

# Engineering a transposon-associated TnpB- $\omega$ RNA system for efficient gene editing and phenotypic correction of a tyrosinaemia mouse model

Received: 26 May 2023

Accepted: 17 January 2024

Published online: 27 January 2024

 Check for updates

Zhifang Li<sup>1,10</sup>, Ruochen Guo<sup>1,2,10</sup>, Xiaozhi Sun<sup>1,3,10</sup>, Guoling Li<sup>4,10</sup>, Zhuang Shao<sup>1</sup>, Xiaona Huo<sup>1,5</sup>, Rongrong Yang<sup>1,5</sup>, Xinyu Liu<sup>2</sup>, Xi Cao<sup>1,6</sup>, Hainan Zhang<sup>4</sup>, Weihong Zhang<sup>4</sup>, Xiaoyin Zhang<sup>1,5</sup>, Shuangyu Ma<sup>7</sup>, Meiling Zhang<sup>8</sup>, Yuanhua Liu<sup>2</sup>, Yinan Yao<sup>2</sup>, Jinqi Shi<sup>1</sup>, Hui Yang<sup>2,4,5</sup>, Chunyi Hu<sup>9</sup> , Yingsi Zhou<sup>4</sup>  & Chunlong Xu<sup>1,3,5</sup> 

Transposon-associated ribonucleoprotein TnpB is known to be the ancestry endonuclease of diverse Cas12 effector proteins from type-V CRISPR system. Given its small size (408 aa), it is of interest to examine whether engineered TnpB could be used for efficient mammalian genome editing. Here, we showed that the gene editing activity of native TnpB from *Deinococcus radiodurans* (ISDra2 TnpB) in mouse embryos was already higher than previously identified small-sized Cas12f1. Further stepwise engineering of noncoding RNA ( $\omega$ RNA or reRNA) component of TnpB significantly elevated the nuclease activity of TnpB. Notably, an optimized TnpB- $\omega$ RNA system could be efficiently delivered in vivo with single adeno-associated virus (AAV) and corrected the disease phenotype in a tyrosinaemia mouse model. Thus, the engineered miniature TnpB system represents a new addition to the current genome editing toolbox, with the unique feature of the smallest effector size that facilitate efficient AAV delivery for editing of cells and tissues.

The TnpB proteins represent a family of transposon-associated RNA-guided endonucleases. Recent biochemical studies<sup>1,2</sup> revealed that TnpB proteins are ancestry predecessors of Cas12 effector proteins in the type-V CRISPR system, and a 247-nucleotides (nt) noncoding RNA (termed  $\omega$ RNA or reRNA) derived from the right end of transposon

element is the required component for ISDra2 TnpB to recognize and cleave target DNA. The size of TnpB proteins, with ~400 amino acid (aa) residues, is much smaller than their evolutionary progeny Cas12 proteins (mostly ~1000 aa). Furthermore, in vitro studies<sup>1,2</sup> demonstrated that TnpB exhibited double-strand DNA cleavage activity

<sup>1</sup>Lingang Laboratory, Shanghai, China. <sup>2</sup>Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China.

<sup>3</sup>School of Life Sciences and Technology, ShanghaiTech University, Shanghai, China. <sup>4</sup>HuidaGene Therapeutics Inc, Shanghai, China. <sup>5</sup>Shanghai Center for Brain Science and Brain-Inspired Technology, Shanghai, China. <sup>6</sup>College of Animal Science and Technology, Northwest A&F University, Yangling, China.

<sup>7</sup>Department of Histoembryology, Genetics and Developmental Biology, Shanghai Key Laboratory of Reproductive Medicine, Shanghai JiaoTong University School of Medicine, Shanghai, China. <sup>8</sup>Center for Reproductive Medicine, International Peace Maternity and Child Health Hospital, Innovative Research Team of High-level Local Universities in Shanghai, School of Medicine, Shanghai Jiao Tong University, Shanghai, China. <sup>9</sup>Department of Biological Sciences, National University of Singapore, Singapore, Singapore. <sup>10</sup>These authors contributed equally: Zhifang Li, Ruochen Guo, Xiaozhi Sun, Guoling Li.

 e-mail: [hu\\_dbs@nus.edu.sg](mailto:hu_dbs@nus.edu.sg); [yingsizhou@huidagene.com](mailto:yingsizhou@huidagene.com); [xucl@lglab.ac.cn](mailto:xucl@lglab.ac.cn)

guided by  $\omega$ RNA. Therefore, there is potential for the use of this TnpB system in genome editing and therapeutic applications.

Gene editing using Cas9 or Cas12 systems has been widely used in animal models and recently applied in clinical trials. At present, AAV is the most commonly used delivery system and shown to be safe in gene therapy<sup>3</sup>. However, the maximal cargo size of AAV was limited to be 4.7 kilobase (kb) pairs, hindering efficient *in vivo* delivery of the large Cas9 or Cas12 protein via single AAV injection. This size problem is exacerbated in the use of base and prime editors comprising Cas9 (or Cas12) and fusion enzymes. Recent identification of compact CRISPR effector proteins Cas12f1 (~500 aa)<sup>4</sup> and Cas13 (~700 aa)<sup>5,6</sup> represent potential solutions. However, the gene editing efficiency of Cas12f1 was relatively low<sup>7–11</sup>, whereas Cas13 exhibited collateral RNA cleavage activity with uncertain safety profile<sup>12,13</sup>.

In this study, we demonstrated that genome editing activity of TnpB was markedly higher than that of Cas12f1 in cultured cells and mouse embryos. To further optimize the TnpB system, we engineered TnpB-associated  $\omega$ RNA in a stepwise manner to identify the optimal  $\omega$ RNA variant with the shortest sequence length and elevated gene editing activity. Importantly, we showed that the optimized TnpB- $\omega$ RNA system could be effectively delivered *in vivo* via a single AAV injection in tyrosinaemia model mice, leading to the correction of the disease phenotype. Thus, we have shown the applicability of the engineered hypercompact TnpB for genome editing *in vivo*.

## Results

### TnpB exhibited gene editing activity higher than Cas12f1

Previous studies have shown the endonuclease activity of several Cas12f1 orthologs from type-V-U CRISPR family that have small sizes. As the ancestry enzyme of Cas12 proteins, TnpB (~400 aa) represents the smallest programmable nuclease among common single effector Cas proteins, including SpCas9, LbCas12a, Un1Cas12f1, and IscB (Fig. 1a). However, the mammalian genome editing potential of TnpB remained to be fully characterized. Thus, we selected several genomic loci to evaluate the editing activity of ISDra2 TnpB (hereafter as TnpB) from *Deinococcus radiodurans* in mouse embryos. First, we *in vitro* transcribed  $\omega$ RNA that targets the mouse *Tyr* gene (Fig. 1b), and injected  $\omega$ RNA together with TnpB mRNA into mouse embryos. The injected embryos were then transferred into surrogate female mice to generate gene-modified offspring. Since *Tyr* gene encodes the black coat color of C57BL/6 mice, we estimated the efficiency of TnpB-induced gene disruption by directly examining the coat color change in TnpB-injected mice. We found that TnpB treatment completely converted black coat color into albino white in all newborn mice (Fig. 1c). In contrast, similar embryo injection of Un1Cas12f1 together with sgRNA targeting the *Tyr* gene did not change the black coat color in the newborn mice (Fig. 1c), suggesting a much lower *Tyr* gene disruption efficiency of Un1Cas12f1 than that of TnpB. Further deep-sequencing for *Tyr* gene showed that 20% and 90% of indel mutations were induced by Un1Cas12f1 and TnpB, respectively (Fig. 1b and Supplementary Fig. S1). Although Cas12f1 and TnpB have different requirements for target adjacent motif (TAM, also known as PAM) that recognizes the target sequence, we have chosen the targeted sequence in *Tyr* gene to have 17-bp overlap (among 20 bp) for both enzymes (Fig. 1b). Thus, the higher editing efficiency of TnpB as compared to Cas12f1 was largely due to its intrinsic activity.

