

A Pandas complex adapted for piRNA-guided transcriptional silencing and heterochromatin formation

Kang Zhao^{1,2,11}, Sha Cheng^{2,3,11}, Na Miao^{1,2,11}, Ping Xu^{1,4,11}, Xiaohua Lu^{1,2,11}, Yuhua Zhang^{3,5,11}, Ming Wang¹, Xuan Ouyang^{1,2}, Xun Yuan^{2,3}, Weiwei Liu¹, Xin Lu^{1,2}, Peng Zhou^{1,2}, Jiaqi Gu^{1,5}, Yiqun Zhang¹, Ding Qiu^{1,2}, Zhaohui Jin^{2,3}, Chen Su⁶, Chao Peng⁶, Jian-Hua Wang⁷, Meng-Qiu Dong^{8,9}, Youzhong Wan⁴, Jinbiao Ma^{1,5}, Hong Cheng^{2,3}, Ying Huang^{1,2,3,10*} and Yang Yu^{1,2*}

The repression of transposons by the Piwi-interacting RNA (piRNA) pathway is essential to protect animal germ cells. In *Drosophila*, Panoramix enforces transcriptional silencing by binding to the target-engaged Piwi-piRNA complex, although the precise mechanisms by which this occurs remain elusive. Here, we show that Panoramix functions together with a germline-specific paralogue of a nuclear export factor, dNxf2, and its cofactor dNxt1 (p15), to suppress transposon expression. The transposon RNA-binding protein dNxf2 is required for animal fertility and Panoramix-mediated silencing. Transient tethering of dNxf2 to nascent transcripts leads to their nuclear retention. The NTF2 domain of dNxf2 competes dNxf1 (TAP) off nucleoporins, a process required for proper RNA export. Thus, dNxf2 functions in a Panoramix-dNxf2-dependent TAP/p15 silencing (Pandas) complex that counteracts the canonical RNA exporting machinery and restricts transposons to the nuclear peripheries. Our findings may have broader implications for understanding how RNA metabolism modulates heterochromatin formation.

Transposons are highly abundant in eukaryotes and make up nearly half of the human genome. To maintain the integrity of the eukaryotic genome, transposons are often silenced by PIWI-clade Argonaute proteins, guided by 23–30-nucleotide small RNA partners (Piwi-interacting RNA, piRNA)^{1–5}. The nuclear PIWI proteins in animal gonads are thought to recognize nascent transposon transcripts for transcriptional gene silencing (TGS) and direct sequence-specific heterochromatin formation^{1–5}. As a critical cofactor of *Drosophila* nuclear Piwi, Panoramix (Panx; also known as Silencio) links the target-engaged Piwi-piRNA complex to the general silencing machinery^{6,7}. Enforced tethering of Panx to nascent transcripts leads to cotranscriptional silencing and correlates with the deposition of histone H3 lysine 9 trimethylation (H3K9me3) marks^{6,7}. However, the mechanism by which Panx mediates this repression remains unknown.

Genetically, both the H3K9 methyltransferase Eggless/SetDB1 and the H3K4 demethylase Su(var)3-3/Lsd1 are required for Panx-mediated TGS^{6,7}. Moreover, HP1a/Su(var)2-5, a reader protein for H3K9me2/3, is critical for this process^{6,7}. It is unclear how Panx is connected to these general heterochromatin factors. Unlike other piRNA pathway-specific effector step factors (for example, Asterix/Gtsf1 or Maelstrom), Panx is the only protein known to be capable

of driving heterochromatin formation if tethered via either RNA or DNA^{6–10}. In contrast, direct tethering of HP1a to RNA failed to induce silencing, although the recruitment of HP1a to DNA is sufficient to do so⁷. It is unclear what makes Panx capable of inducing heterochromatin formation from nascent RNAs while HP1a cannot. An equally puzzling question is why H3K9me3 marks are not always sufficient for transposon silencing⁸.

Here, we identified *Drosophila* Nxf2 (dNxf2) as a critical partner of Panx and showed that dNxf2 is essential for Panx-mediated silencing. The dNxf2 protein belongs to an evolutionarily conserved family of proteins called nuclear RNA-export factors. Structural and biochemical analyses revealed that dNxf2 binds to Panx through its ubiquitin-associated (UBA) domain. Unlike the canonical family member *Drosophila* Nxf1 (dNxf1), dNxf2 lost the capability to bind to nuclear pore complexes. We further demonstrated that dNxf2 can interact with dNxf1 directly, which facilitates Panx-mediated silencing by counteracting transposon RNA export, at least in part by competing with dNxf1 for nucleoporin.

Results

Panx interacts with dNxf2. To understand the molecular function of Panx, we cross-examined proteins co-immunoprecipitated

¹Key Laboratory of RNA Biology, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China.

²University of Chinese Academy of Sciences, Beijing, China. ³State Key Laboratory of Molecular Biology, Shanghai Key Laboratory of Molecular Andrology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, China.

⁴National Engineering Laboratory of AIDS Vaccine, Key Laboratory for Molecular Enzymology and Engineering of Ministry of Education, School of Life Sciences, Jilin University, Changchun, China. ⁵State Key Laboratory of Genetic Engineering, Department of Biochemistry, School of Life Sciences, Fudan University, Shanghai, China. ⁶National Facility for Protein Science in Shanghai, Zhangjiang Lab, Shanghai Science Research Center, Chinese Academy of Sciences, Shanghai, China. ⁷Graduate School of Peking Union Medical College and Chinese Academy of Sciences of Medical Sciences, Beijing, China.

⁸National Institute of Biological Sciences, Beijing, China. ⁹Tsinghua Institute of Multidisciplinary Biomedical Research, Tsinghua University, Beijing, China.

¹⁰Shanghai Key Laboratory of Biliary Tract Disease Research, Shanghai Research Center of Biliary Tract Disease, Department of General Surgery, Xinhua Hospital, affiliated with Shanghai Jiao Tong University School of Medicine, Shanghai, China. ¹¹These authors contributed equally: Kang Zhao, Sha Cheng, Na Miao, Ping Xu, Xiaohua Lu, Yuhua Zhang. *e-mail: huangy@sibcb.ac.cn; yuyang@ibp.ac.cn

with Panx (Supplementary Fig. 1a) and piRNA pathway candidate genes from RNA interference (RNAi) screens^{11–14}. Unexpectedly, dNxf2 was identified as a potential cofactor of Panx (Supplementary Fig. 1a–c). Although dNxf2 belongs to an evolutionarily conserved family of nuclear RNA-export factor proteins, depletion of dNxf2 has no effect on polyadenylated messenger RNA export^{15,16}. Instead, dNxf2 and its cofactor *Drosophila* Nxt1 (dNxt1, also known as p15) have both been identified as essential for transposon silencing in two published RNAi screens^{11,12}. Like Panx, dNxf2 is specifically expressed in female gonads (Supplementary Fig. 1d). To validate the interaction between Panx and dNxf2, immunoprecipitation of green fluorescent protein (GFP) was performed from ovaries expressing the GFP–Panx fusion protein under its native promoter. The results of mass-spectrometry (Supplementary Fig. 1b) and western-blot analyses demonstrated that endogenous dNxf2 is associated with Panx (Fig. 1a). Similarly, dNxf2–Halo with a Halo tag integrated into the C terminus was able to precipitate endogenous Panx from the lysates of ovarian somatic cells (OSC; Fig. 1b and Supplementary Fig. 4a). Although Piwi can be consistently detected in the Panx immunoprecipitates by western blots^{6,7}, we were unable to detect significant enrichment for Piwi from the GFP–Panx pull downs compared with the negative controls in our current mass-spectrometry analysis (Supplementary Fig. 1b), suggesting that the Piwi–Panx complex may be present at low levels. It is likely that most Piwi proteins are not associated with Panx in a background where transposons remain largely silenced^{6,7}.

dNxf2 is required for Panx-mediated silencing. We next tested whether dNxf2 is functionally required for Panx-mediated silencing. The luciferase transcripts with BoxB sites in their 3' untranslated regions (UTR) are repressed if λ N–Panx is tethered^{6,7}. The expression levels of luciferase were measured following the germline-specific knockdown of either dNxf2 or dNxt1 (Fig. 1c). Loss of either dNxf2 or dNxt1 significantly weakened the ability of Panx to repress the reporter, as compared with the controls (Zuc or attp2; Fig. 1c). Consistent with this, transposon transcripts were elevated following dNxf2 RNAi (Supplementary Fig. 1e). Together, our data suggest that dNxf2 and dNxt1 may function as a heterodimer, either with or downstream of Panx, to suppress transposon expression.

We then used RNA sequencing (RNA-seq) to examine the global effects on transposon expression with germline-specific knockdowns of dNxf2 compared with Panx RNAi (Fig. 1d–f). Knockdown of dNxf2 triggered a dramatic increase in transposon transcripts (Fig. 1d), similar to that of Panx (Fig. 1e,f), suggesting that dNxf2 is specifically required for the silencing of transposons repressed by Panx. To rule out off-target effects of RNAi, a loss-of-function dNxf2 mutant was generated using clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR associated protein 9 (Cas9) (Supplementary Fig. 2a)¹⁷. The dNxf2-mutant female flies carrying a deletion of 20 amino acids at the N terminus were completely sterile (Supplementary Fig. 2b), similar to other core piRNA-pathway mutants⁶. Yet, loss of dNxf2 showed little effect on Piwi nuclear localization or stability (Supplementary Fig. 2c,d), indicating that dNxf2 functions as an effector protein rather than in piRNA biogenesis. Consistent with this, the dNxf2 mutants showed derepression of the Panx-tethered reporter (Fig. 1g) and global upregulation of transposons (Supplementary Fig. 2e–g). The absence of dNxf2 unexpectedly reduced the protein levels of endogenous Panx (Supplementary Fig. 2d and Fig. 1h). To rule out the possibility that dNxf2 may indirectly affect transposons via Panx stability, the levels of transposon expression were measured following the overexpression of λ N–FLAG–Panx (Supplementary Fig. 2h). The dNxf2-mutant female flies lost transposon controls (Supplementary Fig. 2h,i) and were still completely sterile (Fig. 1i), as if Panx did not exist.

Panx forms a ternary complex with dNxf2 and dNxt1. The striking phenotypic similarities between dNxf2 and Panx prompted us to test whether these two proteins interact directly. We used yeast two-hybrid (Y2H) assays to determine the interacting regions. The domain architecture of dNxf2 is similar to that of the canonical RNA-export factor dNxf1 (Fig. 2a). Both proteins contain an RNA recognition motif (RRM), leucine-rich repeats (LRR), a nuclear transport factor 2 (NTF2)-like domain and a UBA domain. Panx specifically interacted with the UBA domain of dNxf2 but not that of dNxf1 (Fig. 2b). Surprisingly, neither the full-length nor the NTF2+UBA fragment of dNxf2 could bind Panx (Fig. 2b), suggesting that dNxf2^{UBA} might adopt a 'closed' conformation in the presence of the NTF2 domain. Consistent with this, the interactions between dNxt1 and the NTF2 domains (dNxf2^{NTF2} or dNxf1^{NTF2}) were weakened in the presence of its UBA domain (Fig. 2c). As dNxt1 itself is absent in the Y2H system, we tested whether dNxt1 might release the UBA domain from the NTF2 domain to permit Panx binding. Ectopic expression of dNxt1 was indeed sufficient to allow full-length dNxf2 to interact with Panx in a Y2H assay (Fig. 2d). We then mapped the minimum region of Panx down to residues 315–343 (dNxf2-interacting region, NIR) as sufficient for UBA binding (Fig. 2e,f). Consistent with the fact that dNxt1 forms a heterodimer with dNxf2 (ref. 15), dNxt1 co-migrated with a fusion protein of dNxf2^{NTF2+UBA}–(Gly-Ser)₄–Panx^{NIR} in size-exclusion chromatography (Fig. 2g and Supplementary Fig. 7a,b), suggesting that Panx, dNxf2 and dNxt1 exist as a ternary complex. We failed to crystallize the dNxf2^{NTF2} domain; instead, we crystallized dNxf1^{NTF2} in complex with dNxt1 and determined the structure (Supplementary Fig. 3a). Residues that may be involved in the binding of dNxf2^{NTF2} to dNxt1 were modelled according to the sequence alignment result and the structure of dNxf1^{NTF2} (Supplementary Fig. 3b). Co-immunoprecipitation and Y2H assays validated that dNxf2^{NTF2} maintained most, if not all, residues that interact with dNxt1 (Fig. 2c and Supplementary Fig. 3c), indicating that the interaction mode of the dNxf2^{NTF2}–dNxt1 complex is similar to that of dNxf1^{NTF2}–dNxt1 (Supplementary Fig. 3a–e).

Purified Panx^{NIR} forms a stable complex with dNxf2^{UBA}, with a dissociation constant of approximately 3.2 μ M (Fig. 2h). We then determined the crystal structure of the dNxf2^{UBA}–Panx complex (Fig. 3a–d). Despite many efforts, only the fusion protein of dNxf2^{UBA}–(Gly-Ser)₄–Panx^{NIR} could be successfully crystallized and solved at a resolution of 1.5 Å (Supplementary Table 1). *Drosophila* Nxf2^{UBA} forms a compact three-helix bundle (α 1– α 3) with a 3_{10} -helix (η 1) at the C terminus (Fig. 3b). The Panx^{NIR} is folded into a long α -helix and lays on the hydrophobic surface formed by α 2 and α 3 (Fig. 3b–d). In contrast to the highly charged surface of the Nxf1-type UBA (human Nxf1 or yeast Mex67), dNxf2^{UBA} favours hydrophobic binding with Panx (Fig. 3c). Key residues on the interacting interface are highly conserved among different *Drosophila* species but altered in the Nxf1-type UBA (Supplementary Fig. 3f). A324, A328, V331, L332 and I335 on Panx interact with V80, F819, F826, F840, L823 and I827 on dNxf2 via hydrophobic interactions (Fig. 3d). Moreover, R321 and R327 on Panx form salt bridges with D837 and E799, respectively, on Nxf2^{UBA} (Fig. 3d). Key residues on the interacting interface were mutated to validate the intermolecular interactions between dNxf2 and Panx (Fig. 3e). Whereas the single point mutations L823A and D837A both affected the binding between Panx and dNxf2^{UBA}, the double point mutation of dNxf2^{UBA} (F826A + I827A) nearly abolished its interaction with Panx in both Y2H and co-immunoprecipitation assays (Fig. 3e,f).

