Article

A male-essential miRNA is key for avian sex chromosome dosage compensation

https://doi.org/10.1038/s41586-025-09256-9

Received: 25 February 2024

Accepted: 9 June 2025

Published online: 16 July 2025

Open access

Check for updates

Amir Fallahshahroudi[1,2,9]✉, Sara Yousefi Taemeh[2,9], Leticia Rodríguez-Montes[1], Nils Trost[1], Dana Frank[1], Pascal Lafrenz[1], Jiri Koubek[1], Guillermo Tellez Jr.[3], Maeve Ballantyne[3], Alewo Idoko-Akoh[4], Lorna Taylor[3], Adrian Sherman[3], Megan Davey[3], Cheng Ma[2], Enrico Sorato[5], Martin Johnsson[6], Christina Grozou[2], Ying Xue[7], Long Liu[7], Guenter Kramer[1], Carl-Johan Rubin[2], Margarida Cardoso-Moreira[8], Mike J. McGrew[3,10]✉ & Henrik Kaessmann[1,10]✉

Birds have a sex chromosome system in which females are heterogametic (ZW) and males are homogametic (ZZ)[1]. The differentiation of avian sex chromosomes from ancestral autosomes entails the loss of most genes from the W chromosome during evolution[1,2]. However, the extent to which mechanisms evolved that counterbalance this substantial reduction in female gene dosage remains unclear. Here we report functional in vivo and evolutionary analyses of a Z-linked microRNA (miR-2954) with strong male-biased expression, previously proposed to mediate avian sex chromosome dosage compensation[3]. We knocked out miR-2954 in chicken, which resulted in early embryonic lethality in homozygous knockout males, probably driven by specific upregulation of dosage-sensitive Z-linked target genes. Evolutionary gene expression analyses further revealed that these dosage-sensitive target genes underwent both transcriptional and translational upregulation on the single Z in female birds. Altogether, this work unveils a scenario in which evolutionary pressures following W gene loss drove transcriptional and translational upregulation of dosage-sensitive Z-linked genes in females but also their transcriptional upregulation in males. The resulting excess of transcripts in males, resulting from the combined activity of two upregulated dosage-sensitive Z gene copies, was in turn offset by the emergence of a highly targeted miR-2954-mediated transcript degradation mechanism during avian evolution. This study uncovered a unique sex chromosome dosage compensation system in birds, in which a microRNA has become essential for male survival.

The emergence of sex chromosomes from ancestral autosomes during amniote evolution (Fig. 1a) involved extensive gene loss on the sex-specific chromosomes, which are the Y in male-heterogametic XY systems and the W in female-heterogametic ZW systems[1,2,4]. To counter gene dosage reductions in the heterogametic sex, compensatory mechanisms evolved[1,4,5]. In placental and marsupial (therian) mammals (Fig. 1a), male dosage reductions are primarily offset by an approximately twofold increase in the expression level of many genes on the single X chromosome through upregulation at both transcriptional and translational layers[4–6], restoring ancestral expression levels in males, along with extra mechanisms[4,5]. The resulting overexpression of X-linked genes in females, owing to two upregulated X chromosomes (therian upregulation mechanisms are not male-specific), is secondarily compensated by X inactivation, mediated by *XIST* and *RSX* long non-coding RNAs (lncRNAs) in placentals and marsupials, respectively[7,8]. By contrast, the XY system of green anole lizards features a male-specific twofold transcriptional upregulation, akin to that of fruitflies[9], making female X inactivation unnecessary. Across these systems, the combined effects of different mechanisms result in similar expression outputs between the sexes for most X-linked genes[1,4].