To further evaluate the gene editing activity of TnpB, we chose six additional loci in the mouse *Dmd* gene (Fig. 1d) for targeting in mouse embryos, by injecting  $\omega$ RNA targeting these loci with TnpB mRNA. As shown by deep-sequencing results, TnpB exhibited an average of 90% editing efficiency for all six targeted loci in the *Dmd* gene (Fig. 1d and Supplementary Figs. S2 and S3). Furthermore, the gene editing outcome was verified by immunostaining of dystrophin protein encoded by *Dmd* gene that is specifically expressed in muscle tissues. In contrast to wild-type mice, TnpB-treated mice showed undetectable dystrophin

expression in heart, Diaphragm (DI) and Tibialis anterior (TA) muscles (Fig. 1e and Supplementary Fig. S4), suggesting the complete disruption of *Dmd* gene by TnpB and  $\omega$ RNA injection. Finally, these immunostaining results were confirmed by western blotting of dystrophin protein of various muscle tissues (Fig. 1f). Consequently, rotarod and grip strength assessment of TnpB-treated DMD mice found functional dysfunction of muscle (Supplementary Fig. S5). Thus, our finding indicated more robust gene editing activity of TnpB than that of Un1Cas12 f1 in mammalian tissues.

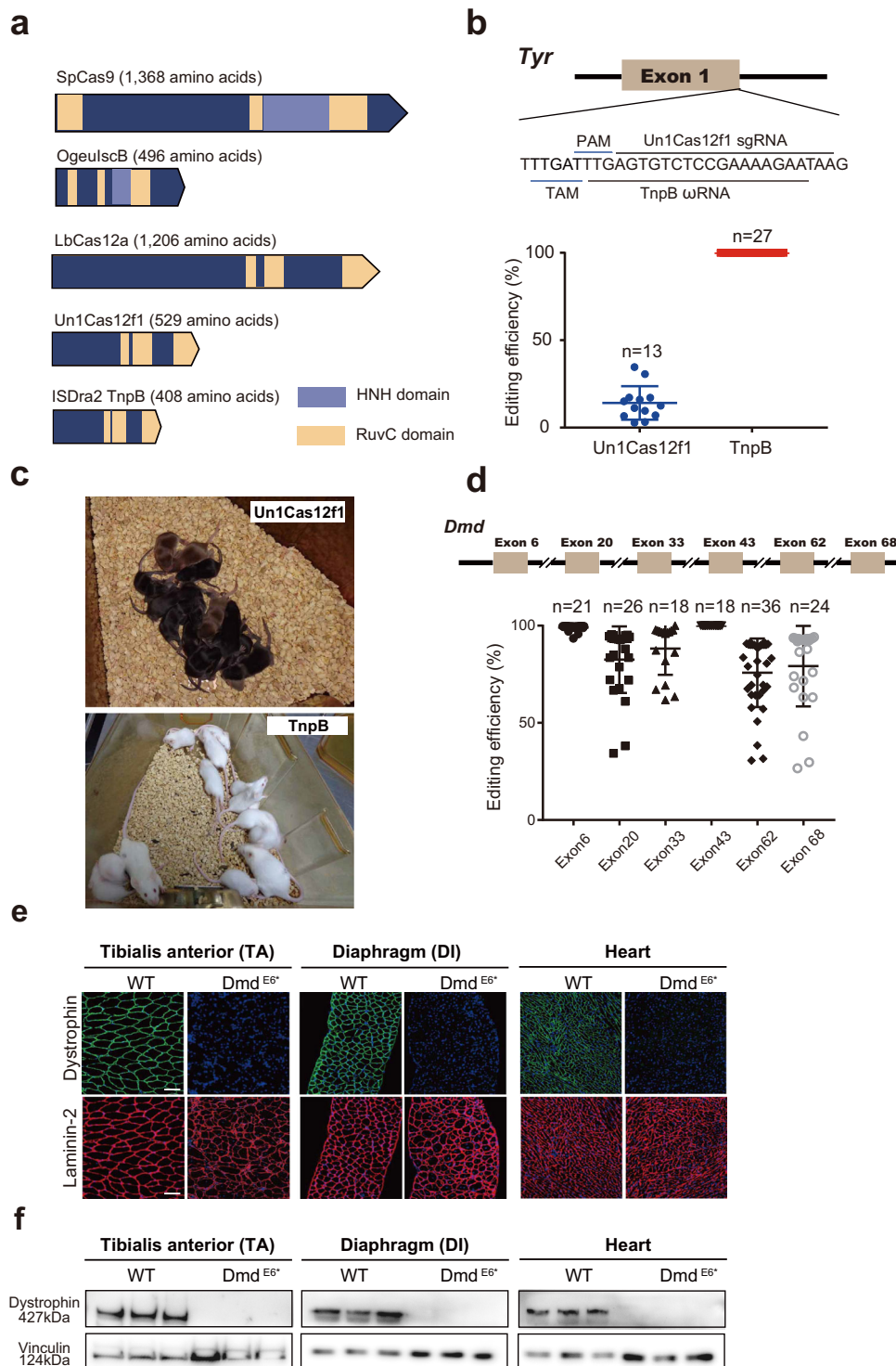
### Engineered TnpB-associated $\omega$ RNA with elevated editing efficiency

Cognate  $\omega$ RNA scaffold associated with TnpB is 231 nt, much longer than sgRNA scaffold for most single effector Cas proteins. Previous findings reported that the sgRNA engineering could improve the performance of gene editing enzymes<sup>14</sup>. We thereby hypothesized that  $\omega$ RNA truncation and optimization might be helpful for enhancing TnpB activity in mammalian cells. To this end, we predicted the secondary structure of  $\omega$ RNA and formulated a stepwise strategy to truncate  $\omega$ RNA (Fig. 2a). Based on the stem loops in predicted structure, we divided  $\omega$ RNA into six segments, named as S1 to S6 for the truncation experiment (Fig. 2b). To facilitate screen of  $\omega$ RNA variants, we designed a gene editing reporter with TnpB target DNA placed within a split and frameshift GFP gene which could only be repaired via the single-strand annealing (SSA) pathway<sup>15</sup> after disruption of TnpB target sequence to express GFP (Fig. 2a). We tested the reporter with cognate  $\omega$ RNA to prove the conditional activation of GFP after treatment of TnpB guided by  $\omega$ RNA targeting frameshift mutation in GFP gene (Fig. 2a). At first, we deleted S1 to S6 one by one and run the reporter assay. It showed that only deletion of S4 and S6 ablated the activity of TnpB (Fig. 2c), suggesting the dispensable role of S1, S2, S3, and S5 for normal  $\omega$ RNA function. Furthermore, sequence deletion of S1 slightly increase TnpB activity (Fig. 2c).

To interrogate combined deletion effect of S1 to S6, we added S2 to S5 deletion in the S1 deletion variant of  $\omega$ RNA to conduct reporter assay. It found that simultaneous deletion of S1, S2, and S3 in  $\omega$ RNA-v1 not only supported the normal function of TnpB but also significantly increased the gene editing efficiency (Fig. 2d). These results implied that the  $\omega$ RNA sequence from S4 to S6 dictated the enzymatic activity of TnpB. Secondary structure of  $\omega$ RNA after combined truncation of S1, S2, and S3 showed typical stem-loop conformations with three distinguishable and consecutive stem-loop (SL) domains, termed as SL1, SL2, and SL3 (Fig. 2e). To further determine the effect of these three SL domains on TnpB activity, we iteratively remove SL1, SL2, and SL3 for reporter test. In addition, we also generated two other  $\omega$ RNA variants with partial deletion of SL2 subdomain or substitution of G:U with G:C pairs (Fig. 2e). We found that SL1, SL2 and SL3 are necessary for the normal function of TnpB since deletion variants lack of any single SL fully blocked the reporter activation (Fig. 2f). However, partial replacement of SL3 subdomain with 5'-GAAA-3' loop sequence actually enhanced the TnpB activity (Fig. 2f). G:C substitution for G:U pair exhibited no additive effect on the performance of TnpB (Fig. 2f). Based on these results, we finally identified an optimal  $\omega$ RNA variant  $\omega$ RNA-v2 or  $\omega$ RNA\* that improved TnpB performance. The predicted secondary structure of  $\omega$ RNA\* presented with three compact stem-loop domains in contrast to the loose organization of cognate  $\omega$ RNA structure (Fig. 2g).