The UBA domain of dNxf2 is essential for silencing. To validate the importance of the direct interactions between dNxf2 and Panx in vivo, a previously described λ N/BoxB luciferase reporter system⁶ was used to test whether artificial tethering of dNxf2 could lead to repression. Significant repression following tethering of a λ N–dNxf2

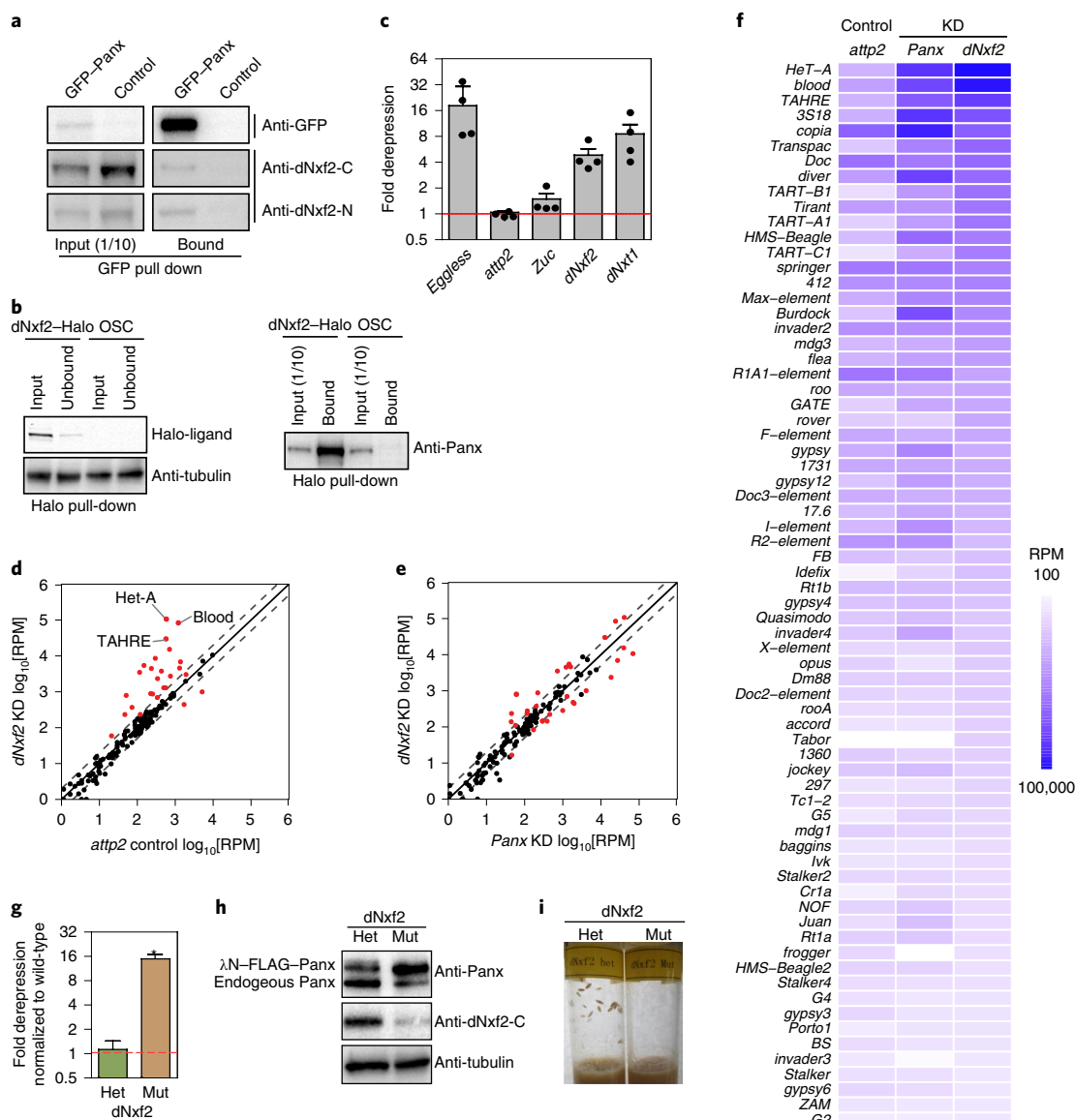


Fig. 1 | dNxf2 functions as a cofactor of Panx in the piRNA pathway. **a**, Western blots showing the co-immunoprecipitation of dNxf2 with GFP-Panx from ovary lysates. Two different rabbit polyclonal monospecific dNxf2 (dNxf2-N and dNxf2-C) antibodies were used. **b**, Halo-ligand staining and western blots showing co-immunoprecipitation of Panx with dNxf2-Halo from the knock-in OSC lysates. A rabbit polyclonal monospecific Panx antibody was used. The depletion of dNxf2-Halo proteins in the unbound sample of Halo pull downs is shown (left); the anti-tubulin blots were used as loading controls; the endogenous Panx proteins are shown on the right. **c**, The effects of germline (*nanos*-GAL4) knockdown of the indicated genes on *Renilla*-normalized firefly luciferase activity of the reporter while tethering λ N-Panx. The relative value of the *attp2* control was used for normalization. Data shown as the mean \pm s.d. ($n = 4$ biologically independent replicates). **d**, The comparison of the steady-state RNA levels is shown as reads per million (RPM) mapping to the sense strand of each transposon consensus from the germline knockdown (KD) of *dNxf2* versus the control. The dashed lines indicate the two-fold changes. The average of two replicates is shown. The red dots indicate transposons with significant changes ($P < 0.05$; Wald test). **e**, Same as in **d**, except that *dNxf2* KD was compared with *Panx* KD. The red dots indicate transposons with significant changes from **d**. **f**, Heat map displaying the steady-state RNA levels (RNA-seq) as RPM of the top-70 detected transposons from the germline knockdowns of the indicated genes. **g**, Expression of the total-protein-normalized firefly luciferase activity of the reporter under the indicated background (Het, heterozygous dNxf2 and Mut, *trans*-heterozygous dNxf2 mutant) while overexpressing λ N-FLAG-Panx. Data show the mean \pm s.d. ($n = 15$ biologically independent replicates; $*P = 1.41387 \times 10^{-7}$, Wilcoxon rank-sum test). **h**, Western blots showing the protein expression levels of dNxf2 and Panx from either heterozygous or mutant dNxf2 ovary lysates while overexpressing λ N-FLAG-Panx. Beta-tubulin was used as a loading control. **i**, Vials showing the pupae from an equal number of either heterozygous or mutant dNxf2 female flies overexpressing Panx and mated with OregonR males. The red dashed line marks the one-fold change. **a, b, h, i**, Images are representative of $n = 3$ biologically independent experiments with similar results. The source data for **c, g** are provided in Supplementary Table 3. The unprocessed gels for **a, b, h** are provided in Supplementary Fig. 8.

fusion protein was observed (Fig. 3g), unlike the negative controls (λ N-YFP and GFP-dNxf2 lacking a λ N tag). The level of repression mediated by λ N-dNxf2 was found to be dosage dependent,

correlating with the number of BoxB sites (Fig. 3h). Most importantly, repression was dependent on the presence of the dNxf2^{UBA} domain (dNxf2- Δ UBA, Fig. 3g; also see Fig. 7h,i).

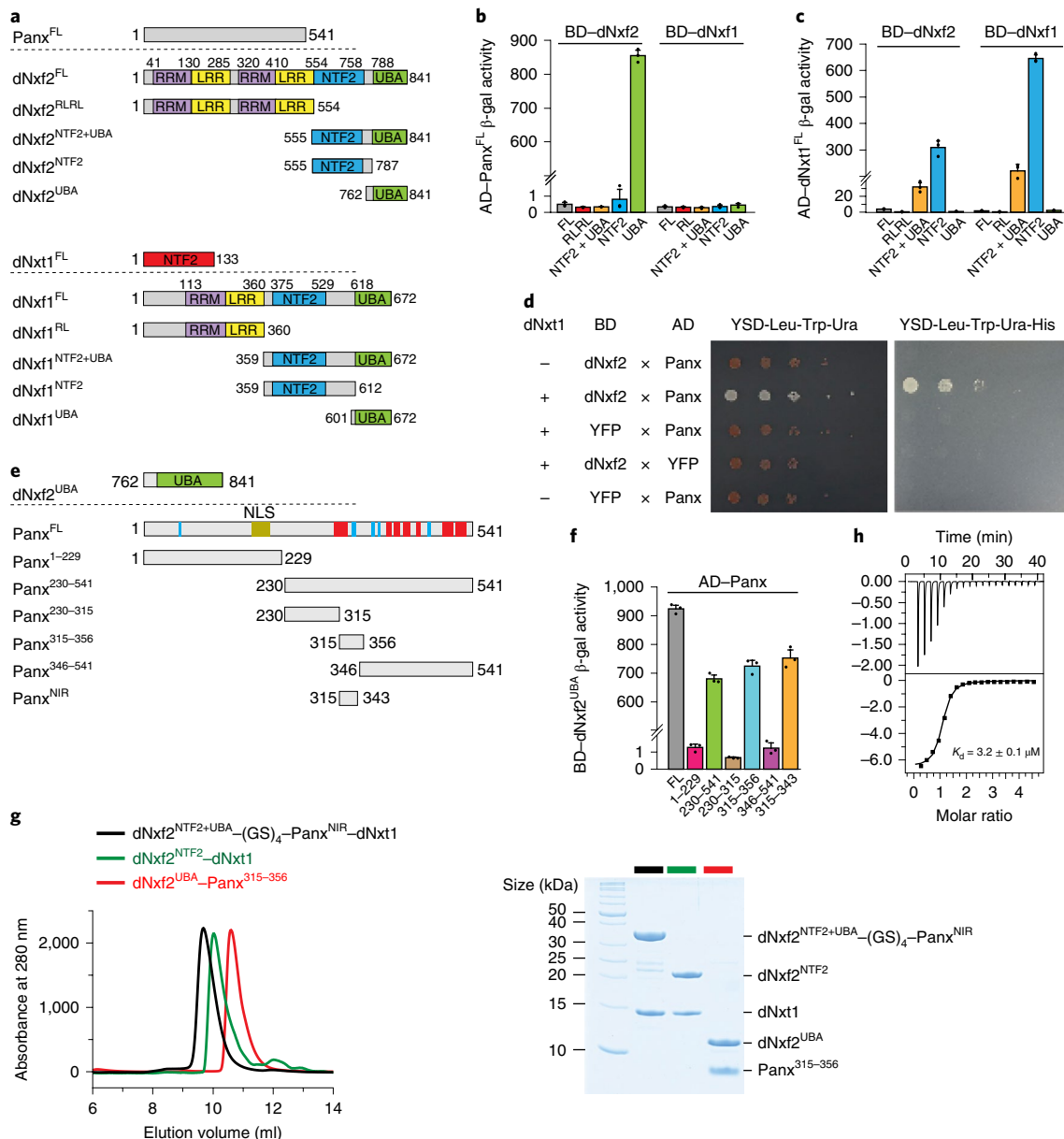


Fig. 2 | The UBA domain of dNxf2 interacts directly with Panx. **a**, Domain architectures of Panx, dNxf2, dNxf1 and dNxt1. The amino-acid residue numbers are indicated for each protein. The domain names are abbreviated within the respective coloured regions. **b–d,f**, The interacting regions between *Drosophila* Nxf1/2 and Panx or dNxt1 were mapped using Y2H assays. The interactions were determined by either measuring the β-galactosidase (β-gal) activity produced by the reporter gene (**b,c,f**) or growth on YSD media lacking the indicated essential amino acid or uracil (**d**). Data shown as the mean ± s.d. ($n=3$ biologically independent replicates). The proteins or fragments shown above the graphs were used as the prey in the assays. **b**, Mapping of the regions of *Drosophila* Nxf1/2 that interact with Panx. **c**, Mapping of the regions of *Drosophila* Nxf1/2 that interact with dNxt1. **d**, Yeast three-hybrid assay determining the requirements of dNxt1 for a Nxf2-Panx interaction. **e**, Schematic of full-length Panx and its truncated fragments used in the Y2H assays. The positions of the amino acid residues at either the N or C terminus of each protein fragment are indicated. Red and blue color indicate an α-helix and β-strand, respectively. **f**, Mapping of the minimum Panx regions that interact with dNxf2^{UBA}. **g**, The size-exclusion-chromatography profile of the NTF2 and UBA domains of dNxf2 forming heterodimers with dNxt1 and Panx^{NIR} in solution, respectively (left). A dNxf2 fragment spanning the NTF2 and UBA domains, which is covalently linked to Panx^{NIR}, forms a stable complex with dNxt1. The components of the peak in the elution profile by SDS-PAGE (right). The colour schemes used for the three complexes are indicated in the legend. **h**, Quantification of the dissociation constant (K_d) of the interaction between dNxf2^{UBA} and Panx^{NIR}, as measured by an isothermal titration calorimetry assay ($n=3$ biologically independent experiments with similar results). One dataset is shown. The K_d value is indicated. **d,g**, $n=3$ biologically independent experiments were repeated with similar results. The source data for **b,c,f** are provided in Supplementary Table 3. FL, full length.

