. Chen, X. Zhu, W. You, F. Shi, X. Yang, Z. Gong, J. Huang, Y. Jin, Intron-targeted mutagenesis reveals roles for Dscam1 RNA pairing architecture-driven splicing bias in neuronal wiring. *Cell Rep.* **36**, 109373 (2021).
44. H. Pan, Y. Shi, S. Chen, Y. Yang, Y. Yue, L. Zhan, L. Dai, H. Dong, W. Hong, F. Shi, Y. Jin, Competing RNA pairings in complex alternative splicing of a 3′ variable region. *RNA* **24**, 1466–1480 (2018).
45. D. D. Pervouchine, E. E. Khrameeva, M. Y. Pichugina, O. V. Nikolaienko, M. S. Gelfand, P. M. Rubtsov, A. A. Mironov, Evidence for widespread association of mammalian splicing and conserved long-range RNA structures. *RNA* **18**, 1–15 (2012).
46. X. O. Zhang, R. Dong, Y. Zhang, J. L. Zhang, Z. Luo, J. Zhang, L. L. Chen, L. Yang, Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* **26**, 1277–1287 (2016).
47. S. Kalmykova, M. Kalinina, S. Denisov, A. Mironov, D. Skvortsov, R. Guigo, D. Pervouchine, Conserved long-range base pairings are associated with pre-mRNA processing of human genes. *Nat. Commun.* **12**, 2300 (2021).
48. M. Kalinina, D. Skvortsov, S. Kalmykova, T. Ivanov, O. Dontsova, D. D. Pervouchine, Multiple competing RNA structures dynamically control alternative splicing in the human ATE1 gene. *Nucleic Acids Res.* **49**, 479–490 (2021).

49. F. C. Oberstrass, S. D. Auweter, M. Erat, Y. Hargous, A. Henning, P. Wenter, L. Reymond, B. Amir-Ahmady, S. Pitsch, D. L. Black, F.-H. T. Allain, Structure of PTB bound to RNA: Specific binding and implications for splicing regulation. *Science* **309**, 2054–2057 (2005).
50. X. Ren, J. Sun, B. E. Housden, Y. Hu, C. Roesel, S. Lin, L. P. Liu, Z. Yang, D. Mao, L. Sun, Q. Wu, J. Y. Ji, J. Xi, S. E. Mohr, J. Xu, N. Perrimon, J. Q. Ni, Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 19012–19017 (2013).
51. W. B. Grueber, L. Y. Jan, Y. N. Jan, Tiling of the Drosophila epidermis by multidendritic sensory neurons. *Development* **129**, 2867–2878 (2002).
52. X. Ren, Z. Yang, J. Xu, J. Sun, D. Mao, Y. Hu, S. J. Yang, H. H. Qiao, X. Wang, Q. Hu, P. Deng, L. P. Liu, J. Y. Ji, J. B. Li, J. Q. Ni, Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in Drosophila. *Cell Rep.* **9**, 1151–1162 (2014).
53. M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, D. G. Higgins, Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
54. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
55. M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
56. M. G. Reese, F. H. Eeckman, D. Kulp, D. Haussler, Improved splice site detection in Genie. *J. Comput. Biol.* **4**, 311–323 (1997).
57. N. Grillenzoni, J. van Helden, C. Dambly-Chaudiere, A. Ghysen, The iroquois complex controls the somatotopy of Drosophila notum mechanosensory projections. *Development* **125**, 3563–3569 (1998).

Acknowledgments: We thank Z. Wang and L. Huang for comments on the manuscript and members of the Jin laboratory for suggestions and discussion during the course of this work. **Funding:** This work was supported by research grants from the National Natural Science Foundation of China (91940303 and 31630089 to Y.J.; 31671510 and 31871461 to H.H.), the National Key Research and Development Program of China (2021YFE0114900 to Y.J.), the Natural Science Foundation of Zhejiang Province (LD21C050002 to Y.J.; LQ22C050001 to H.D.), and the Starry Night Science Fund at Shanghai Institute for Advanced Study of Zhejiang University (SN-ZJU-SIAS-009 to Y.J.). **Author contributions:** Y.J. conceived of this project. H.D., B.X., P.G., J.Z., X.Y., L.L., J.S., F.Z., Y.F., S.Z., and L.B. designed and performed the experiments. Y.S., J.Z., B.X., and H.D. performed RNA secondary structure prediction and evolutionary conserved elements analysis. H.D., L.L., J.S., F.Z., Y.F., S.Z., P.G., and L.B. carried out vector construction and mutant flies screening. H.D., L.L., J.S., P.G., and Y.Z. prepared high-throughput sequencing samples of mutant flies. H.D., B.X., J.Z., W.Y., and Y.S. conducted the analysis of isoform expression. H.D., P.G., J.Z., X.Y., L.L., J.H., and H.H. analyzed the developmental defects of RNA secondary structure mutant flies. Y.J., F.S., H.D., B.X., J.Z., Y.S., F.Z., X.Z., J.H., H.H., and X.F.Y. analyzed the data. Y.J., H.D., and B.X. wrote the manuscript. All authors discussed the results and commented on the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 1 September 2021
Accepted 3 December 2021
Published 26 January 2022
10.1126/sciadv.abm1763