### Characterization of endogenous gene editing and off-target activity for TnpB- $\omega$ RNA system

To verify the reporter assay results for  $\omega$ RNA\*, we selected 14 endogenous genomic loci for further evaluation of gene editing performance in HEK293T (Fig. 3a). Among 14 human loci tested, 10 individual target sites showed significant increase of TnpB gene editing efficiency with  $\omega$ RNA\* compared to original  $\omega$ RNA (Fig. 3b). Moreover, we also

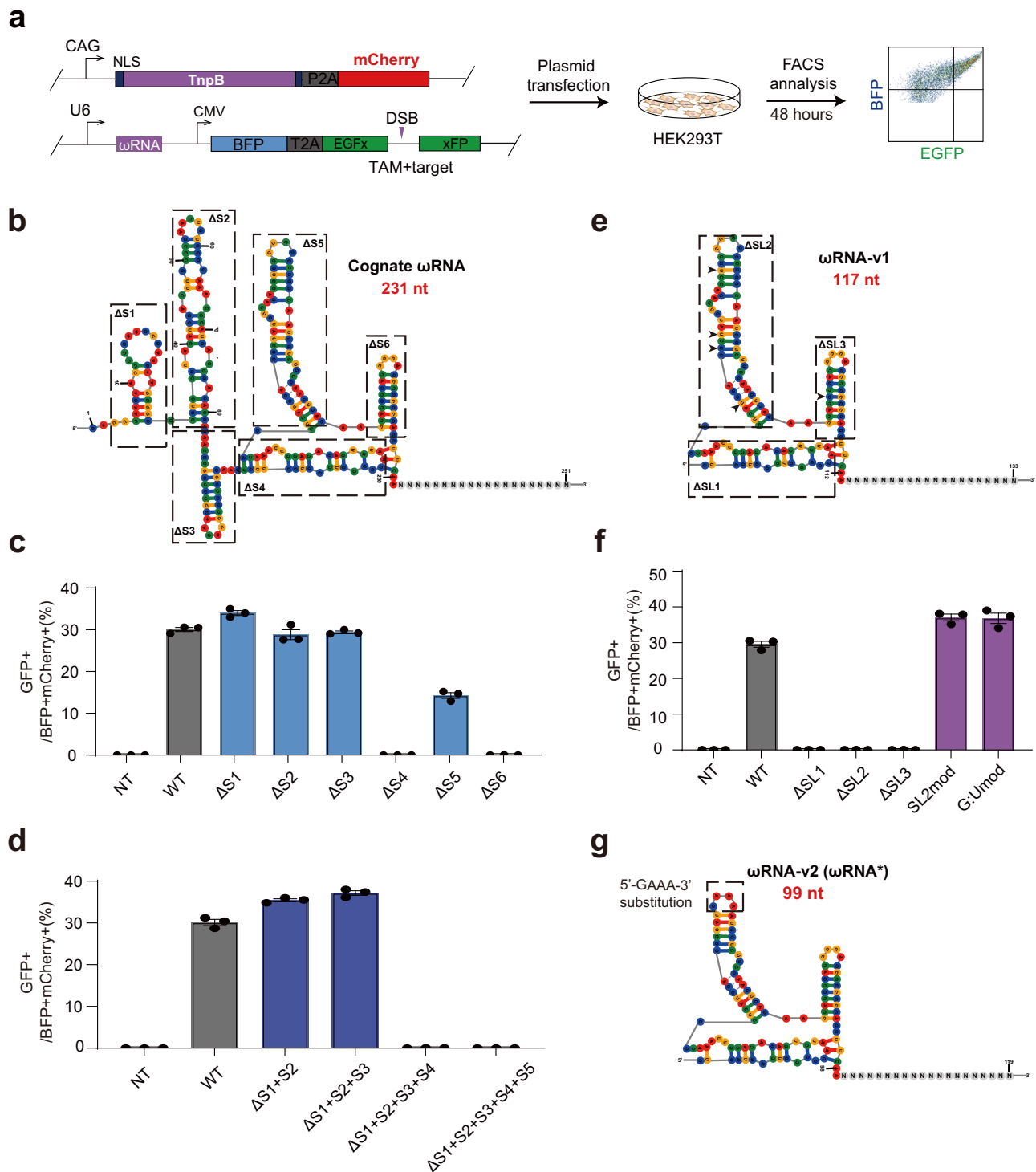


**Fig. 1 | Mouse embryo injection of TnpB- $\omega$ RNA induced efficient gene editing.** **a** Characteristics of SpCas9, IscB, LbCas12a, Un1Cas12f1 and ISDra2 TnpB nucleases. **b** Comparison of editing efficiency between TnpB and Cas12f1 on *Tyr* gene in mice. **c** Coat color phenotype of *Tyr* gene-edited mice by Un1Cas12f1 and TnpB. **d** TnpB-mediated gene editing efficiency for *Dmd* gene. **e** Dystrophin and laminin-2 immunostaining for TA, DI, and heart muscle tissues in wild-type and *Dmd*-edited

mice by TnpB. **f** Western blotting of dystrophin and vinculin protein for three muscle tissues in wild-type and *Dmd*-edited mice by TnpB. Data are represented as means  $\pm$  SEM. A dot represents a biological replicate ( $n = 3$  or more). Unpaired two-tailed Student's  $t$  tests. Significant differences between conditions are indicated by asterisk (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , NS non-significant). Scale bars, 100  $\mu$ m. Source data are provided as a Source Data file.

compared the gene editing activity of TnpB- $\omega$ RNA and - $\omega$ RNA\* with that of SaCas9 and SpCas9 in both mouse N2a and human HEK293T cells. Our results showed that TnpB- $\omega$ RNA\* exhibited similarly high efficiency with SaCas9 but only slightly lower activity than SpCas9 (Fig. 3c and Supplementary Figs. S6–9). To investigate broad

improvement effect of  $\omega$ RNA\* in mammalian cells, we further performed the gene editing in mouse N2a cells targeting four disease-relevant genes, including *Klkb1*, *Tyr*, *Hpd* and *Pcsk9*. It found that all genomic sites exhibited significantly increased gene editing efficiency for  $\omega$ RNA\* compared to the original  $\omega$ RNA (Supplementary Fig. 10).



**Fig. 2 | Stepwise engineering of TnpB-associated ωRNA improved gene editing efficiency.** **a** Reporter assay schematics of detecting cleavage activity of TnpB-ωRNA. **b** Predicted secondary structure of cognate ωRNA (231 nt). Cognate ωRNA was divided into 6 segments, named from S1 to S6. **c** Reporter assay results using engineered ωRNA by one-by-one truncation of S1 to S6. **d** Reporter assay results with engineered ωRNA by different combined truncations of S1 to S5. **e** Predicted

secondary structure of a ωRNA-v1 variant with simultaneous truncation of S1, S2, and S3. **f** Reporter assay results for ωRNA variants with different SL deletion and modifications. **g** Predicted secondary structure of final optimized ωRNA-v2 (ωRNA\*) variant. Data are represented as means ± SEM. A dot represents a biological replicate ( $n = 3$ ). Source data are provided as a Source Data file.