Nucleoporin binding activity is lost in dNxf2^{UBA}. A hydrophobic pocket is formed on the opposite surface of the Nxf1-type UBA to accommodate the FxFG peptide of the nuclear pore complex (NPC; Fig. 4a,b). However, this pocket is missing in dNxf2^{UBA} due to a salt

bridge formed between K829 and E814 (Fig. 4b, dNxf2). In addition, the bulky L825 side chain on dNxf2^{UBA} may hinder FG binding (Fig. 4b). In contrast, the corresponding amino acids in human Nxf1 (A602) or yeast Mex67 (G583) contained much smaller side

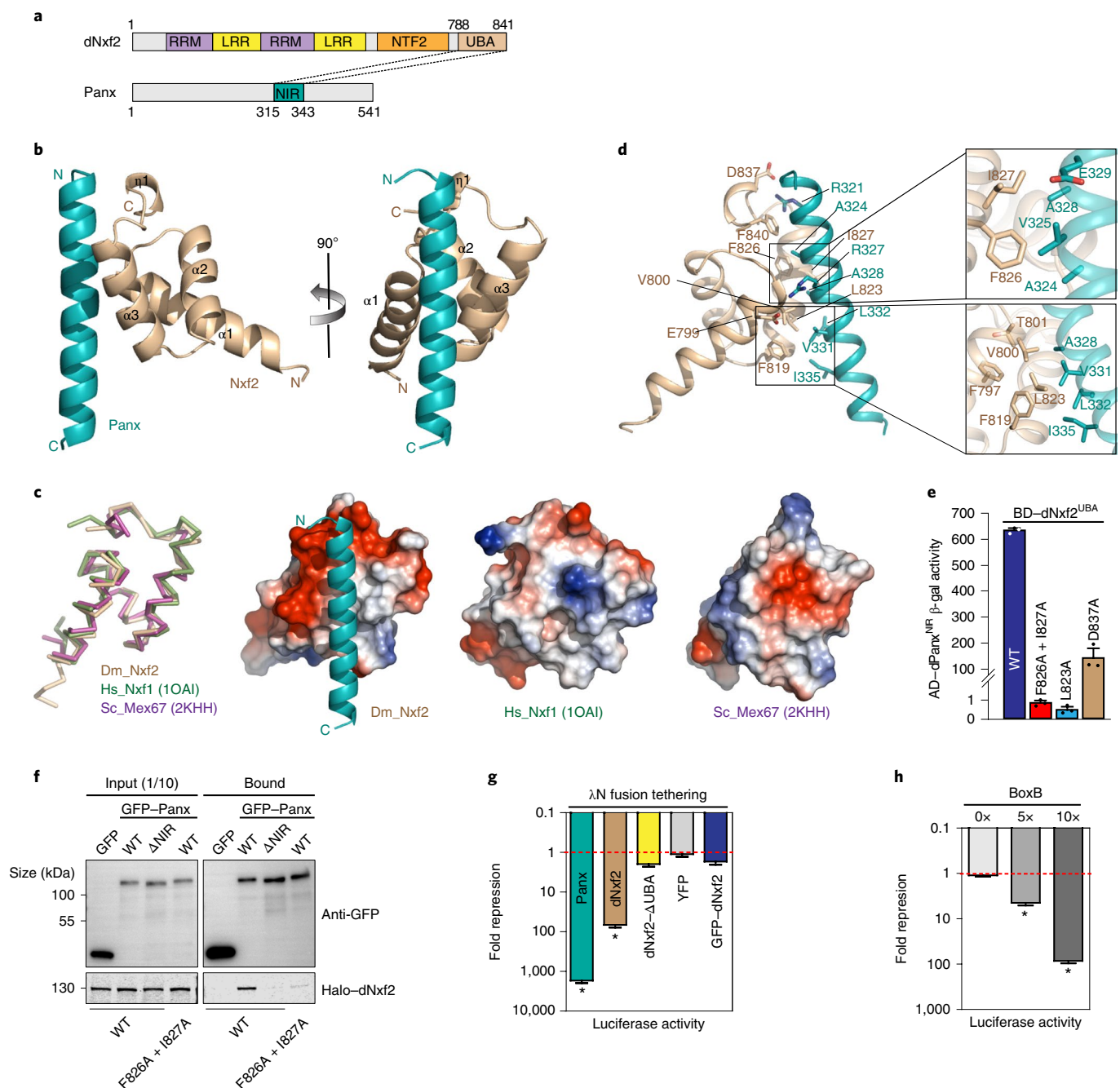


Fig. 3 | Structure of dNxf2^{UBA} in complex with Panx^{NIR}. **a**, Schematic of the interacting region between dNxf2 and Panx. The numbers above and below the diagrams indicate the positions of the amino acid residues in dNxf2 and Panx, respectively. The domain names are abbreviated within the respective coloured regions. **b**, Cartoon representation of dNxf2^{UBA} in complex with Panx^{NIR} (left). View of the complex rotated 90° around the vertical axis (right). **c**, Superimposition of the UBA domains of dNxf2, human Nxf1 (Protein Data Bank (PDB) accession code 1OAI) and yeast Mex67 (PDB code 2KHH; left). Electrostatic potential analysis showing that the corresponding surfaces in the UBA domains of dNxf2 (left), human Nxf1 (middle) and yeast Mex67 (right) are positively and negatively charged, respectively. **d**, Detailed view of the interactions between dNxf2^{UBA} and Panx^{NIR} are shown (right). The key residues involved in binding are shown as sticks. Close-up views of the hydrophobic interactions between dNxf2^{UBA} and Panx^{NIR} are shown (right). **e**, Binding of wild-type (WT) and mutant dNxf2^{UBA} to Panx^{NIR}, as measured by Y2H assays. Data shown as the mean ± s.d. ($n=3$ biologically independent replicates). **f**, Western blots and Halo-ligand staining showing co-immunoprecipitation of GFP-tagged Panx and its NIR deletion mutant (Δ NIR) with Halo-tagged dNxf2 or its F826A + I827A double mutant from OSCs. GFP serves as a negative control. Images are representative of $n=3$ biologically independent experiments with similar results. **g**, Effects of the indicated λ N fusion proteins and a non-tethering control (GFP-dNxf2) on luciferase activity, normalized to the total protein, of the reporters integrated into the attP2 landing site. Data shown as the mean ± s.d. ($n=15$ biologically independent replicates; $*P=1.41387 \times 10^{-7}$). **h**, The effects of λ N-dNxf2 tethering on luciferase activity of reporters with an increasing number of BoxB sites. All reporters were integrated into the same genomic locus (attP2 landing site). Fold repression was calculated as the firefly luciferase luminescence values, normalized to the total protein, of the control (no tethering) divided by that of the indicated experiments. Data shown as the mean ± s.d. ($n=15$ biologically independent replicates; $*P=1.41387 \times 10^{-7}$). The red dashed line marks the one-fold change. The source data for **e, h, g** are provided in Supplementary Table 3. The unprocessed gels for **f** are provided in Supplementary Fig. 8.

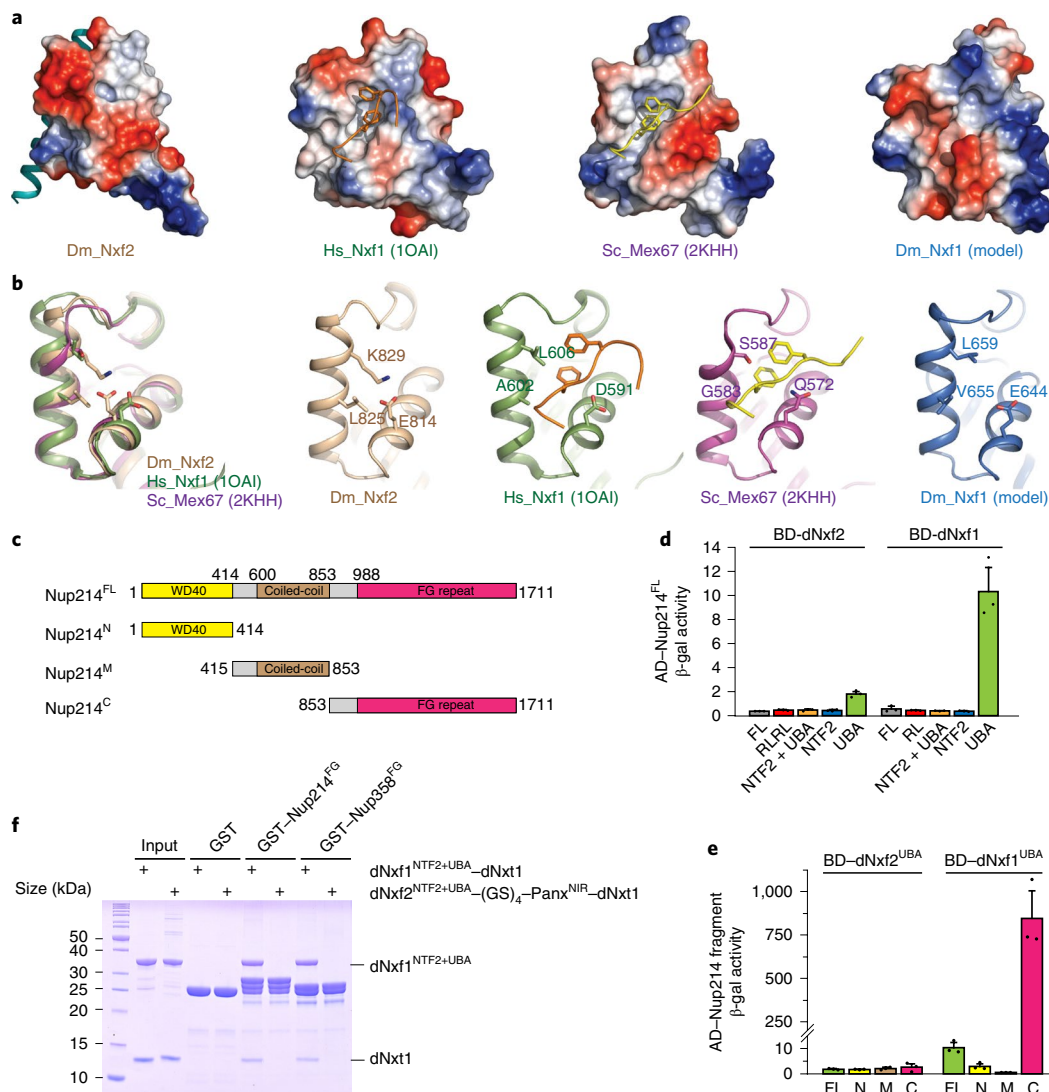


Fig. 4 | The UBA domain of dNxf2 lost NPC binding capability. **a**, Electrostatic potential analysis showing a closed surface on dNxf2 UBA that blocks the entry of FG-nup. In the corresponding surface of dNxf2 UBA, the passage for the peptide is blocked (left). Human Nxf1 UBA (PDB: 1OAI) and yeast Mex67 UBA (PDB: 2KHH) show a hydrophobic pocket to accommodate the two phenylalanines of the FxFG peptides (middle). A model of dNxf1 UBA (right) indicates that a peptide binding pocket may exist on the dNxf1 surface. **b**, Detailed view, left to right, of the UBA domains of dNxf2, human TAP, yeast Mex67, and dNxf1 (model). The dNxf1 model was generated by SWISS-MODEL server. In human Nxf1 and yeast Mex67, the FG peptides bind in the concave passage formed by $\alpha 2$ and $\alpha 3$. **c**, Schematic of full-length Nup214 and its truncated fragments used in the Y2H assays. Numbers corresponding to amino acid residues at either the N or C terminus of each protein fragment. **d**, Y2H assays show that dNxf1 UBA domain but not dNxf2 UBA interacts with full-length Nup214. Data are shown as the mean \pm s.d. (biologically independent replicates $n = 3$). **e**, Y2H assays show that dNxf1 UBA domain but not dNxf2 UBA interacts with the C-terminus FG-repeat region of Nup214 (Nup214C). Interactions were determined by either measuring the β -galactosidase activity produced by the reporter gene. Data are shown as the mean \pm s.d. (biologically independent replicates $n = 3$). Proteins or fragments of Nup214 are shown above the dashed line. **f**, SDS-PAGE showing pull-down results of the dNxf1NTF2+UBA/dNxt1 complex and dNxf2NTF2+UBA-(GS)₄-PanxNIR/dNxt1 complex by either GST-tagged Nup214FG or Nup358FG, respectively, compared to a GST control ($n = 3$ biologically independent experiments with similar results). One data set is shown. The beads were washed and aliquots of the bound fraction (20%) were analysed. 2 μ g of each input protein was loaded. Positions of molecular weight markers are indicated on the left in kDa. The source data for **d,e** are provided in Supplementary Table 3.

chains (Fig. 4b), therefore providing space for the FxFG interactions. Consistent with this, the dNxf2^{UBA} domain was unable to bind to the FG repeats of Nup214, a NPC component known to interact with dNxf1^{UBA} in both Y2H and glutathione S-transferase (GST) pull-down assays (Fig. 4c–f)¹⁶. Similar GST pull-down results were obtained using the FG repeats from Nup358 (Fig. 4f, GST–Nup358^{FG}). Given that a minimum of two copies of the FG binding domains (NTF2 and UBA) is required for proper RNA export¹⁸, dNxf2 lacks at least one copy of the FG-binding domain (UBA) and thus cannot export RNAs.

The dNxf2UBA lost nucleoporin binding activity. Piwi-targeted transposon transcripts are preferentially associated with dNxf2. The presence of RNA-binding domains (RBDs) at the N terminus of dNxf2 (Fig. 3a) implies that dNxf2 might bind to transposon RNAs¹⁵. We thus performed gel-omitted ligation-dependent cross-linking and immunoprecipitation (GoldCLIP) and quantitative PCR with reverse transcription (RT–qPCR) experiments (Fig. 5a), which rely on the covalent attachment of a Halo tag to the beads to allow denaturing purification of cross-linked protein–RNA complexes¹⁹. A Halo tag was inserted at the C terminus of dNxf2 (dNxf2–Halo)