In the avian ZW system (Fig. 1a), the Z-linked gene expression output is substantially higher in ZZ males than in ZW females[5,10–12], and both the extent and mechanisms of dosage compensation remain poorly understood. Previous studies have shown incomplete transcriptional upregulation of Z-linked genes in females, probably because of partial upregulation across many genes or full upregulation of only a subset, resulting in transcript levels below those of ancestral autosomal genes[5,9]. By contrast, males retain ancestral expression levels from their two Z chromosomes[5,9], leaving it unclear whether they have been affected by any upregulation mechanism. However, the relatively balanced expression of dosage-sensitive genes between the sexes suggests that avian dosage compensation specifically targets these genes[13]. We recently identified a Z-linked microRNA (miRNA; miR-2954) that is predominantly expressed in males (approximately 5-fold to 10-fold higher

[1]Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, Heidelberg, Germany. [2]Department of Medical Biochemistry and Microbiology, Biomedical Center (BMC), Uppsala University, Uppsala, Sweden. [3]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, UK. [4]School of Biochemistry, Faculty of Life Sciences, University of Bristol, Bristol, UK. [5]Reneco International Wildlife Consultants, Abu Dhabi, United Arab Emirates. [6]Department of Animal Biosciences, Swedish University of Agricultural Sciences, Uppsala, Sweden. [7]College of Animal Science and Technology, Yangzhou University, Yangzhou, China. [8]Evolutionary Developmental Biology Laboratory, Francis Crick Institute, London, UK. [9]These authors contributed equally: Amir Fallahshahroudi, Sara Yousefi Taemeh. [10]These authors jointly supervised this work: Mike J. McGrew, Henrik Kaessmann. ✉e-mail: amirshahroudi@gmail.com; mike.mcgrew@roslin.ed.ac.uk; h.kaessmann@zmbh.uni-heidelberg.de