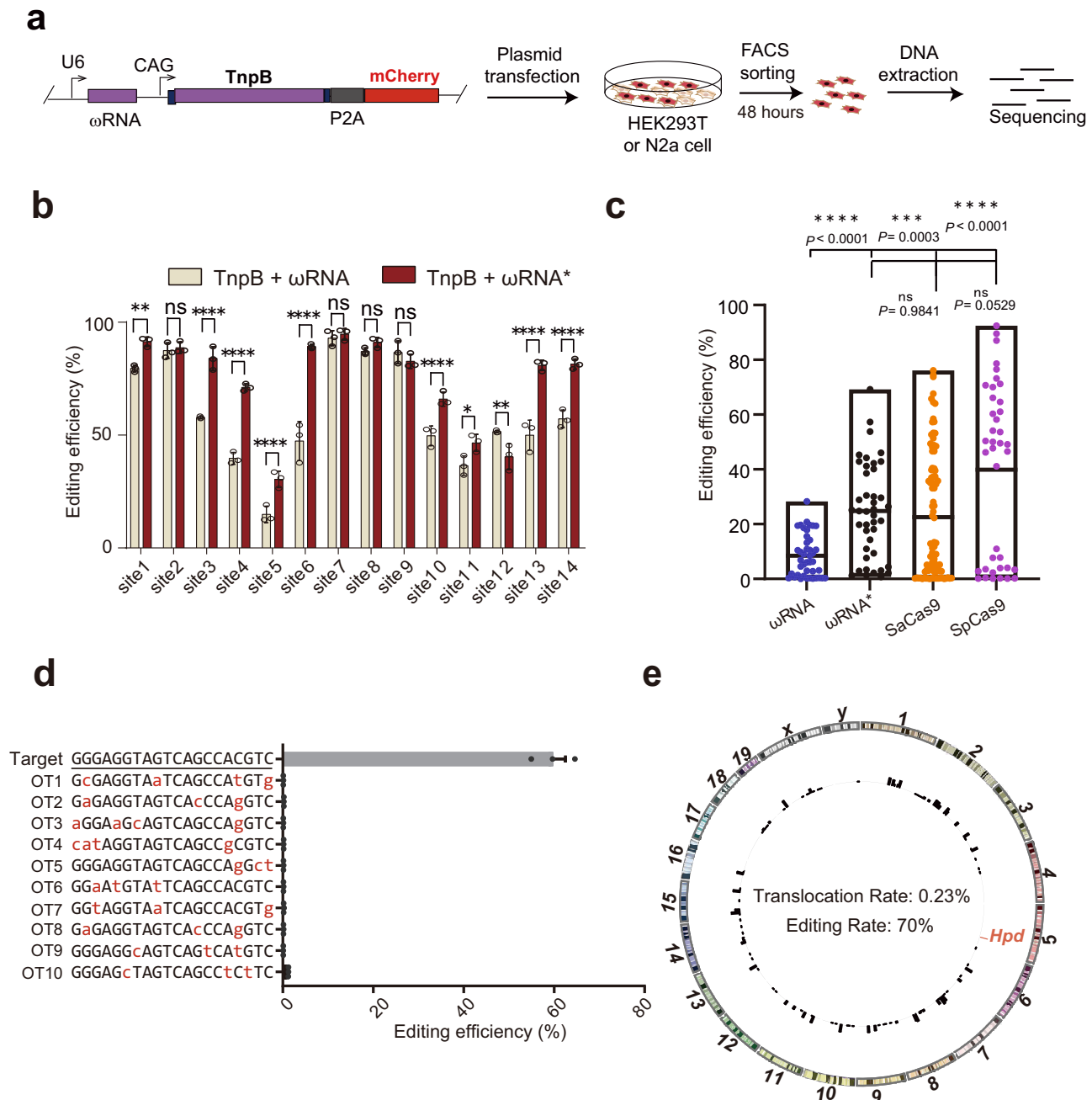
Quantitative analysis revealed twofold increase of gene editing efficiency in N2a for ωRNA\* versus wild-type ωRNA (Supplementary Fig. 10). In particular, the ωRNA\* even supported TnpB editing of some loci that are edited with very low efficiency using the cognate ωRNA (Supplementary Fig. 10). Furthermore, we packaged TnpB-ωRNA and

ωRNA\* with AAV to evaluate their gene editing efficiency in mice, which also confirmed the significantly improved performance with TnpB-ωRNA\* in vivo (Supplementary Fig. 11). Therefore, we demonstrated the enhanced TnpB activity in both mammalian cells and mice via the identification of ωRNA\* after stepwise engineering.

To examine the off-target effect of TnpB, we carried out prediction of potential off-target genomic loci with Cas-OFFinder<sup>16</sup> for off-target analysis when designing  $\omega$ RNA against a target site in *Hpd* gene (Fig. 3d). For the top ten predicted off-target sites, no gene editing events was detected for TnpB- $\omega$ RNA targeting *Hpd* (Fig. 3d). Furthermore, we also performed genome-wide off-target analysis by PEM-seq<sup>17</sup> to identify potential translocation between on-target and off-target loci. Our PEM-seq results showed that there is no induction of translocation events related to gene editing of *Hpd* gene by the engineered TnpB- $\omega$ RNA treatment (Fig. 3e).

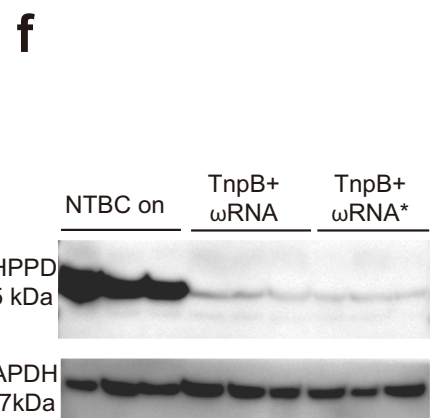
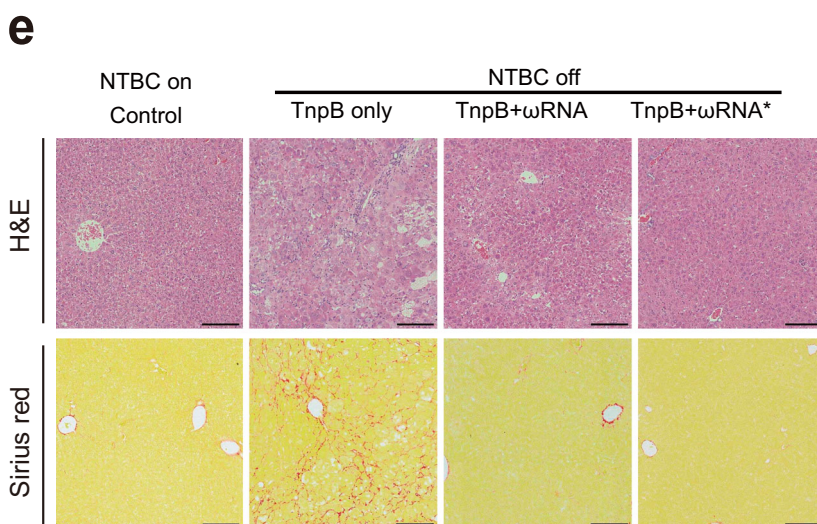
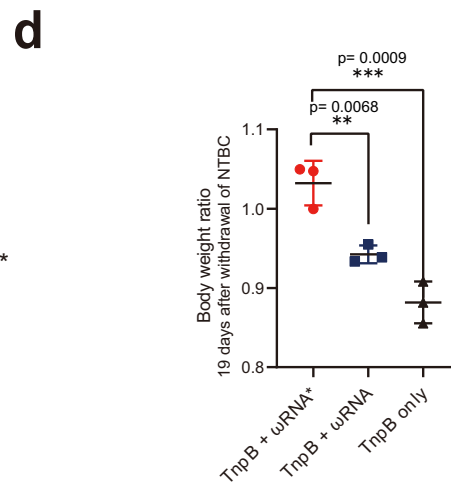
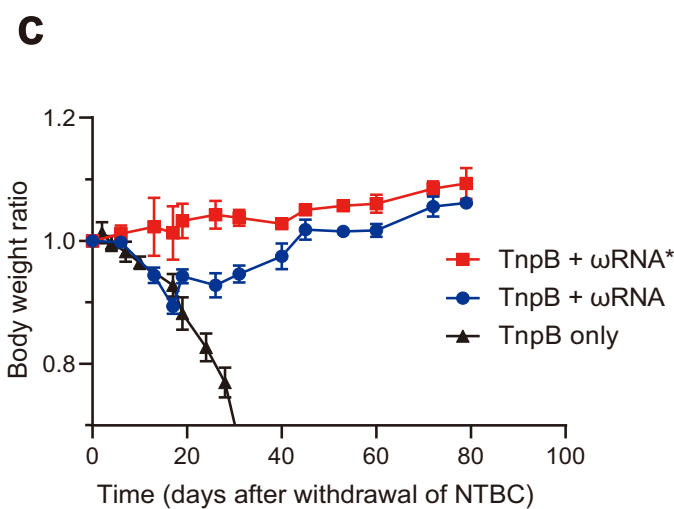
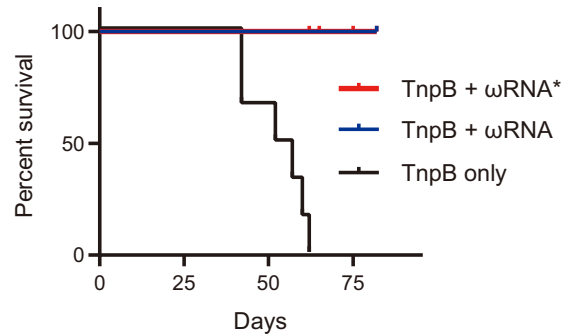
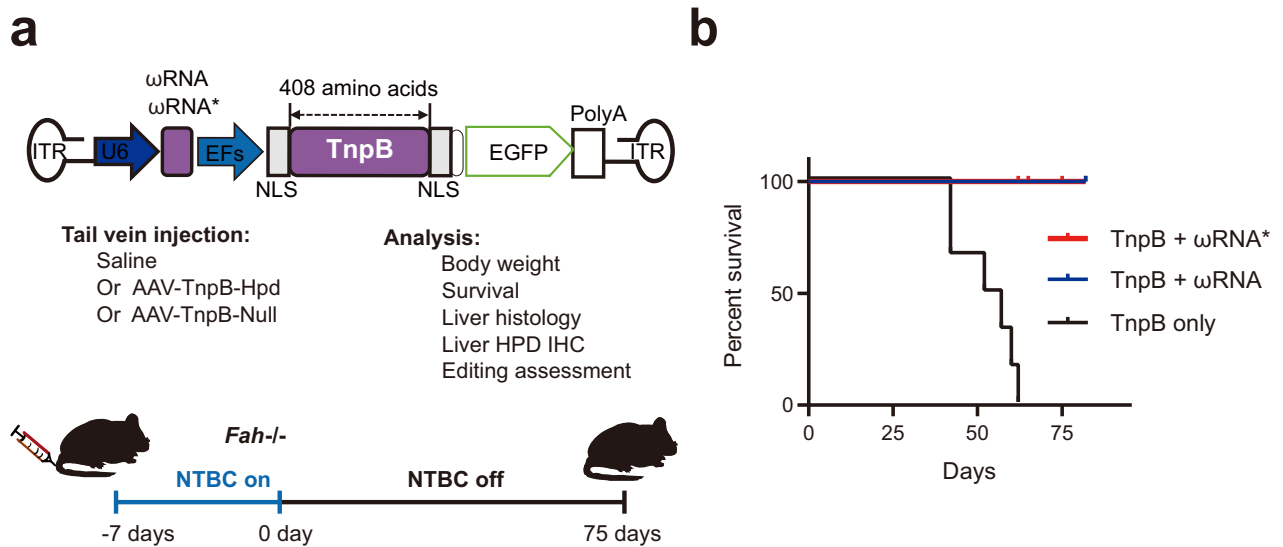
### Correction of fatal liver disease with in vivo delivery of TnpB- $\omega$ RNA via single AAV