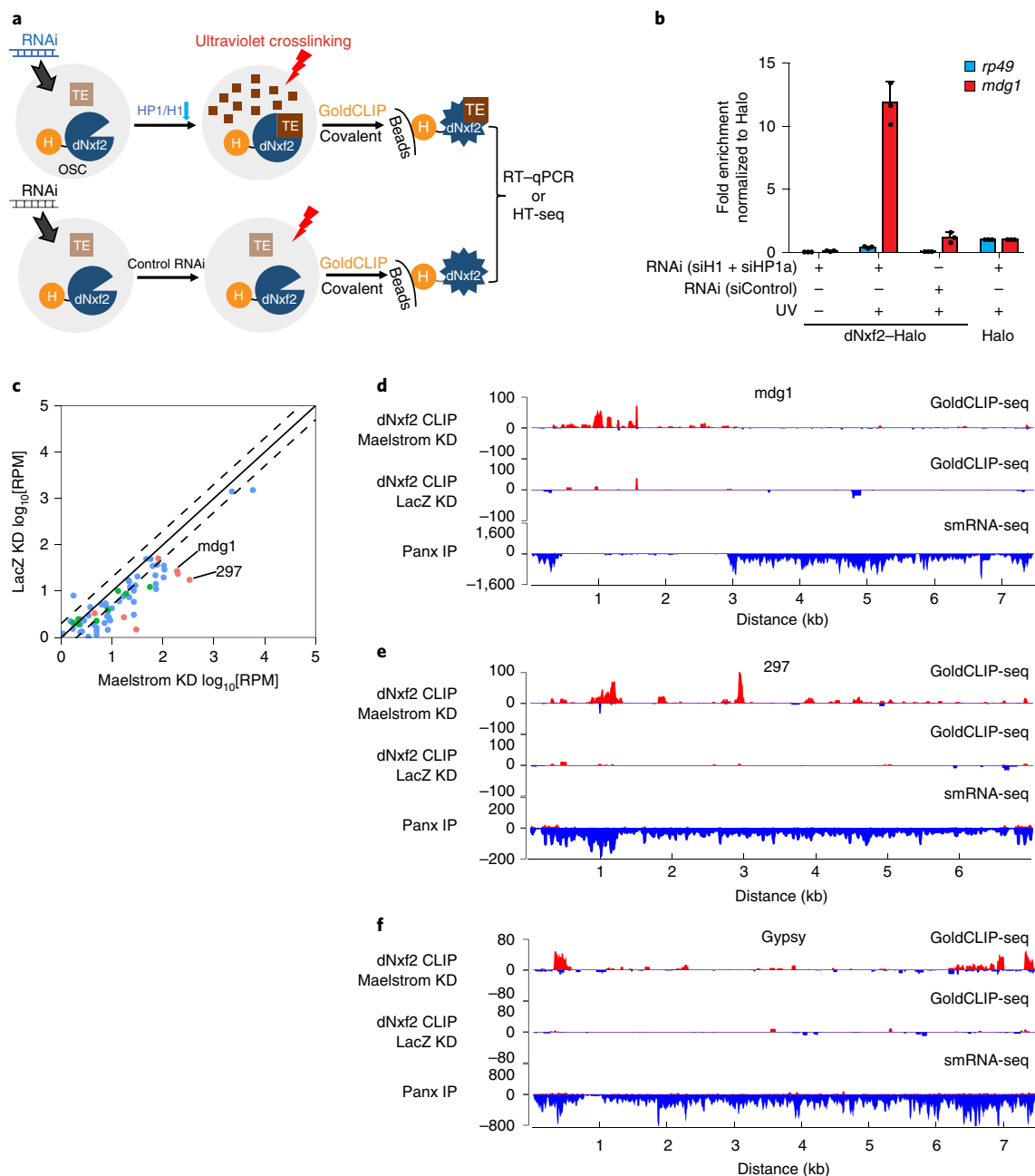


Fig. 5 | dNxf2 preferentially binds to piRNA-targeted transposons. **a**, Schematic representation of the GoldCLIP experiments. TE, transposon. H, Halo tag. HT-seq, high-throughput sequencing. **b**, RT-qPCR results showing relative levels of the *mdg1* transposon or *rp49* for dNxf2-Halo GoldCLIP samples, compared with either no ultraviolet treatment or a Halo control. Fold enrichments were calculated as the normalized mammalian spike-in RNA levels divided by that of the input samples. For comparison, the value of the Halo control was used for normalization. Data shown as the mean \pm s.d. ($n=3$ biologically independent replicates). **c**, Comparison of GoldCLIP sequencing data shown as RPM mapping to the sense strand of each transposon consensus for Maelstrom knockdown versus a LacZ control knockdown. A mammalian spike-in was used to normalize the different samples. The dashed lines indicate the two-fold changes. The red dots indicate the Piwi-targeted transposons, blue dots indicate the H1-dependent transposons and green dots indicate other transposons. **d**, Density plots of the dNxf2-Halo GoldCLIP sequencing reads from Maelstrom knockdown OSCs (top), dNxf2-Halo GoldCLIP reads from LacZ knockdown OSCs (middle) and Panx co-immunoprecipitated small modulatory RNA (smRNA) reads (bottom; from previously published data⁷) over the transposon consensus of *mdg1*. **e, f**, Same as in **d**, except for the transposon 297 (**e**) and Gypsy (**f**). IP, immunoprecipitation. **c–f**, The pooled reads from two replicates are shown. The source data for **b** are provided in Supplementary Table 3.

using CRISPR-Cas9 (Supplementary Fig. 4a)²⁰. *Mdg1* is one of the transposon families targeted by Piwi-piRNAs in OSCs^{8,21}. Following denaturing washes, the transcripts of *mdg1*, but not the house-keeping gene *rp49*, remained attached to the dNxf2-Halo fusion protein (Fig. 5b). This association was dependent on ultraviolet-

cross-linking, demonstrating direct binding between *mdg1* and dNxf2 (Fig. 5b). Interestingly, the interactions were only observed when both histone H1 and HP1a were depleted by RNAi, but not in the control knockdown (Fig. 5b), providing support to the idea that the majority of transposon transcripts remain suppressed in a

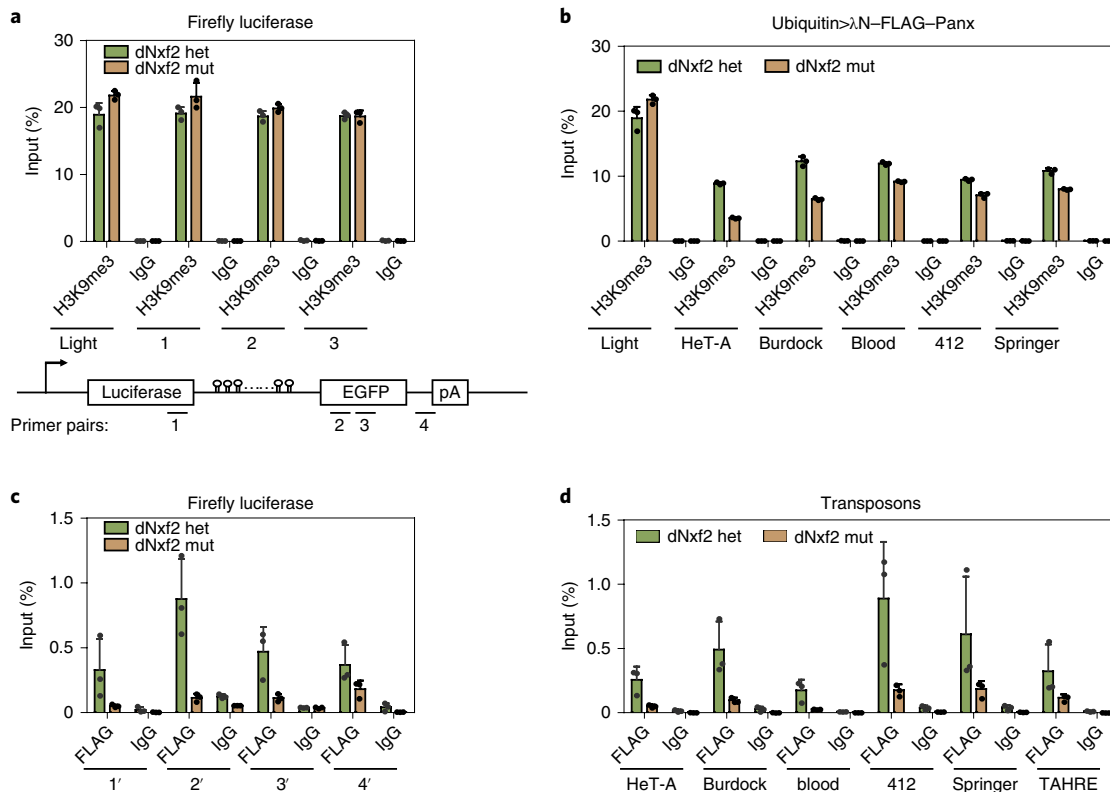


Fig. 6 | dNxf2 is required for Panx to associate with chromatin. a, H3K9me3-mark enrichments at the Firefly-10xBoxB reporter tethered with λ N-FLAG-Panx comparing heterozygous (het) and mutant (mut) dNxf2 when overexpressing Panx driven by a ubiquitin promoter (top). A schematic of the qPCR primer pair positions is shown (bottom). **b**, H3K9me3-mark enrichments at the indicated transposons in the same genetic background as in **a**. **a, b**, Rabbit polyclonal antibody specific to H3K9me3 and pre-immune rabbit IgG was used. **c**, FLAG-Panx enrichments at the Firefly-10xBoxB reporter tethered with λ N-FLAG-Panx in the same genetic background as in **a**. **d**, FLAG-Panx enrichments at the indicated transposons in the same genetic background as in **a**. **c, d**, Mouse monoclonal antibody specific to FLAG (M2) and pre-immune mouse IgG were used. Data shown as the mean \pm s.d. ($n=3$ biologically independent replicates). The source data are provided in Supplementary Table 3.

wild-type background. Low steady-state levels of transposon nascent RNAs make it difficult to obtain significant signals in an already inefficient ultraviolet-cross-linking experiment. Nevertheless, following the removal of the downstream silencing factors (H1 and HP1a), transposon transcripts accumulated and were bound by dNxf2-Halo (Fig. 5b). In contrast, Frogger—a transposon known not to be targeted by piRNAs²¹—did not show any detectable CLIP signal, although its transcripts were dramatically upregulated when depleting both H1 and HP1a (Supplementary Fig. 4c). This result suggests that the binding of dNxf2 to transposons could be correlated with piRNA targeting²¹. Furthermore, we performed GoldCLIP-sequencing experiments using dNxf2-Halo knock-in OSCs depleted of Maelstrom (Fig. 5c,d), a piRNA pathway effector component either parallel or downstream of H3K9me3 establishment on transposons^{8,10}. The CLIP sequencing data provided support for the idea that dNxf2 preferentially binds to the Piwi-targeted transposons, especially in the absence of Maelstrom (Fig. 5c–f).

Removal of the first RNA-binding unit (RRM + LRR, amino acids 1–285) of the dNxf2 RBDs only weakened silencing in the tethering assays (Δ 1st RBD unit; Supplementary Fig. 4d). The tethering effect of λ N-dNxf2- Δ 1st RBD was much weaker than that observed by either Brennecke or Siomi and colleagues^{22,23}. It is possible that either the location of BoxB (intronic versus 3' UTR) or their relative distance to the promoter could lead to the different results seen here (Supplementary Fig. 4d)^{22,23}. The BoxB sites in our reporter are located in the 3' UTR, whereas the BoxB sites were inserted into the intronic region by other groups. In contrast, deleting either

the second (Δ 2nd RBD; amino acids 286–554) or the entire RBD region (Δ RBD; amino acids 1–554) completely abolished silencing (Supplementary Fig. 4d; also see Fig. 7h,i). Therefore, the first RNA-binding unit is likely to be involved in the Piwi-dependent recruitment of dNxf2, whereas the second may be required for the effector step of silencing. Collectively, our data are consistent with the model in which Panx and dNxf2/dNxt1 function together to directly suppress transposons that are targeted by Piwi-piRNAs.

Uncoupling of H3K9me3 and silencing without dNxf2. As the loss of Panx leads to a significant reduction of H3K9me3 marks on transposons^{6,7}, we next tested whether the removal of dNxf2 could result in a similar phenomenon. Considering that Panx is unstable in the absence of dNxf2 (Supplementary Fig. 2d), we performed H3K9me3 chromatin immunoprecipitation (ChIP)–qPCR assays while overexpressing λ N-FLAG-Panx over several transposons as well as the Firefly-10xBoxB reporter. H3K9me3 showed marginal changes following the removal of dNxf2 (Fig. 6a,b and Supplementary Fig. 5a). In contrast, transposon transcripts were dramatically upregulated in the absence of dNxf2 (Supplementary Fig. 2h,i). This result suggests that transposon silencing and H3K9me3 deposition could somehow be uncoupled in the dNxf2 mutant. Interestingly, λ N-FLAG-Panx was unable to stay bound to the derepressed transposon chromatin, as measured by FLAG ChIP–qPCR (Fig. 6c). The FLAG-ChIP signals over the tethered reporter were also diminished, although λ N-FLAG-Panx was constitutively tethered to the RNAs (Fig. 6d). Given that Panx is thought to be recruited by Piwi-piRNAs via