Given the hypercompact size of TnpB, it would greatly facilitate in vivo delivery via single AAV for gene editing therapy. To demonstrate the potential of TnpB in disease intervention, we chose the *Hpd* as therapeutic target for gene editing therapy of type I hereditary tyrosinaemia (HT1) in *Fah*<sup>-/-</sup> mouse model. Adult *Fah*<sup>-/-</sup> was administrated with AAV-TnpB (TnpB only without  $\omega$ RNA), AAV-TnpB- $\omega$ RNA, or AAV-TnpB- $\omega$ RNA\* (Fig. 4a) and kept without nitisinone (NTBC) drug, an inhibitor of 4-hydroxyphenylpyruvate dioxygenase (HPD) to examine



**Fig. 3 | Characterization of endogenous gene editing activity and off-target effect with optimized TnpB- $\omega$ RNA system.** **a** The experimental workflow for detecting editing efficiency of original and optimized TnpB- $\omega$ RNA in HEK293T cells. **b** Gene efficiency comparison results for 14 human endogenous gene loci targeted by wild-type and optimized TnpB- $\omega$ RNA in HEK293T. **c** Gene editing activity comparison among TnpB- $\omega$ RNA, TnpB- $\omega$ RNA\*, SaCas9 and SpCas9. All data points with median at center line and 25th and 75th quartile lines. Detailed

statistical results in Source Data file. **d** Off-target analysis for top predicted off-target genomic loci via Cas-OFFinder. **e** Genome-wide off-target analysis with PEM-seq for the engineered TnpB- $\omega$ RNA. Data are represented as means  $\pm$  SEM. A dot represents a biological replicate ( $n = 3$  or more). Unpaired two-tailed Student's  $t$  tests. Significant differences between conditions are indicated by asterisk ( $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ ,  $****P < 0.0001$ , NS non-significant.). Source data are provided as a Source Data file.



**Fig. 4 | Correction of fatal liver disease with in vivo delivery of TnpB- $\omega$ RNA via single AAV. a** Diagram of AAV-TnpB- $\omega$ RNA/ $\omega$ RNA\* vector and gene therapy schematics in *Fah*<sup>-/-</sup> mouse model of type I hereditary tyrosinaemia. **b** Survival curve for disease mice treated with AAV-TnpB- $\omega$ RNA/ $\omega$ RNA\* or AAV-TnpB without  $\omega$ RNA (TnpB only group). **c** Body weight change during the observation period for disease mice in different treatment groups. **d** Body weight ratio for TnpB- $\omega$ RNA or TnpB only versus TnpB- $\omega$ RNA\*-treated mice in 19-day after NTBC withdrawal.

**e** Histology analysis with H&E and Sirius red staining for mouse liver from different treatment groups. **f** Western blot for HPD protein from untreated and TnpB- $\omega$ RNA-treated HT1 mice. Data are represented as means  $\pm$  SEM. A dot represents a biological replicate ( $n = 3$  or more). Unpaired two-tailed Student's *t* tests. Significant differences between conditions are indicated by an asterisk (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , NS non-significant.). Scale bars, 200  $\mu$ m. Source data are provided as a Source Data file.

the therapeutic effect after gene editing. We observed that AAV-TnpB- $\omega$ RNA- and - $\omega$ RNA\*-treated *Fah*<sup>-/-</sup> mice was still alive after 75 days without NTBC, but all mice in TnpB only group died at about 65 days (Fig. 4b). Furthermore, *Fah*<sup>-/-</sup> mice subjected to AAV-TnpB- $\omega$ RNA treatment gained body weight after experiencing a short period of weight loss (Fig. 4c, d). Notably, mice with AAV-TnpB- $\omega$ RNA\* treatment did not experience weight loss after NTBC withdrawal (Fig. 4c, d), indicating better therapeutic efficacy than that of wild-type TnpB- $\omega$ RNA. In addition, we found that UniCas12f1-treatment could also rescue HT1 mice but with significantly decreased therapeutic efficacy compared to AAV-TnpB- $\omega$ RNA and - $\omega$ RNA\* in terms of body weight loss (Supplementary Fig. 12). Contrarily, mice treated without  $\omega$ RNA exhibited rapid weight loss until death (Fig. 4c and Supplementary Fig. 12). These results suggested the importance for improving gene-editing activity with the engineered  $\omega$ RNA to increase the therapeutic effect for disease correction. To examine the histological correction by gene editing treatment, we performed H&E and Sirius red staining to find dramatically reduced fibrotic pathology in TnpB- $\omega$ RNA/ $\omega$ RNA\* treated mice, whereas massive liver fibrosis in mice treated without  $\omega$ RNA (Fig. 4e).

Furthermore, we also analyzed the HPD expression in treated versus untreated mice. It showed the remarkable decrease of HPD protein (Fig. 4f) and large HPD-negative liver region in AAV-TnpB- $\omega$ RNA-treated mice (Supplementary Fig. 13a, b). To investigate the in vivo gene editing outcomes, we collected liver tissue from mice for deep-sequencing analysis. We found 20% and 60% of indel rate for AAV-TnpB- $\omega$ RNA- and AAV-TnpB- $\omega$ RNA\*-treated mice in 1 month, respectively (Supplementary Fig. 13c, d). Consistently, liver metabolic functions were significantly ameliorated after AAV-TnpB- $\omega$ RNA treatment as indicated by the blood biochemical profiling results of alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin and tyrosine (Supplementary Fig. 14a–d). Moreover, we analyzed the succinylacetone level in TnpB- $\omega$ RNA-treated mice and confirmed that TnpB- $\omega$ RNA treatment largely reduced succinylacetone in both plasma and urine of treated mice (Supplementary Fig. 14e, f), consistent with phenotypic correction of HT1 mice after TnpB- $\omega$ RNA administration. Therefore, our results showed the proof-of-concept for applying the engineered TnpB- $\omega$ RNA system in disease control via single AAV delivery in vivo.