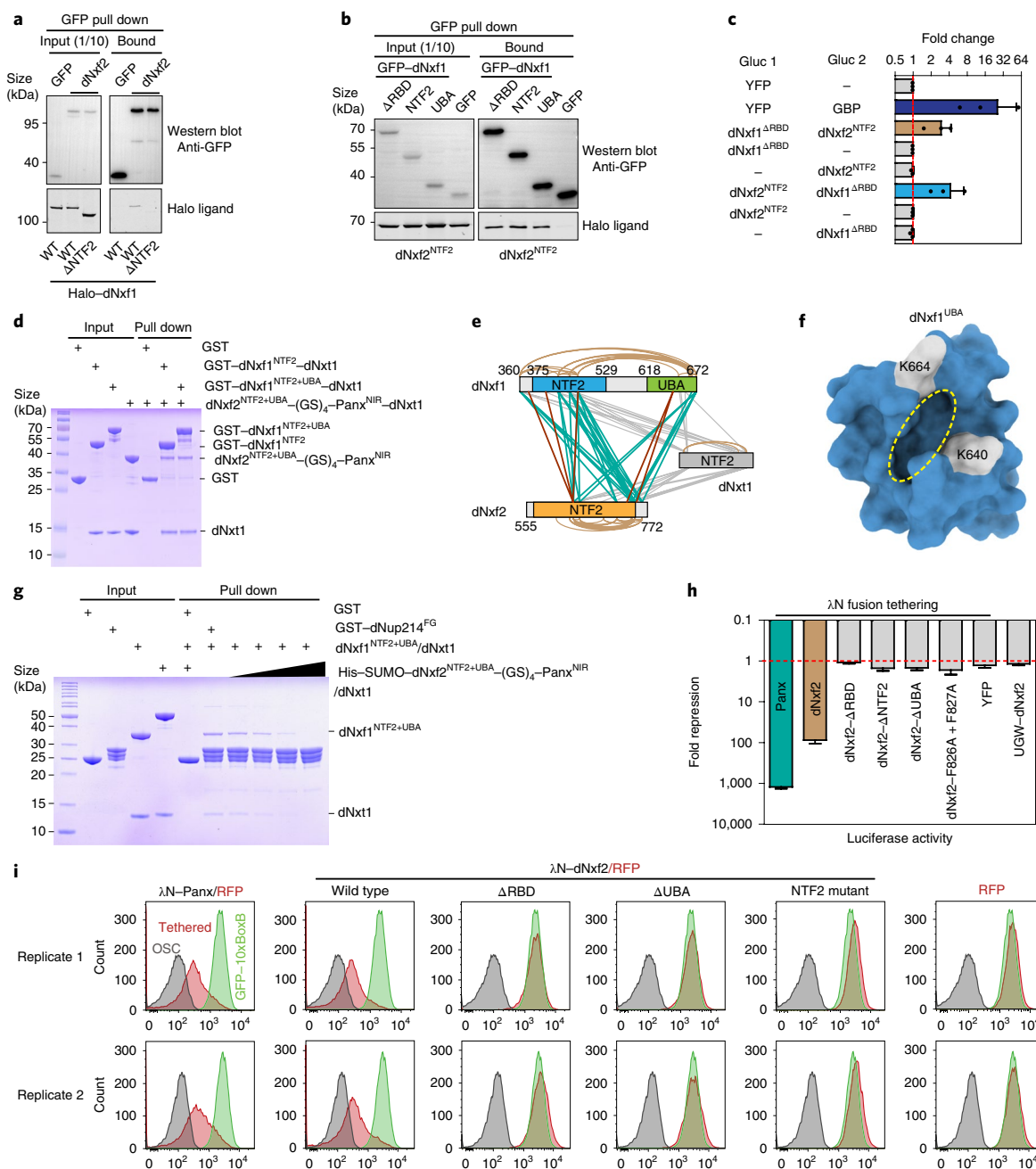


Fig. 7 | The NTF2 domain of dNxf2 interacts directly with dNxf1. **a**, Western blots and Halo-ligand staining showing co-immunoprecipitation of GFP-tagged dNxf2 with Halo-tagged dNxf1 from OSCs. Δ NTF2, dNxf1 lacking the NTF2 domain; GFP served as a negative control. **b**, Same as in **a**, except showing Halo-dNxf2^{NTF2} co-immunoprecipitated with different GFP-dNxf1 truncations. **c**, Split *Gaussia* luciferase complementation assays showing fold changes (total-protein-normalized luciferase values divided by that of the corresponding controls) in coexpression of the indicated proteins in S2 cells. Data shown as the mean \pm s.d. ($n=3$ biologically independent experiments). YFP, yellow fluorescent protein. **d**, SDS-PAGE gel showing the GST-tagged dNxf1^{NTF2}-dNxf1 and dNxf1^{NTF2+UBA}-dNxf1 pull-down results of the Nxf2^{NTF2+UBA}-(GS)₄-Panx^{NIR}-dNxf1 complex compared with a GST control. **e**, Schematic summary of statistically significant cross-linking residues identified between dNxf1^{NTF2+UBA}, dNxf2^{NTF2} and dNxf1 in a recombinant complex reconstituted in vitro. Intermolecular cross-links are shown as straight lines in either teal (dNxf1-dNxf2) or grey (dNxf1-dNxf1 or dNxf2-dNxf2) for DSS cross-linking and in maroon (dNxf1-dNxf2) for EDC cross-linking. Intramolecular cross-links are shown as brown curves. **f**, Schematic of the dNxf2^{NTF2}-cross-linking residues K664 and K640 on dNxf1^{UBA}. The dNxf1 model was generated by the SWISS-MODEL server. The FG-binding pocket is indicated by a dashed yellow circle. **g**, SDS-PAGE gel showing the competition assay of the proteins precipitated by GST-Nup214^{FG} pre-loaded with the dNxf1^{NTF2+UBA}-dNxf1 complex, incubated with increasing concentrations of the dNxf2^{NTF2+UBA}-(GS)₄-Panx^{NIR}-dNxf1 complex. **d,g**, The positions of the molecular weight markers are indicated on the left; 2 μ g of each input protein was loaded. **h**, Effects of the indicated λ N fusion proteins or a non-tethering control (GFP-dNxf2) on the luciferase activity of the reporter as in Fig. 3g. Data shown as the mean \pm s.d. ($n=10$ biologically independent replicates). **i**, Fluorescence-activated cell sorting (FACS) showing the counts and fluorescence of GFP-positive cells gated by OSCs expressing high levels of red fluorescent protein (RFP), which were electro-transfected with the indicated plasmids expressing different λ N fusion proteins. OSC, autofluorescence of OSCs; tethered, GFP reporter OSCs transfected with the indicated expression plasmids; GFP-10xBoxB, GFP reporter OSCs with no transfection. Two representative replicates are shown. The red dashed line marks the one-fold change. **a,b,d,g,i**, $n=3$ biologically independent experiments were repeated with similar results. The source data for **c,h** are provided in Supplementary Table 3. The unprocessed gels for **a,b** are provided in Supplementary Fig. 8.

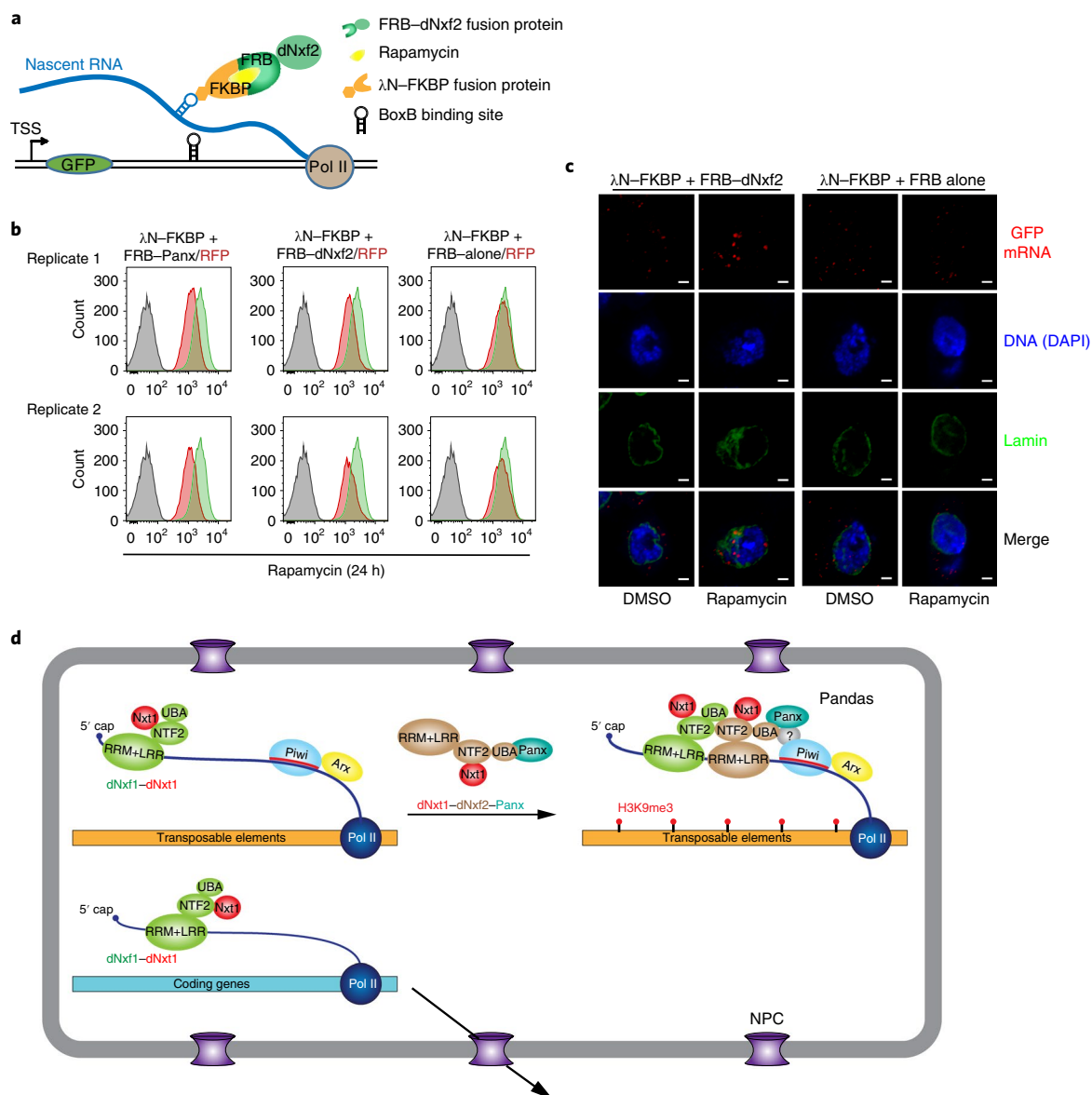


Fig. 8 | The Pandas complex is required for piRNA-guided transposon silencing. **a**, Schematic representation of a rapamycin-inducible system for λN-FKBP-FRB tethering. **b**, FACS showing counts and fluorescence of GFP-positive cells gated by high-RFP-expressing OSCs transfected with the plasmids expressing both λN-FKBP and the indicated FRB fusion proteins. The cells were treated with rapamycin 24 h after transfection for an additional 24 h. Two biologically independent replicates are shown. **c**, Confocal microscopy of RNA FISH by hybridization chain reaction (HCR). Comparison of the localizations of the reporter mRNAs tethered with λN-FKBP following transient recruitment of either FRB-dNxf2 or FRB alone when treated with rapamycin for 2 h to induce FKBP-FRB dimerization. Dimethylsulfoxide (DMSO) was used as a control. Scale bars, 2 μm. **d**, Model of the function of the Pandas complex in piRNA-guided silencing, preventing nascent transposon RNA export when targeted by Piwi-piRNAs. **b,c**, $n = 3$ biologically independent experiments were repeated with similar results.

nascent transposon transcripts, our data suggest that, without dNxf2, Panx together with nascent RNAs would leave the transcription loci and fail to efficiently suppress transposon expression.

Panx-dNxf2-dNxt1 and dNxf1-dNxt1 form a Pandas complex. The loss of Piwi leads to the dissociation of HeT-A chromatin from the nuclear peripheries²⁴. Consistent with this, we found that HeT-A chromatin left the nuclear peripheries without either Panx or dNxf2 (Supplementary Fig. 5b,c and Supplementary Videos 1–4). In this regard, I-element transcripts, targeted by piRNAs, have been shown to accumulate in the nucleus²⁵. We thus propose that certain RNA-export machinery may be modulated by the piRNA pathway to

retain transposons in the nucleus. As human Nxf1 has been reported to dimerize with other members of the nuclear RNA-export factor family to regulate RNA export^{26,27}, dNxf1 (TAP) could be one such candidate. We tested whether dNxf2 interacts with dNxf1 and thus counteracts the RNA-exporting function of dNxf1. Indeed, GFP-tagged dNxf2 was able to co-immunoprecipitate dNxf1-Halo from OSC lysates (Fig. 7a). *Drosophila* Nxf2^{NTF2} was sufficient to pull down either dNxf1^{NTF2} or dNxf1^{UBA} (Fig. 7b and Supplementary Fig. 6a–c). The interaction between dNxf2 and dNxf1 was sufficient to bring together a split *Gaussia* luciferase, demonstrating that dNxf2 and dNxf1 can be in close proximity in vivo (Fig. 7c). Consistent with this, the GST pull-down assays demonstrated direct binding

between dNxf1 and dNxf2 (Fig. 7d and Supplementary Fig. 6d–f). Using chemical cross-linking of proteins coupled with mass spectrometry²⁸, key residues in close proximity were identified as the intermolecular cross-links (Fig. 7e). Interestingly, K640 and K664 from dNxf1^{UBA}, located on each side of the FG-binding pocket, were cross-linked with dNxf2^{NTF2} (Fig. 7f). This raised the possibility that dNxf2 binding might interfere with the FG-binding pocket of dNxf1^{UBA}. In fact, an excess amount of dNxf2^{NTF2} can compete dNxf1^{NTF2+UBA} off the FG repeats of Nup214 (Fig. 7g). Together, our results demonstrate that the dNxf2^{NTF2} domain can directly block the access of the FG binding pockets on either the NTF2 or UBA domains of dNxf1. As the binding of dNxf1 to the NPC FG repeats is essential for RNA export¹⁵, dNxf2 may inhibit transposon export via its NTF2 domain to block the ability of dNxf1 to bind NPC. Moreover, the NTF2 domain of dNxf2 is essential for the silencing induced by tethering (dNxf2-ΔNTF2, Fig. 7h), demonstrating the functional importance of this interaction.

We next sought to directly visualize potential transient changes in RNA localization caused by dNxf2 using a rapamycin-inducible tethering system (Fig. 8a,b). The GFP mRNAs containing ten copies of BoxB-binding sites were mostly localized in the cytoplasm, despite constitutive λN-FKBP tethering (Fig. 8c). Following rapamycin treatment, λN-FKBP bound to the FRB–dNxf2 fusion protein, allowing a transient association of dNxf2 to the GFP mRNAs. Intriguingly, the GFP mRNAs started to accumulate in nuclei or at the nuclear peripheries following the binding of FRB–dNxf2 (Fig. 8c). This effect was specific to dNxf2, as tethering of FRB alone or other nucleus-localized piRNA factors (Piwi or Asterix) failed to cause any change (Fig. 8c, FRB alone, and Supplementary Fig. 7c). No GFP fluorescence could be detected following H₂O₂ bleaching, demonstrating that the signals in the fluorescence in situ hybridization (FISH) experiments came from lamin staining (Supplementary Fig. 7d). Given the involvement of Panx, dNxf2, dNxf1 (TAP) and dNxt1 (p15) in transposon silencing, we named this multi-protein complex Pandas (Panx–dNxf2-dependent TAP/p15 silencing). Our data raise the possibility that deterring the function of dNxf1 in transposon RNA export may be a key event in piRNA-guided silencing (Fig. 8d).

Discussion

Without dNxf2, Panx was unable to efficiently repress either transposons or the RNA-tethered reporters (Supplementary Fig. 2h,i). However, the mutated Panx lacking the dNxf2 binding region can induce silencing if tethered to DNA^{22,29}. Correspondingly, HP1a (also SetDB1 and Lsd1) can induce heterochromatin formation through DNA tethering, whereas direct tethering of HP1a to nascent RNAs fails to do so^{7,22,30–33}. Therefore, RNA-mediated heterochromatin formation may require the synergistic action of multiple factors rather than just a simple recruitment of general silencing machinery. We demonstrated that the function of dNxf2 in a Pandas complex is essential for piRNA-guided silencing. Our data provide a mechanistic insight that sequestering nascent transcripts in the nucleus may be important for the full establishment of heterochromatin to enforce silencing. By removing dNxf2, we may have uncovered an intermediate state of silencing during heterochromatin formation.

Like any coding mRNA, transposon transcripts would probably be transported into the cytoplasm by the general RNA-exporting machinery (dNxf1–dNxt1) if not restrained by Piwi–piRNAs^{16,25}. In piRNA-guided TGS, dNxf2 may function together with Panx as a stable complex to counteract this process (Fig. 8d). Our structure provides mechanistic insights into why dNxf2^{UBA} prefers to bind the silencing factor Panx rather than the FG repeats of NPCs (Fig. 4). Consistent with our finding, three other groups (Brennecke, Hannon, Siomi and colleagues) independently identified that dNxf2, in complex with Panx, is required for piRNA-guided transposon silencing^{22,23,29}. We further showed that dNxf2 binds directly

to dNxf1 by preventing its association with nucleoporin (Fig. 7g), thereby explaining how dNxf2 blocks dNxf1-mediated general RNA export. As dNxf2 is preferentially associated with the piRNA-targeted transcripts (Fig. 5 and Supplementary Fig. 4), only a subset of dNxf1-associated transcripts could be affected by dNxf2. Accordingly, part of the silencing function of dNxf2 may be locally hijacking the RNA-export machinery and repurposing dNxf1 into a ‘dead-end’ complex, hence trapping transposon transcripts within the nuclear peripheries, most likely still attached to chromatin (Fig. 8c,d). A previous study has shown that both Piwi and NPCs contact chromatin at similar regions by DNA adenine methyltransferase identification (DamID)³⁴. In this regard, dNxf1 has been found to be localized to the nuclear peripheries where most constitutive heterochromatin resides^{15,35–37}. Together, our data suggest that sequestering transposons to the nuclear peripheries via the Pandas complex may help to establish/maintain their heterochromatic state (Fig. 6 and Supplementary Videos 1–4)^{36,37}.

Similar to piRNA-guided silencing, Xist can relocate the silent X chromosome to the nuclear rim during X-chromosome inactivation in mammalian cells, indicating that similar principles may be applied to the formation of facultative heterochromatin marked with H3K27me3 (ref. 38). Recent evidence has demonstrated that hsNxf1 is required for the efficient elongation of RNA polymerase II (ref. 39). It is tempting to speculate that the Pandas complex may also inhibit transcriptional elongation of transposons by neutralizing dNxf1, consistent with the model of piRNA-guided transcriptional repression. In summary, we have uncovered an unexpected role for the transposon RNA-export blockage required for TGS. Our results will have broader implications in understanding how RNA metabolism modulates epigenetic gene silencing and heterochromatin formation (Fig. 8d).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at <https://doi.org/10.1038/s41556-019-0396-0>.

Received: 24 July 2019; Accepted: 22 August 2019;

Published online: 30 September 2019

References

- Ge, D. T. & Zamore, P. D. Small RNA-directed silencing: the fly finds its inner fission yeast? *Curr. Biol.* **23**, R318–R320 (2013).
- Martienssen, R. & Moazed, D. RNAi and heterochromatin assembly. *Cold Spring Harb. Perspect. Biol.* **7**, a019323 (2015).
- Czech, B. & Hannon, G. J. One loop to rule them all: the ping-pong cycle and piRNA-guided silencing. *Trends Biochem. Sci.* **41**, 324–337 (2016).
- Ozata, D. M., Gainetdinov, I., Zoch, A., O’Carroll, D. & Zamore, P. D. PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* **20**, 89–108 (2018).
- Gainetdinov, I., Colpan, C., Arif, A., Cecchini, K. & Zamore, P. D. A single mechanism of biogenesis, initiated and directed by PIWI proteins, explains piRNA production in most animals. *Mol. Cell* **71**, 775–790 (2018).
- Yu, Y. et al. Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science* **350**, 339–342 (2015).
- Siensi, G. et al. Silencio/CG9754 connects the Piwi–piRNA complex to the cellular heterochromatin machinery. *Genes Dev.* **29**, 2258–2271 (2015).
- Siensi, G., Dönertas, D. & Brennecke, J. Transcriptional silencing of transposons by Piwi and Maelstrom and its impact on chromatin state and gene expression. *Cell* **151**, 964–980 (2012).
- Ohtani, H. et al. DmGTSF1 is necessary for Piwi–piRISC-mediated transcriptional transposon silencing in the *Drosophila* ovary. *Genes Dev.* **27**, 1656–1661 (2013).
- Chang, T. H. et al. Maelstrom represses canonical polymerase II transcription within bi-directional piRNA clusters in *Drosophila melanogaster*. *Mol. Cell* **73**, 291–303.e6 (2019).
- Czech, B., Preall, J. B., McGinn, J. & Hannon, G. J. A transcriptome-wide RNAi screen in the *Drosophila* ovary reveals factors of the germline piRNA pathway. *Mol. Cell* **50**, 749–761 (2013).

12. Muerdter, F. et al. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in *Drosophila*. *Mol. Cell* **50**, 736–748 (2013).
13. Handler, D. et al. The genetic makeup of the *Drosophila* piRNA pathway. *Mol. Cell* **50**, 762–777 (2013).
14. Guruharsha, K. G. et al. A protein complex network of *Drosophila melanogaster*. *Cell* **147**, 690–703 (2011).
15. Herold, A., Klymenko, T. & Izaurralde, E. NXF1/p15 heterodimers are essential for mRNA nuclear export in *Drosophila*. *RNA* **7**, 1768–1780 (2001).
16. Katahira, J. Nuclear export of messenger RNA. *Genes* **6**, 163–184 (2015).
17. Port, F., Chen, H.-M., Lee, T. & Bullock, S. L. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc. Natl Acad. Sci. USA* **111**, E2967–E2976 (2014).
18. Braun, I. C., Herold, A., Rode, M. & Izaurralde, E. Nuclear export of mRNA by TAP/NXF1 requires two nucleoporin-binding sites but not p15. *Mol. Cell Biol.* **22**, 5405–5418 (2002).
19. Gu, J. et al. GoldCLIP: gel-omitted ligation-dependent CLIP. *Genom. Proteom. Bioinform.* **16**, 136–143 (2018).
20. Savić, D. et al. CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. *Genome Res.* **25**, 1581–1589 (2015).
21. Iwasaki, Y. W. et al. Piwi modulates chromatin accessibility by regulating multiple factors including histone H1 to repress transposons. *Mol. Cell* **63**, 408–419 (2016).
22. Batki, J. et al. The nascent RNA binding complex SFINX licenses piRNA-guided heterochromatin formation. *Nat. Struct. Mol. Biol.* **26**, 720–731 (2019).
23. Murano, K. et al. Nuclear RNA export factor variant initiates piRNA-guided co-transcriptional silencing. *EMBO J.* **257**, e102870 (2019).
24. Radion, E. et al. Key role of piRNAs in telomeric chromatin maintenance and telomere nuclear positioning in *Drosophila* germline. *Epigenetics Chromatin* **11**, 40 (2018).
25. Chambeyron, S. et al. piRNA-mediated nuclear accumulation of retrotransposon transcripts in the *Drosophila* female germline. *Proc. Natl Acad. Sci. USA* **105**, 14964–14969 (2008).
26. Matzat, L. H., Berberoglu, S. & Lévesque, L. Formation of a Tap/NXF1 homotypic complex is mediated through the amino-terminal domain of Tap and enhances interaction with nucleoporins. *Mol. Biol. Cell* **19**, 327–338 (2008).
27. Aibara, S., Katahira, J., Valkov, E. & Stewart, M. The principal mRNA nuclear export factor NXF1:NXT1 forms a symmetric binding platform that facilitates export of retroviral CTE-RNA. *Nucleic Acids Res.* **43**, 1883–1893 (2015).
28. Combe, C. W., Fischer, L. & Rappsilber, J. xiNET: cross-link network maps with residue resolution. *Mol. Cell Proteom.* **14**, 1137–1147 (2015).
29. Fabry, M. H. et al. piRNA-guided co-transcriptional silencing coopts nuclear export factors. *eLife* **8**, e47999 (2019).
30. Danzer, J. R. & Wallrath, L. L. Mechanisms of HP1-mediated gene silencing in *Drosophila*. *Development* **131**, 3571–3580 (2004).
31. Hines, K. A. et al. Domains of heterochromatin protein 1 required for *Drosophila melanogaster* heterochromatin spreading. *Genetics* **182**, 967–977 (2009).
32. Li, Y., Danzer, J. R., Alvarez, P., Belmont, A. S. & Wallrath, L. L. Effects of tethering HP1 to euchromatic regions of the *Drosophila* genome. *Development* **130**, 1817–1824 (2003).
33. Azzaz, A. M. et al. Human heterochromatin protein 1 α promotes nucleosome associations that drive chromatin condensation. *J. Biol. Chem.* **289**, 6850–6861 (2014).
34. Ilyin, A. A. et al. Piwi interacts with chromatin at nuclear pores and promiscuously binds nuclear transcripts in *Drosophila* ovarian somatic cells. *Nucleic Acids Res.* **45**, 7666–7680 (2017).
35. Kerkow, D. E. et al. The structure of the NXF2/NXT1 heterodimeric complex reveals the combined specificity and versatility of the NTF2-like fold. *J. Mol. Biol.* **415**, 649–665 (2012).
36. van Steensel, B. & Belmont, A. S. Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell* **169**, 780–791 (2017).
37. Towbin, B. D., Meister, P. & Gasser, S. M. The nuclear envelope—a scaffold for silencing? *Curr. Opin. Genet. Dev.* **19**, 180–186 (2009).
38. Chen, C.-K. et al. Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* **354**, 468–472 (2016).
39. Chen, S. et al. The mRNA export receptor NXF1 coordinates transcriptional dynamics, alternative polyadenylation, and mRNA export. *Mol. Cell* **74**, 118–131.e7 (2019).

Acknowledgements

We thank the National Facility for Protein Science in Shanghai Zhangjiang Lab and the Shanghai Science Research Center for their instrumental support and technical assistance. We thank the staff from the BL19U1 beamline at the Shanghai Synchrotron Radiation Facility for assistance with data collection. We thank S. Li from the Center for Biological Imaging and Y. Wang from the Protein Science Core Facility at the Institute of Biophysics, CAS for their technical support and assistance with data collection. This work was supported in part by grants from the Ministry of Science and Technology of China (grant no. 2017YFA0504200 to Y.Y.), the Strategic Priority Research Program of the Chinese Academy of Sciences (grant no. XDB19000000 to Y.Y.), the National Natural Science Foundation of China (grant nos 91640105 and 31770875 to Y.Y., and 91640102 and 31870741 to Y.H.), the National Postdoctoral Program for Innovative Talents (grant no. BX20190081 to Y.H.Z.) and the China Postdoctoral Science Foundation Grant (grant no. 2019M653166 to J.Q.G.).

Author contributions

Y.Y. and Y.H. conceived the project and wrote the manuscript. K.Z. constructed the dNxf2–Halo knock-in OSCs and established the dNxf2 mutant. K.Z., P.X., W.L., X.L. and D.Q. performed the co-immunoprecipitations, tethering assays, transgenic fly constructions, RNA-seq and RT-qPCR experiments. S.C. and X.Y. performed the structural studies, β -gal activity assays and isothermal titration calorimetry experiments. S.C. and Yuhang Zhang performed the GST pull-down assays. K.Z., Z.J., P.Z., X.O., J.G. and P.X. performed the cloning. S.C. and X.L. performed the Y2H assays. K.Z. and X.O. performed the FACS analysis. N.M. performed the FISH and RNAi experiments. M.W. and Yiqun Zhang performed the bioinformatics analysis. C.S., C.P., J.-H.W. and M.-Q.D. performed the mass spectrometry and analysed the data. Y.W., J.M. and H.C. provided critical reagents and advice. All authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41556-019-0396-0>.

Correspondence and requests for materials should be addressed to Y.H. or Y.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

Methods

Fly strains and husbandry. All fly strains were incubated at 25 °C and were used for analysis at the age of 3–5 d. Short hairpin RNA lines targeting *dNxf1* (THU0463), *dNxf2* (THU1328), *Eggless* (THU0892), *Zuc* (THU1831) and *attp2* (TB00072) were obtained from the Tsinghua Fly Center. Short hairpin RNAs against *dNxt1* (GL00414), *nanos-GAL4* (25751), *dNxf2* deficiency strain (25124) and *actin-Cas9* (54590) were obtained from the Bloomington Stock Collection. An equal number of females with the indicated genetic backgrounds were mated to wild-type OrgerR males with fresh yeast paste for fertility tests. The pupae were checked one week later. The *dNxf2* mutants were generated by CRISPR-Cas9 as described¹⁷. The guide sequences of the gRNAs cloned into pCFD4 (Addgene, cat. no. 49411) were CAGATGAGAGTGCTGGACTT and GATTGCAAGATCTTCAGCAA.

Cell culture. Ovarian somatic cells were maintained at 25 °C as described previously⁴⁰.

Plasmid construction, generation of transgenic flies and an OSC stable line.

Fragments spanning different structural domains of *Drosophila melanogaster* *Panx*, *dNxf2*, *dNxf1* and *dNxt1* were cloned into the pENTR4 vector (Thermo Fisher Scientific, A10465) as described previously⁶. To generate the 10×BoxB reporter constructs, a fragment containing RanGAP-T2A-Puro-10×BoxB-tRNA was cloned into the pENTR4 vector and subsequently recombined into the UGW (DGRC1283) or URW (DGRC1282) vectors. The resulting vector was transfected into OSCs and selected with puromycin. The cells that were GFP-positive were sorted by FACS to obtain the ubiquitin-GFP-10×BoxB reporter cell line. To generate the λN-tethering constructs, a fragment containing the ubiquitin promoter-RFP-blasticidin-polyA was inserted into the UNF-Gateway destination vector⁶. A ubiquitin-RFP fragment was inserted into the Ubi-λN-3×FLAG-Gateway. Then, pENTRY vectors were constructed containing the following open reading frames: *Panx*, *dNxf2*, *dNxf2*-ΔRBD (residues 1–554), *dNxf2*-Δ1st RBD (residues 1–285), *dNxf2*-Δ2nd RBD (residues 286–554), *dNxf2*-ΔNTF2 (residues 585–758), *dNxf2*-ΔUBA (residues 762–841), *dNxf2*-mutant (F826A + I827A), *dNxf1*, *dNxf1*-ΔRBD (residues 1–345), *dNxf1*-ΔNTF2 (residues 366–545) and *dNxf1*-ΔUBA (residues 618–672). A promoter (Piwi)-λN-FKBP-T2A-Bla fragment was inserted into the Ubi-FRB-HA-Gateway. Finally, pENTRY vectors containing the indicated open reading frames (*Panx* or *dNxf2*) were recombined into the appropriate destination vectors. The resulting plasmids were used for the tethering experiments in OSCs or transgenic flies. Transgenic flies were generated by micro-injection of the indicated constructs into fly embryos (UniHuaii) using either the P-element or the PhiC31 system. For the split *Gaussia* luciferase assay, the pENTRY vectors containing either *dNxf1*-ΔRBD or *dNxf2*^{NTF2} were recombined with the destination plasmids containing either *Gluc1* or *Gluc2* complementary DNA fragments of *Gaussia* luciferase⁴¹.

The OSC Halo knock-in stable lines were generated as described previously²⁰. The correct integration was confirmed by genomic PCR and western blotting or Halo-ligand staining. The gRNA sequences cloned into pCFD4 (Addgene, cat. no. 49411) were GGTATGTGTAGACGTTTCGT and GAAGCTCCTGCTCTTCAAG.

GoldCLIP. Both *dNxf2*-Halo and Halo-alone OSC stable lines were transfected with either siH1 + siHP1a or control siRNAs as described before²¹. For the Maelstrom knockdowns, DNA templates containing T7 promoters were in vitro transcribed using T7 polymerase⁴². The purified double-stranded RNAs were co-transfected with plasmids overexpressing RFP-Dicer2 into OSCs using FuGENE HD. The cells were harvested 72 h after electroporation and Halo pull down was performed as described previously, except that no nuclease was added to the lysates¹⁹. The RNAs were reverse transcribed using random primers and the resulting cDNAs were quantified using the pre-validated qPCR primers (Supplementary Table 2). For the GoldCLIP-sequencing experiments, we followed the eCLIP protocol to construct libraries following Halo pull down⁴³.