## Discussion

Diverse CRISPR-Cas systems evolved from immune battle between microbe and mobile genetic elements (MGE), providing us abundant resources for the identification of gene editing enzymes<sup>18</sup>. In the past years, various single effector Cas proteins including Cas9<sup>19</sup>, Cas12<sup>20</sup>, and Cas13<sup>21</sup> were found to deploy DNA or RNA editing activity in different organisms for both research and therapeutic purpose<sup>22</sup>. Recently, TnpB-like proteins, including IscB and TnpB associated with microbe transposon element, were identified to be active ancestry endonuclease for Cas9 and Cas12<sup>1,2</sup>. Given the hypercompact size of TnpB and IscB, they are excellent candidates for developing miniature gene editing tools that would facilitate in vivo delivery via AAV. To this end, our present study demonstrated the potential of TnpB for robust genome editing in both cultured cells and animal tissues. Although Kim et al. recently

reported engineering base editor from a 557-aa “TnpB”<sup>23</sup>, both Siksny and Doudna group lately demonstrated that “TnpB” used by Kim et al. study should be actually annotated as Cas12f1 that works as dimer unlike monomer TnpB<sup>24,25</sup>. Thus, our work was the first study to extensively show the rational optimization of TnpB to achieve excellent in vitro and in vivo performance for gene editing. Furthermore, we also showed the effectiveness of TnpB based gene editing therapy to correct fatal genetic disease in mouse model of tyrosinaemia via in vivo single AAV delivery of TnpB and  $\omega$ RNA. Interestingly, we performed stepwise truncation of cognate  $\omega$ RNA to generate a  $\omega$ RNA variant with short sequence and high efficiency. Our study represents a good start point to optimize TnpB or even IscB for more broad and convenient use in research and therapeutic scenario.

Endonuclease activity of TnpB was only shown with limited data in 2021 by Karvelis et al. study<sup>1</sup>. Extensive characterization of TnpB activity in mammalian cell and tissue were currently needed. Our finding corroborated the results from Karvelis et al. study, revealing unexpected higher activity of TnpB than that of Cas12f1 without further engineering. Moreover, we showed that deletion of 5'-end and partial internal sequence in  $\omega$ RNA could enhance the gene editing performance of TnpB both in vitro and in vivo. Intriguingly, such deletion strategy was supported by two structural studies<sup>26,27</sup> of TnpB- $\omega$ RNA-DNA ternary complex published lately, suggesting the potential useful applicability of our  $\omega$ RNA engineering strategy for more TnpB-like systems. In addition, the TnpB structure could accelerate the rational engineering of such compact enzyme with more demanding properties such as relaxed limitation of target-adjacent motif (TAM), enhanced editing activity and specificity etc.

Gene editing therapy was partly impeded by the limited AAV cargo capacity of only ~4.7 kb considering the fact that common Cas9, Cas12 and their derived base or prime editors have protein size beyond 1000 aa<sup>3,28</sup>. TnpB with less than 500 aa are highly desired gene editing enzymes for AAV delivery in vivo. Our results with TnpB in treating fatal tyrosinaemia in mice signify the advantage of reducing gene editing cargo size despite the modest modification efficiency for *Hpd* target gene after TnpB- $\omega$ RNA optimization. Besides, compact TnpB size could permit using sophisticated regulatory sequences for switchable gene editing and reducing the AAV administration dose for high expression to enable safe therapeutic applications. Furthermore, our optimized  $\omega$ RNA\* variant with less than 100 nt would also be easy for synthesizing chemically modified  $\omega$ RNA, which is very useful for ribonucleoprotein (RNP)-based gene editing applications. To investigate the anti-TnpB immunity in human population, we performed the ELISA and western blot analysis using human blood samples to find that there exists some but not all individuals without carrying antibody against TnpB, similar to SpCas9 (Supplementary Figs. 15 and 16). Therefore, to extend the applicability of TnpB for broad population in the future, it warranted the identification of more TnpB orthologs given the high diversity of TnpB in nature.

Overall, our study demonstrated the enhanced gene editing activity of TnpB via  $\omega$ RNA engineering in cultured cells and showed its disease correction ability in animal models, indicating the potential of hypercompact TnpB- $\omega$ RNA system as effective miniature gene editing modality for more AAV-based disease treatment in animal models and even human patients.

## Methods

### Study approval

The objectives of the present study were to show proof-of-concept for in vivo TnpB-mediated gene editing in wild-type and disease mice. All animal experiments were performed and approved by the Animal Care and Use Committee of Shanghai Center for Brain Science and Brain-Inspired Technology, Shanghai, China. All animal experiments complied with the ARRIVE guidelines for reporting animal experiments. De-identified blood samples were obtained with the patients' written consent in strict observance of the legal and institutional ethical regulations and approved as a non-human study by the Institutional Review board of International Peace Maternity and Child Health Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China. No recruitment criteria was used and demographic information about our blood samples were also de-identified for the patient's privacy.

### Plasmid constructions

The pCBh-TnpB-hU6-Bpil plasmid encoded a human codon-optimized TnpB driven by CBh promoter, and hU6-driven  $\omega$ RNAs with *Bpil* cloning site. The sgRNA and  $\omega$ RNA were designed to be suitable for Un1Cas12f1, TnpB or other Cas proteins, then synthesized as DNA oligonucleotides and cloned into pCBh-Un1Cas12f1, pCBh-TnpB or other Cas plasmids to get the gene editing plasmids. Related plasmids were deposited to Addgene.

### Cell culture, transfection, and flow cytometry analysis

HEK293T were maintained in Dulbecco's modified eagle medium (DMEM) (Gibco, 11965092) supplemented with 10% fetal bovine serum at 37 °C and 5% CO<sub>2</sub> in a humidified incubator. For gene editing analysis, 1  $\mu$ g TnpB or CRISPR plasmids and reporter plasmids were co-transfected using polyethylenimine (PEI) transfection reagent. After transfected cells were cultured for 48 h, we carefully resuspended the cell pellet, and then analyzed or sorted by BD FACSAria II for deep sequencing. Flow cytometry results were analyzed with FlowJo X (v.10.0.7).

### In vitro transcription of TnpB and $\omega$ RNA

TnpB mRNA was transcribed using the mMACHINE T7 Ultra Kit (Invitrogen, AM1345). T7 promoter was added to  $\omega$ RNA template by PCR amplification of pCX2280 using forward and reverse primers in the supplementary files. The PCR products purified with Omega gel extraction Kit (Omega, D2500-02) as templates were transcribed using the MEGAscript Kit (Invitrogen, AM1354). The TnpB mRNA and  $\omega$ RNA were purified by MEGAclear Kit (Invitrogen, AM1908), eluted with RNase-free water and stored at -80 °C.

### Zygote injection and embryo transplantation

Eight-week-old B6D2F1 female mice were superovulated and mated with B6D2F1 male mice, and fertilized embryos were collected from the oviduct. The mixture of TnpB mRNA (50 ng/ $\mu$ L) and  $\omega$ RNA (100 ng/ $\mu$ L) was injected into the cytoplasm of fertilized eggs using a Femtojet microinjector (Eppendorf). The injected embryos were cultured in KOSM medium with amino acids at 37 °C under 5% CO<sub>2</sub> in a humidified incubator overnight and then transferred into oviducts of pseudo-pregnant 8-week-old ICR foster mothers at 0.5-d.p.c. (day post coitus).

### $\omega$ RNA engineering

At first, we performed the secondary structure prediction of  $\omega$ RNA with online RNAfold webserver (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>). Based on the predicted structure, we divided  $\omega$ RNA into six subdomains or segments and tested the influence of each segment on gene editing activity by deleting each segment one by one. In the first-round engineering, we found that deletion

of segment #1, #2, and #3 (corresponding to  $\Delta$ S1,  $\Delta$ S2, and  $\Delta$ S3 in Fig. 3b) does not affect gene editing efficiency of TnpB as shown in Fig. 2c, d. In the second-round engineering, we demonstrated segment #4, #5, and #6 (corresponding to  $\Delta$ SL1,  $\Delta$ SL2 and  $\Delta$ SL3 in Fig. 2e) necessary for the normal gene editing activity of TnpB. In addition, substitution of terminal stem-loop structure with 5'-GAAA-3' resulted in the final  $\omega$ RNA\* variant shown in Fig. 2g with improved gene editing efficiency and small RNA size.

### Targeted deep sequencing

To analyze TnpB gene editing efficiency, the DNA of successfully transfected cells or AAV-TnpB- $\omega$ RNA treated tissues were extracted with TIANamp Genomic DNA Kit (TIANGEN,) according to the manufacturer protocol. DNA was amplified with Phanta max super-fidelity DNA polymerase (Vazyme, P505-d1) for Sanger or deep sequencing methods. Deep-sequencing libraries were generated by adding Illumina flow cell binding sequences and specific barcodes on the 5' and 3' end of the primer sequence. The products were pooled and sequenced with 150 bp paired-end reads on an Illumina HiSeq instrument. FASTQ format data were demultiplexed using the Cutadapt (v.2.8)41 according to the assigned barcode sequences. CRISPResso2 was used for gene editing analysis<sup>29</sup>.

### PEM-seq analysis

Genome-wide off-target analysis was performed following PEM-seq protocol<sup>17</sup>. The 20  $\mu$ g genomic DNA from TnpB edited or control samples were fragmented with Covaris sonicator to generate 300–700 bp DNA. DNA fragments was tagged with biotin at 5'-end by one-round PCR extension using a biotinylated primer, primer leftover removed by AMPure XP beads and purified by streptavidin beads. The single-stranded DNA on streptavidin beads is ligated with a bridge adapter containing 14-bp random molecular barcode, and PCR product was generated via nested PCR to enrich DNA fragment containing the bait DSB events and tagged with illumine adapter sequences. The prepared sequencing library was sequenced by Hi-seq 2500 with 150 bp pair-end reads. PEM-seq data analysis was performed using PEM-Q pipeline with default parameters.

### AAV virus production

The adeno-associated virus 8 (AAV8) serotype was used in this study. The TnpB plasmids with  $\omega$ RNA was sequenced before packaging into AAV8 vehicle, and the AAV vectors were packaged by transfection of HEK293T cell with helper plasmids. The virus titer was  $5 \times 10^{13}$  (AAV-TnpB- $\omega$ RNA) genome copies/mL as determined by qPCR specific for the inverted terminal repeat.

### Gene editing treatment for tyrosinaemia mouse model

Mice were housed in a barrier facility with a 12-h light/dark cycle and 18–23 °C with 40–60% humidity. Diet and water were accessible at all times. The *Fah*<sup>-/-</sup> mouse model harbors the same homozygous G to A point mutation of the last nucleotide of exon 8, which result in exon skipping and the loss of FAH. *Fah*<sup>-/-</sup> mice were kept on 10 mg/L NTBC (Sigma-Aldrich, Cat. No. PHRI731) in drinking water when indicated. For viral particle injection, AAV8 ( $4 \times 10^{11}$  vg/mouse) in 200  $\mu$ l saline were injected via the tail vein into 8–10 weeks old male and female mice. Mice were kept off NTBC water at 7 days post injection, and their body weights were recorded every 3–5 days. Mice were euthanized by CO<sub>2</sub> asphyxiation and harvested at 75 days after NTBC water withdrawal for histology and DNA analysis. Control mice off NTBC water were harvested when reaching >20% weight loss.

### Histological analysis, immunohistochemistry, and immunofluorescence

For histological analysis, liver sections embedded in paraffin were initially deparaffinized in xylene, followed by rehydration with a



gradient of ethanol ranging from 100% to 50%. Subsequently, the sections were washed in distilled water and stained with hematoxylin and eosin (H&E) and picosirius red solution (0.1%) for histological examination.

Immunohistochemistry was conducted on deparaffinized sections, starting with the inhibition of endogenous peroxidase activity using a 0.6% hydrogen peroxide/methanol solution. Antigen retrieval was achieved using EDTA antigen retrieval solution. To minimize nonspecific binding, a 3% BSA solution was applied for 30 min, followed by overnight incubation at 4 °C with the primary antibody (HPD antibody, Santa Cruz, sc-390279, dilution 1:100). On the following day, slides were washed and incubated with the secondary antibody (Goat Mouse IgG, Abcam, ab97023, dilution 1:1000). Staining was visualized using the DAB peroxidase substrate kit.

For mouse tibialis anterior, diaphragm, and heart muscle tissues, immunofluorescence was carried out on frozen sections. Nonspecific binding was blocked with 5% normal goat serum in TBST for 1 h. The slides were then incubated overnight at 4 °C with primary antibodies (dystrophin antibody, Abcam, ab15277, dilution 1:100; laminin-2 antibody, Sigma-Aldrich, L0663, dilution 1:100) in 5% normal goat serum in TBST. Subsequently, slides were washed in TBST and incubated with secondary antibodies (Alexa Fluor 488 Goat Rabbit, Invitrogen, A-11008, dilution 1:1000; Alexa Fluor 594 Goat Rat IgG, Invitrogen, A-11007, dilution 1:1000) for 1 h. Imaging was performed using an FV3000 confocal microscope.

### Serum and urine biochemistry

Mouse blood was collected using retro-orbital puncture before mice were sacrificed. Mouse urine was collected every 24 h for a total of three times. Mouse plasma tyrosine levels were measured on a high-performance liquid chromatograph (HPLC1200) according to the standard protocols. ALT, AST, and bilirubin levels in plasma were determined using diagnostic ELISA Kits (Abcam). Succinylacetone levels were measured using High-Performance Liquid Chromatography (ACQUITY UPLC I-Class) coupled to tandem mass spectrometer (AB Sciex API 4000 LC-MS/MS).

### TnpB protein expression and purification

To express TnpB in *E. coli*, a prokaryotic TnpB expression plasmid (pCX2691) was constructed by cloning TnpB fragment in a pET2b-derived plasmid. TnpB expression plasmid was then transformed into Rosetta (DE3) competent cells for growth in LB medium with ampicillin (100 µg/ml) at 37 °C. After the OD<sub>600</sub> reaches 0.6–0.8, TnpB expression was induced with IPTG. Induced cells were further cultured at 16 °C for 14–16 h. Next, we pelleted bacteria by centrifugation and resuspended in lysis buffer (20 mM Tris-HCl pH 7.0, 250 mM NaCl, 5 mM 2-mercaptoethanol, 25 mM imidazole, 2 mM PMSF and 5% (v/v) glycerol). After sonication, the cell lysate was centrifugated to remove cell debris, and the supernatant was load onto Ni<sup>2+</sup>-charged HiTrap chelating HP column (GE Healthcare). TnpB protein was eluted with imidazole. Pooled TnpB fractions were dialyzed against storage buffer (20 mM Tris-HCl (pH 8.0 at 25 °C), 250 mM NaCl, 2 mM DTT and 50% (v/v) glycerol) and stored at –20 °C for further use.

### Human serum antibody analysis for TnpB and SpCas9 by ELISA

Adult serum was provided by Center for Reproductive Medicine, International Peace Maternity and Child Health Hospital, Innovative Research Team of High-level Local Universities in Shanghai, School of Medicine, Shanghai Jiao Tong University, Shanghai, China. About 2 ml of adult blood samples were collected from veins and left at room temperature for 30 min. Subsequently, they were centrifuged at 1800 × g for 10 min at room temperature, and the resulting supernatants were carefully transferred to tubes. The transferred samples were then subjected to a second centrifugation at 1300 × g for 2 min and stored at –80 °C for future use.

The ELISA protocol was adapted from the standard methodologies<sup>30,31</sup>. In a concise summary, each antigen (TnpB, SpCas9, human albumin) was coated onto a 96-well plate (0.5 µg/well) overnight at 4 °C in a coating buffer. Following this, the plates underwent five washes for 3 min each with TBST wash buffer. Subsequently, the plates were blocked with a 5% bovine serum albumin (BSA) blocking solution for 1 h at room temperature.

Serum samples were diluted 50-fold with 1% BSA Diluent Solution, added to the wells, and incubated for 1 h at 37 °C with shaking (200 rpm). Afterward, the plates were washed three times. Horseradish peroxidase-conjugated goat anti-human IgG Fc secondary antibody (Epigentek, cat#A-9000, dilution 1:5,000) was then applied and incubated for 1 h at room temperature. Following this, 3,3',5,5'-Tetramethylbenzidine substrate solution was added and allowed to develop for 15 min before sulfuric acid was added to stop the reaction. The absorbance at 450 nm was measured using a SpectraMax i3X microplate reader.