Luciferase assays. Dissected fly ovaries with the indicated genotypes were ground independently with the passive lysis buffer (Promega) and measured as previously described⁶. For the fold-change calculations, the luminescence values of *Renilla* or total-protein-normalized Firefly luciferase were divided by that of the control (using a Wilcoxon rank-sum test).

The split *Gaussia* luciferase assays were performed as previously described⁴¹. *Drosophila* S2 cells were transfected using the FuGENE HD transfection reagent (Promega, E2311). The cells were lysed using the Pierce *Gaussia* luciferase flash assay kit (Thermo Fisher Scientific, cat. no. 16159) 2 d after transfection. Luciferase activity was measured with the luminometric measurement optics set to 'normal' by Varioskan Flash (Thermo Fisher Scientific) and normalized at OD₂₈₀ using a second aliquot of the cell lysates by Multiskan Go (Thermo Fisher Scientific).

Tethering assay in OSCs and flow-cytometry analysis. For the tethering experiments in OSCs, GFP-10×BoxB stable lines, vectors expressing different λN-fusion-protein constructs as well as Ubi-RFP were transfected in six-well plates using FuGENE reagent (Promega, E2311) following the manufacturer's

instructions. The pPiwi-λN-FKBP-T2A-Bla/Ubi-FRB-HA-ORF dual expression vectors (*Panx* or *dNxf2*) were co-transfected with Ubi-RFP plasmid at a ratio of 10:1 in six-well plates for the rapamycin-inducible tethering experiments. After transfection (24 h), the cells were treated with 20 nM rapamycin for 24 h before FACS. All analytical procedures were performed on a FACSAria (Becton Dickinson). Data for a minimum of 10,000 events of RFP high-signal cells were collected 4 d after the indicated transfections using a 488 nm (GFP) and a 561 nm (RFP) laser for excitation. The data were analysed on FlowJo X software (Tree Star). For the RNA FISH experiments with rapamycin-inducible tethering, we transfected either the pPiwi-λN-FKBP-T2A-Bla/Ubi-FRB-HA-*dNxf2* plasmid or the FRB-alone control plasmid into GFP-10×BoxB reporter OSCs and selected with blasticidin (10 μg ml⁻¹) until the cells doubled at a normal rate. The stable cell lines were then transfected again with the indicated λN-FKBP-FRB tethering plasmids and treated with 20 nM rapamycin or DMSO control for 2 h before fixing with 4% formaldehyde to visualize the GFP mRNAs.

RNA isolation and qPCR. Total RNA was isolated from dissected ovaries with TRIzol (Thermo Fisher Scientific) and treated with the TURBO DNA-free kit (Ambion). Complementary DNA was synthesized using SuperScript III reverse transcriptase (Thermo Fisher Scientific) and random hexamers. The resulting cDNAs were diluted tenfold to carry out a quantitative PCR assay (SsoAdvanced SYBR Green supermix and CFX384, BioRad). The primer sequences are listed in Supplementary Table 2.

RNA-seq and GoldCLIP analysis. RNA-seq was performed and analysed as described⁴⁴. For the GoldCLIP data analysis, reads shorter than 25 nucleotides were discarded. Potential PCR duplicates in the GoldCLIP libraries were removed using the random barcodes at positions 1–10 of the read. The processed reads were mapped against the spike-in human genome (hg38, download from the UCSC) using the Bowtie⁴⁵ (v1.1.2; bowtie -v 2 -k 1 -best -un -S -p 8 -f), a library of mature transfer RNAs and ribosomal RNAs, and the mitochondrial genome. The remaining unmapped reads were mapped against the *D. melanogaster* genome (dm3, BDGP release 5) using Bowtie (options -v 2 -m 1 -best -strata -S -p 8 -f) and transposon consensus sequences using STAR (v2.5.2a)⁴⁶. For plotting, read-counts were normalized to one million uniquely mapping reads and converted to bigwig tracks using Deeptools⁴⁷ (v3.1.3).

Antibodies, co-immunoprecipitations and western blotting. Rabbit polyclonal antibodies were raised against either the VPIQLDYPDSKIFSKSSYG (N terminus) or CLKLFIQKNADHEIPDLAFA (C terminus) peptides of *dNxf2*, and CVKLSRALAPALAIKGRQP (C terminus) of *Panx* following affinity purification using the corresponding epitope. A 1:2,000 dilution was used for the western blots. The antibodies used in this study include rabbit anti-Piwi (Abcam, ab5207), mouse anti-FLAG (clone M2; Sigma, F3165) and rabbit anti-GFP (Torrey Pines Scientific, TP401). GFP-Trap nanobody magnetic beads were obtained from Allele Biotech (ABP-NAB-GFPM100).

Ovarian somatic cells were seeded in six-well plates at approximately 1 × 10⁶ cells ml⁻¹ and transfected with the FuGENE HD transfection reagent (Promega, E2311). The cells were harvested 48–60 h post transfection and lysed as previously described⁴⁰. The cell lysates were incubated overnight at 4 °C with pre-blocked magnetic beads (Allele biotech) using 1 mg ml⁻¹ BSA for immunoprecipitation. The beads were washed three times in the washing buffer (10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.1% Triton and 1 mM phenylmethylsulfonyl fluoride (PMSF)) and eluted at 95 °C in the washing buffer supplemented with SDS. Images of the western blots were taken using an Odyssey imaging system.

Co-immunoprecipitation and mass-spectrometry analysis. About 200 ovaries were dissected and dounce homogenized in 1 ml lysis buffer (1% Triton X-100, 50 mM Tris-HCl, pH 7.4, 50 mM NaCl, 2 mM CaCl₂, 5% glycerol, 2 mM dithiothreitol, 0.1 mM PMSF, protease inhibitor Cocktail (Roche) and 5 U μl⁻¹ MNase (New England Biolabs, M0247S)). The lysates were incubated on ice for 15 min and cleared by centrifugation at 16,000g for 10 min at 4 °C. About 200 μl GFP-Trap MA bead slurry (Allele Biotech) was used for each immunoprecipitation. The samples were rotated overnight at 4 °C and washed seven times with a buffer containing 10 mM Tris-HCl, pH 7.4, 500 mM NaCl, 0.1% Triton X-100 and 0.1 mM PMSF. Proteins precipitated on the beads were boiled in the sample buffer and analysed by one-dimensional gel electrophoresis. The gel was cut, sliced and subjected to in-gel digestion with trypsin. The mass spectrometry analysis of the purified samples was performed by liquid chromatography with tandem mass spectrometry (LC-MS/MS) using a Linear Ion Trap mass spectrometer (Thermo Fisher Scientific) equipped with a nano-spray source and Eksigent high-performance liquid chromatography. The buffers used in the liquid chromatography separation consisted of buffer A (1.95% acetonitrile, 97.95% H₂O and 0.1% formic acid) and buffer B (79.95% acetonitrile, 19.95% H₂O and 0.1% formic acid). The liquid chromatography separation was achieved by linear gradient elution from 5 to 50% buffer B over 80 min at a flow rate of 300 nl min⁻¹. The Linear Ion Trap mass spectrometer was operated in the data-dependent mode. A full-scan survey mass spectrometry experiment was performed (m/z range of

400–2,000, automatic gain control target of 5×10^5 ions, resolution of 100,000 at 400 m/z). Only the five most abundant ions were analysed by MS2 scan events (automatic gain control target 1×10^4 ions, maximum ion accumulation time of 200 ms). The normalized collision energy was 35%. The peak areas of fragment ions were used to calculate the relative intensity of the precursor ions for the selected peptides. At least two peptides were selected for the quantification of one protein. The means of the relative intensity of selected peptides represented the relative expression level of the proteins. The average enrichments between bait and control samples were calculated using a custom R script. Adjusted *P* values were calculated using the R package *limma*⁴⁸.

GST pull-down assay. Purified GST fusion proteins including GST–dNxf1^{NFT2+UBA} (residues 359–672)/dNxt1, GST–dNxf1^{NFT2} (residues 359–544)/TS–dNxt1, GST–dNxf1^{UBA} (residues 601–672), GST–dNup214^{FG} (residues 1445–1516) and GST–dNup358^{FG} (residues 2395–2426) were immobilized on Glutathione-Sepharose beads (GE Healthcare) for the GST pull-down assays. Purified GST protein was set as a control. To explore the interactions between dNxf1 and dNxf2, the dNxf2^{NFT2} (residues 555–772) and dNxt1 complex was added in increasing quantities and incubated with GST–dNxf1 fragments or GST in a binding buffer (20 mM Tris–HCl, pH 8.0, 500 mM NaCl, 10% glycerol and 0.1% Triton X-100) for 2 h with rotation at room temperature. For the pull-down assay of the dNxf2^{NFT2+UBA}–(GS)₄–Panx^{NIR} and dNxt1 complex, GST–dNxf1^{NFT2+UBA} and GST–dNxf1^{NFT2} were used. After four washes with the binding buffer, the bound proteins were finally eluted with the binding buffer supplied with 10 mM glutathione reduced form. The eluates were subjected to 15% SDS–PAGE followed by Coomassie blue staining. To test the binding ability of dNxf1 and dNxf2 to the Nup proteins, 25 µg GST control or GST–Nup^{FG} protein were first incubated with 30 µl Glutathione-Sepharose beads (GE Healthcare) at room temperature for 2 h. The beads were washed three times with the binding buffer and then incubated with 100 µg of either the dNxf1^{NFT2+UBA}–dNxt1 or dNxf2^{NFT2+UBA}–dNxt1 complex for 2 h at room temperature. An additional four washes were performed using the binding buffer. The bound fractions were boiled in sample buffer and subjected to 15% SDS–PAGE, followed by Coomassie blue staining. For the competition binding assay of dNxf2 versus dNxf1–GST–Nup^{FG}, 25 µg GST control or GST–Nup214^{FG} proteins were incubated with 30 µl Glutathione-Sepharose beads (GE Healthcare) for each reaction at room temperature for 2 h. Next, 100 µg of the dNxf1^{NFT2+UBA}–dNxt1 complex was incubated with the beads pre-loaded with GST–Nup214^{FG} for each reaction and washed with the binding buffer. The beads were aliquoted and incubated with buffer or an increasing amount of purified His–SUMO–dNxf2^{NFT2+UBA}–(GS)₄–Panx^{NIR}–dNxt1 complex at room temperature for another 2 h. After four washes, each bound fraction was boiled in sample buffer and subjected to 15% SDS–PAGE followed by Coomassie blue staining.

Chemical cross-linking of proteins coupled with mass spectrometry. The purified complexes (10 µg), dNxf1^{NFT2+UBA}–dNxt1 and dNxf2^{NFT2}–dNxt1, were incubated with 1 mM disuccinimidyl suberate (DSS; Thermo Fisher Scientific, A39267)—in a reaction buffer containing 50 mM HEPES, pH 7.5 and 500 mM NaCl—or 4 mM 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC; Thermo Fisher Scientific, PG82073)—in a reaction buffer containing 50 mM 2-(*N*-morpholino)ethanesulfonic acid, pH 5.2 and 500 mM NaCl—at room temperature for 2 h. To terminate the reaction, 0.5 M ammonium bicarbonate was added to a final concentration of 20 mM. The reaction was further incubated at room temperature for 20 min. The samples were precipitated by acetone and digested with trypsin for 16 h at 37 °C at an enzyme-to-substrate ratio of 1:50 (w/w). The trypsin-digested peptides were desalted and loaded onto an in-house packed capillary reverse-phase C18 column (40 cm length, 100 µm inner diameter, 360 µm outer diameter, 1.9 µm particle size, 120 Å pore diameter) connected to an Easy LC 1200 system. The samples were analysed with a 120 min high performance liquid chromatography gradient from 6% to 35% of buffer B (buffer A, 0.1% formic acid in H₂O and buffer B, 0.1% formic acid in 80% acetonitrile) at 300 nl min^{−1}. The eluted peptides were ionized and directly introduced into a Q-Exactive mass spectrometer using a nano-spray source. Survey full-scan MS spectra (*m/z* 300–1,800) were acquired in an Orbitrap analyzer with resolution *r* = 70,000 at *m/z* 400.

Cross-linked peptides were identified and evaluated using pLink software⁴⁹. The pLink search parameters were similar to what had been previously described⁵⁰: instrument, HCD; precursor mass tolerance, 20 ppm; fragment mass tolerance, 20 ppm; cross-linker, DSS (mass-shift 138.0680796, mono-link mass-shift 156.0786442), EDC (mass-shift −18.01056); fixed modification, C 57.02146; peptide length, minimum of 6 amino acids and maximum of 60 amino acids per chain; peptide mass, minimum of 600 and maximum of 6,000 Da per chain; enzyme, trypsin, with up to three missed cleavage sites per chain. The results were filtered using the following requirements: FDR < 5%, E-value < 0.00001 and spectra count ≥ 1. The MS2 spectra were annotated using pLabel1, with a mass deviation ≤ 20 ppm required. The interactive maps of the cross-link network were drawn using xINet²⁸.

DNA FISH, RNA FISH, three-dimensional-SIM super-resolution microscopy and image analysis. DNA FISH for Het-A was performed as previously

described²⁴. RNA FISH for GFP was performed as previously described⁵¹. The three-dimensional-structured-illumination-microscopy (SIM) images of the *Drosophila* ovaries were acquired on the DeltaVision OMX V3 imaging system (GE Healthcare) with a ×100 1.40 NA (numerical aperture) oil objective (Olympus UPlanSApo), solid-state multimode lasers (488 nm, 405 nm and 561 nm) and electron-multiplying CCD (charge-coupled device) cameras (Photometrics, Evolve 512 × 512). Serial z-stack sectioning was done at 125 nm intervals for the SIM mode and 250 nm intervals for the conventional mode. Immersion oils with refractive indices of 1.512 were used for dissected ovaries on glass coverslips to obtain optimal images. The microscope was routinely calibrated with 100 nm fluorescent spheres to calculate both the lateral and axial limits of the image resolution. The SIM image stacks were reconstructed using softWoRx 6.1.1 (GE Healthcare) with the following settings: pixel size, 39.5 nm; channel-specific optical transfer functions; Wiener filter constant, 0.0010; discard Negative Intensities background; drift correction with respect to first angle and custom K0 guess angles for camera positions. Conventional image stacks were processed by deconvolution methods using softWoRx 5.0 (GE Healthcare) with the following settings: Wiener filter enhancement, 0.900 and winner filter smoothing, 0.800. The reconstructed images were further processed for maximum-intensity projections with softWoRx 6.1.1. Pixel registration was corrected to be less than one pixel for all channels using 100 nm Tetraspek beads. The three-dimensional renders and image analysis were finished using Imaris software. First, a new surface to match the cell nucleus was constructed and then the FISH signal was captured as spots by diameter. The system recognised the nucleus as a cell and calculated the distance between the DNA FISH signals and the nuclear membrane as the distance between the spots and the cell membrane. The nuclear periphery was stained with lamin antibody (1:100; DSHB, ADL67–10) and second antibody (1:400; A11001). The RanGAP–GFP stable OSC lines were treated with 3% H₂O₂ for 30 min at room temperature immediately following fixation to bleach the GFP signal.

Protein expression and purification. The *D. melanogaster* Nxf2 UBA domain fragment (residues 762–841) was cloned into a modified pCold vector between the EcoRI and XhoI sites, which carried a 6×His–GST tag at the N terminus. The Panx fragment (residues 315–356), dNxf1^{NFT2} domain (residues 359–544), dNxf1^{NFT2+UBA} domain (residues 359–672), full-length dNxt1, dNup214^{FG} (residues 1445–1516) and dNup358^{FG} (residues 2395–2426) of *D. melanogaster*, and the covalently linked constructs dNxf2^{UBA}–(GS)₄–Panx^{NIR} (dNxf2^{UBA}, residues 788–841 and Panx^{NIR}, residues 315–343) were then cloned into the pGEX-6p vector. Full-length dNxt1, TS–dNxt1 (fusion with a Twin-Strep tag at the N terminus), dNxf2^{NFT2+UBA} (residues 569–841) and dNxf2^{NFT2+UBA}–(GS)₄–Panx^{NIR} (dNxf2^{NFT2+UBA}, residues 569–841 and Panx^{NIR}, residues 315–343) were cloned into the pET28-SMT3 vector carrying a 6×His–SUMO tag at the N terminus.

Recombinant proteins were expressed in the *Escherichia coli* strain BL21(DE3) Gold (Agilent). For coexpression, two plasmids, each carrying a protein gene, were transformed into the same competent cell. Cultures were grown in LB media at 37 °C to an OD₆₀₀ of 0.6. After induction with 0.2 mM isopropyl-β-D-thiogalactoside, the cells were cultured at 18 °C overnight. Selenomethionine derivative cells (L795M) were cultured in M9 media as previously described⁵².

The cells were harvested by centrifugation at 4 °C and disrupted by French Press (JNBio) in buffer A (20 mM Tris–HCl, pH 8.0, 500 mM NaCl, 25 mM imidazole and 1 mM PMSF) for His-tagged proteins or buffer S (20 mM Tris–HCl, pH 8.0, 500 mM NaCl, 1 mM EDTA, 0.1% 2-mercaptoethanol and 1 mM PMSF) for GST-tagged proteins. The bacterial lysates were clarified by centrifugation at 16,000g for 1 h. The supernatant of recombinant His-tagged proteins was loaded onto a His-Trap HP column (GE Healthcare). After extensive washing with buffer A, the target protein was eluted using buffer A with a 25–500 mM imidazole gradient. For the recombinant GST-tagged proteins, the supernatant was loaded onto a GSTrap HP column (GE Healthcare). After extensive washing with buffer S, the target protein was eluted using buffer S supplemented with 10 mM glutathione reduced form (GSH). The 6×His–SUMO, GST and 6×His–GST tags were cleaved at 4 °C using Ulp1 protease or 3C protease. The target proteins were further purified by passage through a HiLoad 16/60 Superdex G75 column (GE Healthcare) with buffer GF (10 mM Tris–HCl, pH 8.0, 100 mM NaCl and 1 mM dithiothreitol). The fractions containing the target proteins were pooled and concentrated for further use.

Crystallization and data collection. The native and selenomethionine derivative of dNxf2^{UBA}–(GS)₄–Panx^{NIR} was pooled and concentrated (15 mg ml^{−1}). Crystals were grown from a solution containing 30% (w/v) PEG 3350, 0.1 M Tris–HCl, pH 9.0 and 0.2 M sodium acetate trihydrate at 17 °C using the hanging-drop vapour-diffusion method. The dNxf1^{359–544}–dNxt1 complex (12 mg ml^{−1}) was used for crystallization. Crystals were grown from a reservoir solution containing 16% (w/v) PEG 8,000, 40 mM K₂HPO₄ and 20% glycerol at 17 °C using the hanging-drop vapour-diffusion method. The crystals were flash frozen in liquid nitrogen in the above reservoir solution supplemented with glycerol for data collection. The diffraction datasets were collected at a wavelength 0.9778 Å for the dNxf2–Panx crystals and 0.9785 Å for the dNxf1–dNxt1 crystals on a beamline BL19U1 at the Shanghai Synchrotron Radiation Facility and processed using the HKL-3000 (ref. 53).

Structure determination and refinement. The L795M selenomethione derivative dataset of the dNxf2^{UBA}–(GS)₄–Panx^{NIR} was collected at 1.7 Å. The overall completeness was 98.8%, with 91.0% of the reflections measured at the highest-resolution shell (1.76–1.70 Å). AutoSol, as implemented in the Phenix program suite, was used for selenium site location and phase calculation⁵⁴. The initial figure of merit was 0.48. The final figure of merit reached 0.72 after density modification. The overall architecture of the molecule was then clearly displayed. The initial model was built using the program AutoBuild, also embedded in the Phenix program suite. The interactive model building was carried out using Coot⁵⁵. The structure was further refined in Phenix against the 1.5 Å native dataset. The final R_{work} and R_{free} are 0.164 and 0.173, respectively (Supplementary Table 1). The structure of the dNxf1 NTF2 domain and dNxt1 complex was determined by molecular replacement using the structure of the human TAP and p15 complex (PDB accession code 1JKG). The structure was refined in Phenix until the crystallographic statistics were within the reasonable range (Supplementary Table 1).

The model of the dNxf1 UBA domain was generated by the SWISS-MODEL server⁵⁶. The model of the dNxf2 NTF2 domain was generated in Coot by mutating the corresponding residues in the dNxf1–dNxt1 structure. The graphics were generated using PyMOL (The PyMOL Molecular Graphics System, version 1.6.9, Schrödinger, LLC). Electrostatic potentials were calculated in PyMOL and the surface was coloured continuously from red to blue (–70 to +70 keV). Sequence alignments were analysed using Clustal W⁵⁷ and presented by Espript 3.0 (ref. ⁵⁸).

Y2H assay. The Y2H assays were performed using L40 strains containing the pBTM116 and pACT2 (Clontech) fusion plasmids. Fragments of *D. melanogaster* dNxf2 (residues 1–841, 1–554, 555–841, 555–787 and 762–841), dNxf1 fragments (residues 1–672, 1–360, 359–672, 359–612 and 601–672) were cloned into the pBTM116 vector. Fragments of *D. melanogaster* Panx (residues 1–541, 1–229, 230–541, 230–315, 315–356, 346–541 and 315–343) and the full-length dNxt1 were cloned into the pACT2 vector. All mutants were generated using a site-directed mutagenesis kit (New England Biolabs) according to the manufacturer's instructions. The colonies containing both plasmids were selected on –Leu–Trp plates. To quantify the protein–protein interactions, β-galactosidase assays were performed as previously described⁵⁹.

Analytical size-exclusion chromatography. A Superdex G75 increase column (GE Healthcare) was used at a flow rate of 0.4 ml min^{–1} in buffer containing 20 mM Tris–HCl, pH 8.0 and 500 mM NaCl at 4 °C. The purified and filtrated samples were injected at a concentration of 10 mg ml^{–1} in the running buffer.

Isothermal-titration-calorimetry assay. Isothermal titration calorimetry was used to measure the binding affinities of dNxf2^{UBA} with Panx^{NIR}. All measurements were performed at 25 °C using an iTC200 (Microcal). The samples were centrifuged before each experiment to remove any protein precipitates. Panx^{NIR} was titrated into the experimental cell, which contained 0.1 μM dNxf2^{UBA}, in twenty 2 μl injection increments. The titration data were analysed using the Origin 8.0 software (Microcal Software).

Statistics and reproducibility. All statistical analyses were performed using GraphPad Prism 7 and Excel 2010. Results are shown as the mean ± s.d., as described in the figure legends. The Wald test *P* values were calculated using a two-tailed Student's *t*-test and a cut-off of 0.05 was used to identify statistical significance. Data were collected from individual flies for the tethering experiments ($n = 10$ or 15 biologically independent replicates of distinct samples per indicated condition); the Wilcoxon rank-sum test was used to calculate the *P* values for $n = 15$. Three biologically independent replicates of distinct samples were measured in parallel for the Y2H assays. All experiments were reproduced at least three times independently unless otherwise indicated in the legend.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw files of all sequencing libraries generated for this study have been submitted to the Gene Expression Omnibus under accession number GSE130042. The coordinates for the structures reported in this paper have been deposited in the PDB under accession number 6IEW (dNxt2^{UBA} and Panx^{NIR} complex) and 6IHJ (dNxt1^{NTF2} and dNxt1 complex). The proteomics data of the binding proteins for Panx and the cross-link mass spectrometry data have been deposited in ProteomeXchange with the primary accession codes PXD014926 and PXD014884. The structure data collection and refinement statistics have been provided as Supplementary Table 1. The DNA oligo sequences have been provided as Supplementary Table 2. The source data for Figs. 1–7 and Supplementary Figs. 1,2,4,5 have been provided as Supplementary Table 3. All other data supporting the findings of this study are available from the corresponding authors on request.

References

- Saito, K. et al. A regulatory circuit for *piwi* by the large Maf gene *traffic jam* in *Drosophila*. *Nature* **461**, 1296–1299 (2009).
- Kolkhof, P. et al. A luciferase-fragment complementation assay to detect lipid droplet-associated protein–protein interactions. *Mol. Cell. Proteom.* **16**, 329–345 (2017).
- Chamberlin, M., McGrath, J. & Waskell, L. New RNA polymerase from *Escherichia coli* infected with bacteriophage T7. *Nature* **228**, 227–231 (1970).
- Van Nostrand, E. L. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).
- Armour, C. D. et al. Digital transcriptome profiling using selective hexamer priming for cDNA synthesis. *Nat. Methods* **6**, 647–649 (2009).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
- Ritchie, M. E. et al. *limma* powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
- Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- Lu, S. et al. Mapping native disulfide bonds at a proteome scale. *Nat. Methods* **12**, 329–331 (2015).
- Choi, H. M. T., Beck, V. A. & Pierce, N. A. Next-generation in situ hybridization chain reaction: higher gain, lower cost, greater durability. *ACS Nano* **8**, 4284–4294 (2014).
- Doublié, S. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **276**, 523–530 (1997).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
- Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
- McWilliam, H. et al. Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* **41**, W597–W600 (2013).
- Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
- Mehla, J., Caufield, J. H., Sakhawalkar, N. & Uetz, P. A comparison of two-hybrid approaches for detecting protein–protein interactions. *Methods Enzymol.* **586**, 333–358 (2017).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

HKL-3000 for crystal structures. The RNA-seq data were obtained using illumina standard procedures on HiSeq X Ten and HiSeq 2500. Basecalling were performed using bcl2fastq v1.8.4.

Data analysis

Phenix-1.13 for structural study. The scripts used in this manuscript for RNA-seq datasets were deposited on Github with the following address: https://github.com/bakerwm/2018_nxf2_paper_scripts. In brief, the RNA-seq reads were processed using Cutadapt (v1.16) and aligned to D. melanogaster reference genome (dm3) using STAR (v2.5.2a). Using featureCounts (v1.6.1) and DESeq2 (v1.20.0) for quantification and differential expression analysis. The related plots were generated by in-house R scripts, including ggplot2 (v3.0.0), limma (v3.36.5) packages.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The coordinates for the structures reported in this paper have been deposited in the Protein Data Bank under accession number 6IEW (dNxt2UBA and PanxNIR complex) and 6IHJ (dNxt1NTF2 and dNxt1 complex). Raw files of all sequencing libraries generated for this study were submitted to GEO under accession number GSE130042.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For luciferase assays, we measured 15 individual samples per condition.
Data exclusions	no
Replication	three independent biological replicates
Randomization	no
Blinding	no

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	commercial antibodies: rabbit anti-piwi (Abcam, ab5207), mouse anti-Flag (clone M2, Sigma, F3165), rabbit anti-GFP (Torrey Pines Scientific, TP401). mouse anti-LAM (DSHB #ADL67) Antibodies raised in this study: rabbit anti-dNxf2-N; rabbit anti-dNxf2-C; rabbit anti-Panx.
Validation	LAM (DSHB #ADL67): validated in immunostaining (1:100) and western blotting (1:1000) to identify nuclei in Cell, 2014, 159(4):829-43. Flag (F3165): validated in multiple studies including ChIP, ELISA, EMSA, FC/FACS, ICC, ICC-IF, IF, IHC, IP, In Vitro, PLA, and WB GFP (TP401): validated in multiple studies including ChIP, ELISA, IF, IP, PLA, and WB Piwi (ab5207): gave a positive signal in Drosophila embryo nuclear extract. dNxf2-N, dNxf2-C and Panx rabbit antibodies were validated as follows: 1. Affinity purified using corresponding peptide; 2. Western blotting using mutant ovaries lysates; 3. Western blotting using a GFP-tagged fusion protein.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Ovarian Somatic Cell
Authentication	obtained from Professor Siomi, University of Tokyo
Mycoplasma contamination	We checked mycoplasma contamination once every month using the PCR and never found contaminations
Commonly misidentified lines (See ICLAC register)	no

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Drosophila Melanogaster
Wild animals	no
Field-collected samples	no
Ethics oversight	No Ethical approval is required for this study

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Disassociate the cell with Trypsin, and deactivate the Trypsin with OSC medium.
Instrument	BD FACSaria IIIu , Becton,Dickinson, 5820092313
Software	FlowJo X, Tree Star
Cell population abundance	minimum 10,000 events at the 10E4-10E5 axis scales of Pe-TxR-A cells were collected. Usually RFP positive cells is ~10% of total population.
Gating strategy	Using the polygon to specify the FSC/SSC cells, and specify the non-stick cells in FSW/SSW gate. Specifying 10E4-10E5 axis scales of Pe-TxR-A as the RFP-high positive cells and 0-10E3 of Pe-TxR-A as the RFP-negative cell within non-stick cells. Finally specify the RFP-high cells and change the axis from RFP-high to FITC-A (GFP), analyzing the fluorescence of FITC-A between experimental sample, negative control and GFP-alone cells.

- ☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.