### Immunoblot

For the detection of HPD expression in mouse liver, the murine liver tissues were incubated in RIPA Lysis and Extraction Buffer. Equal amounts of proteins were separated on SDS-polyacrylamide gel electrophoresis gels. Primary antibodies were listed as below: Anti-CRISPR-Cas9 (ET1703, HUABIO; dilution 1:1000), Anti-Human-IgG (ALPVHHS, dilution 1:1000), anti-HPD antibody (Santa Cruz, sc-390279; dilution 1:100 or 1:500), anti-P21 antibody (Abcam, ab109199; dilution 1:200), anti-Vinculin (Cell Signaling Technology, 13901 S; dilution 1:100), anti-dystrophin antibody (Abcam, ab15277; dilution 1:100), anti-laminin-2 antibody (Sigma-Aldrich, L0663; dilution 1:100). For anti-TnpB/SpCas9 antibody analysis, aliquots of 1 µg TnpB and 1 µg SpCas9 were resolved in buffer respectively and then applied to a 10% SDS-polyacrylamide gel electrophoresis. Samples were transferred to a PVDF membrane and blocked with 5% BSA in TBST for 1 h at room temperature. Immunoblots were then incubated overnight in TBST with 0.05% BSA with a 1:10 dilution of serum. Immunoblots were washed three times for 8 min in TBST on a shaker and then incubated with horseradish peroxidase-conjugated goat anti-human IgG Fc secondary antibody for 1 h at room temperature. After the incubation with Chemiluminescent substrates (#WP20005, Invitrogen), the membranes were viewed by the Image Lab™ Software 5.2.

### Statistical analysis

The number of independent biological replicates is shown in the figure legend. The data are presented as means ± SEM. Differences were assessed using unpaired two-tailed Student's *t* tests or one-way ANOVA. Differences in means were considered statistically significant at  $P < 0.05$  (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ ).

### Reporting summary

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

### Data availability

Deep-seq data are deposited to the GEO repository under accession number [PRJNA973546](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=PRJNA973546). Source data are provided with this paper.

### References

1. Karvelis, T. et al. Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. *Nature* **599**, 692–696 (2021).
2. Altae-Tran, H. et al. The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* **374**, 57–65 (2021).
3. Wang, D., Zhang, F. & Gao, G. CRISPR-based therapeutic genome editing: strategies and in vivo delivery by AAV vectors. *Cell* **181**, 136–150 (2020).

4. Harrington, L. B. et al. Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* **362**, 839–842 (2018).
5. Xu, C. et al. Programmable RNA editing with compact CRISPR-Cas13 systems from uncultivated microbes. *Nat. Methods* **18**, 499–506 (2021).
6. Kannan, S. et al. Compact RNA editors with small Cas13 proteins. *Nat. Biotechnol.* **40**, 194–197 (2022).
7. Kim, D. Y. et al. Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus. *Nat. Biotechnol.* **40**, 94–102 (2022).
8. Kong, X. et al. Engineered CRISPR-OsCas12f1 and RhCas12f1 with robust activities and expanded target range for genome editing. *Nat. Commun.* **14**, 2046 (2023).
9. Bigelyte, G. et al. Miniature type V-F CRISPR-Cas nucleases enable targeted DNA modification in cells. *Nat. Commun.* **12**, 6191 (2021).
10. Wu, Z. et al. Programmed genome editing by a miniature CRISPR-Cas12f nuclease. *Nat. Chem. Biol.* **17**, 1132–1138 (2021).
11. Xu, X. et al. Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol. Cell* **81**, 4333–4345 e4334 (2021).
12. Tong, H. et al. High-fidelity Cas13 variants for targeted RNA degradation with minimal collateral effects. *Nat. Biotechnol.* **41**, 108–119 (2023).
13. Li, Y. et al. The collateral activity of RfxCas13d can induce lethality in a RfxCas13d knock-in mouse model. *Genome Biol.* **24**, 20 (2023).
14. Nowak, C. M., Lawson, S., Zerez, M. & Bleris, L. Guide RNA engineering for versatile Cas9 functionality. *Nucleic Acids Res.* **44**, 9555–9564 (2016).
15. Yang, Y. et al. Highly efficient and rapid detection of the cleavage activity of Cas9/gRNA via a fluorescent reporter. *Appl. Biochem. Biotechnol.* **180**, 655–667 (2016).
16. Bae, S., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).
17. Yin, J. et al. Optimizing genome editing strategy by primer-extension-mediated sequencing. *Cell Discov.* **5**, 18 (2019).
18. Makarova, K. S. et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
19. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
20. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
21. Abudayeh, O. O. et al. RNA targeting with CRISPR-Cas13. *Nature* **550**, 280–284 (2017).
22. Wang, J. Y. & Doudna, J. A. CRISPR technology: a decade of genome editing is only the beginning. *Science* **379**, eadd8643 (2023).
23. Kim, D. Y. et al. Hypercompact adenine base editors based on transposase B guided by engineered RNA. *Nat. Chem. Biol.* **18**, 1005–1013 (2022).
24. Karvelis, T. & Siksnyš, V. Mis-annotation of TnpB: case of TaRGET-ABE. *Nat. Chem. Biol.* **19**, 261–262 (2023).
25. Yoon, P. H., Adler, B. A. & Doudna, J. A. To TnpB or not TnpB? Cas12 is the answer. *Nat. Chem. Biol.* **19**, 263–264 (2023).
26. Nakagawa, R. et al. Cryo-EM structure of the transposon-associated TnpB enzyme. *Nature* **616**, 390–397 (2023).
27. Sasnauskas, G. et al. TnpB structure reveals minimal functional core of Cas12 nuclease family. *Nature* **616**, 384–389 (2023).
28. Raguram, A., Banskota, S. & Liu, D. R. Therapeutic in vivo delivery of gene editing agents. *Cell* **185**, 2806–2827 (2022).
29. Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
30. Charlesworth, C. T. et al. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019).
31. Wang, D. et al. Adenovirus-mediated somatic genome editing of Pten by CRISPR/Cas9 in mouse liver in spite of Cas9-specific immune responses. *Hum. Gene Ther.* **26**, 432–442 (2015).

## Acknowledgements

We thank technical support from the laboratory animal center (Y.D., J.S., and T.Z.), optical imaging (L.T., K.S., and W.L.), and gene-editing core (R.Y., X.H., and X.Z.) facility in Shanghai Center for Brain Science and Brain-Inspired Technology as well as Lingang Laboratory. This work was funded by Lingang Laboratory (startup fund to C.X.), and Shanghai City Committee of Science and Technology Project (22QA1412300 to C.X., 20ZR1466600 to X.H.). S.M. was funded by National Natural Science Foundation of China (32100641). H.Y. was funded by National Science and Technology Innovation 2030 Major Program (2021ZD0200900) (H.Y.), Chinese National Science and Technology major project R&D Program of China (2018YFC2000101) (H.Y.), Strategic Priority Research Program of Chinese Academy of Science (XDB32060000) (H.Y.), National Natural Science Foundation of China (31871502, 31901047, 31925016, 91957122, and 82021001) (H.Y.), Basic Frontier Scientific Research Program of Chinese Academy of Sciences From 0 to 1 original innovation project (ZDBS-LY-SM001) (H.Y.), Shanghai Municipal Science and Technology Major Project (2018SHZDZX05) (H.Y.), Shanghai City Committee of Science and Technology Project (18411953700, 18JC1410100, 19XD1424400 and 19YF1455100) (H.Y.) and the International Partnership Program of Chinese Academy of Sciences (153D31KYSB20170059) (H.Y.).

## Author contributions

C.X., Z.L., and R.G. jointly conceived the project and designed experiments. Y.Z. and C.X. supervised the whole project. Z.L. and G.L. generated mouse model. Z.L. and R.G. designed vectors, performing in vitro experiments and scanning confocal imaging. X.H., X.S., X.C., and Z.S. assisted with the construction of plasmids and cell culture. R.Y. and X.Z. prepared AAV virus. R.G., Z.L., Z.S., Y.Y., X.L., and G.L. performed in vivo virus injection, tissue dissection, histological immunostaining, and liver function experiments. H.Z., X.C., J.S., S.M., and W.Z. performed PEM-seq experiments. M.Z. collected adult serum. Y.L. and Y.Z. performed bioinformatics analysis. R.G., G.L., and X.H. assisted with tissue dissection, immunostaining, and animal breeding. Z.L., R.G., G.L., C.H., Y.Z., and C.X. analyzed the data and organized figures. H.Y., C.H., Y.Z., and C.X. wrote the manuscript with data contributed by all authors who participated in the project.

## Competing interests

H.Y. is a founder of HuidaGene Therapeutics. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-45197-z>.

**Correspondence** and requests for materials should be addressed to Chunyi Hu, Yingsi Zhou or Chunlong Xu.

**Peer review information** *Nature Communications* thanks Nicola Brunetti-Pierri, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024