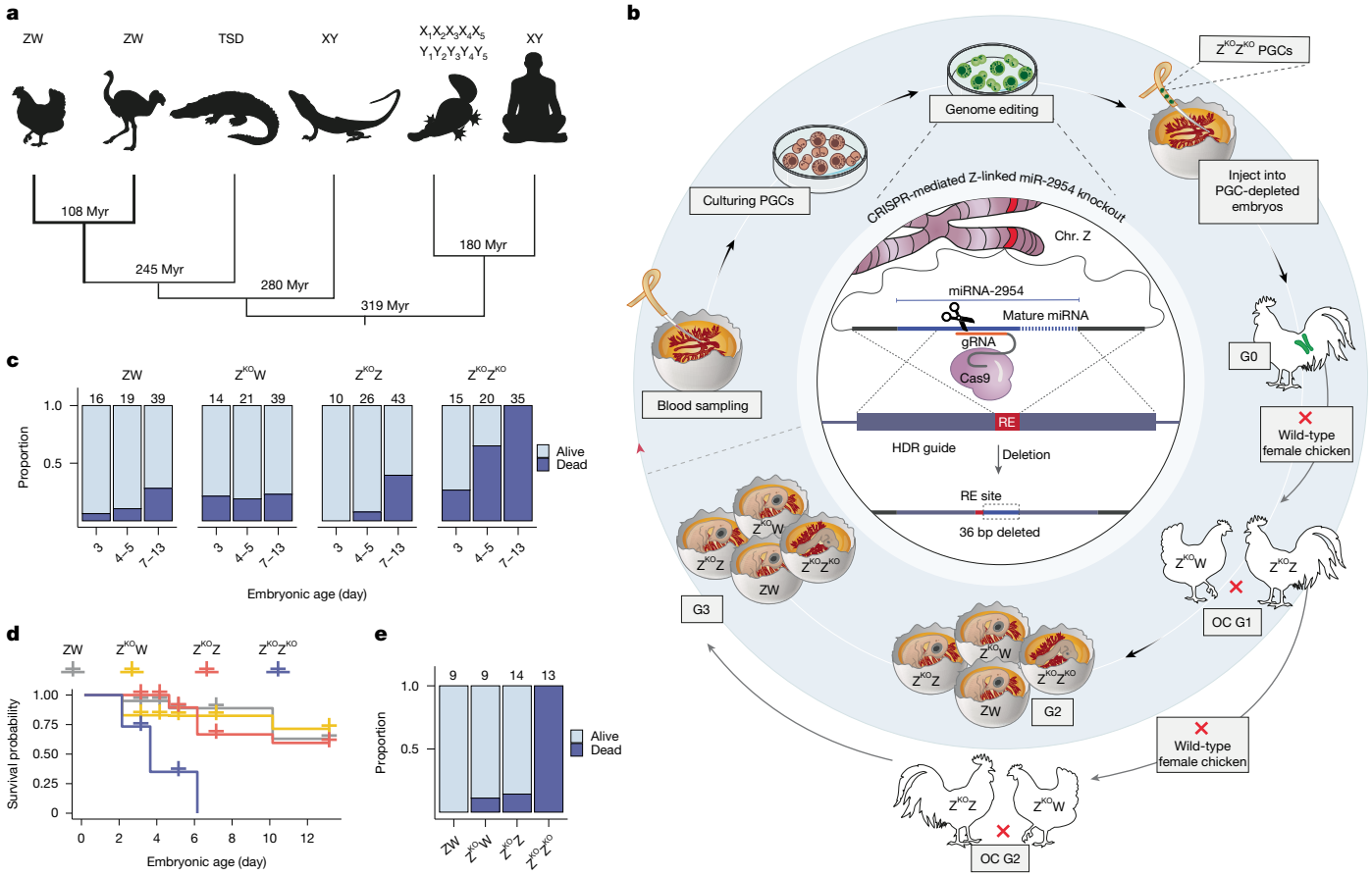


Fig. 1 | Role of miR-2954 in male chicken development. a, Overview of major known sex determination systems in amniotes[2,3,9]: ZW in birds (icons indicate chicken and ostrich, marking the deepest divergence in bird phylogeny), temperature-dependent sex determination (TSD) in crocodiles, XY in Iguania lizards (icon reflects the green anole), multiple (5X and 5Y) sex chromosomes in platypus and XY in humans. The approximate divergence times in million years (Myr) are indicated at the respective nodes. Note that the XY sex chromosomes in lizards and humans have evolved independently from different ancestral autosomes. **b**, Schematic of the experimental design used for the generation of miR-2954 KO chickens across generations (G0–G3) on the basis of genome editing (CRISPR–Cas9 with single guide RNA (sgRNA)) in PGCs and outcrossings (OCs), and the assessment of the resulting phenotypes (see Methods and Extended Data Fig. 1 for details). The restriction enzyme (RE) site used for genotype screening and homology-directed repair (HDR) template are indicated. **c**, Distribution of live and dead second-generation (G2) embryos, categorized by genotype (female wild-type ZW, female hemizygous KO $Z^{KO}W$, male heterozygous KO $Z^{KO}Z$ and male homozygous KO $Z^{KO}Z^{KO}$) and embryonic day of development. Numbers above the bars indicate the total number of embryos analysed for each subgroup. **d**, Kaplan–Meier survival curves showing survival rates for embryos with different genotypes during development. **e**, Distribution of live and dead third-generation (G3) embryos at embryonic day 14, grouped by genotype.

than that in females across tissues and embryos[3,14]) and with predicted targets that are mainly Z-linked dosage-sensitive genes that show more balanced expression between sexes than other Z-linked genes[3], suggesting a key role for miR-2954 in avian dosage compensation.

Here we assessed the function of miR-2954 in vivo by using a chicken knockout (KO) model. Combined with evolutionary genomics analyses, our study clarified the mechanisms, extent and importance of avian sex chromosome dosage compensation.

Generation of miR-2954 chicken KO

To investigate the function of miR-2954, located in an intron of the host gene *XPA* (Extended Data Fig. 1), we generated chicken KO lines by genome editing primordial germ cells[15–18] (PGCs) (Fig. 1b, Extended Data Fig. 1 and Supplementary Table 1). Using a high-fidelity CRISPR–Cas9 system[19], to minimize potential off-target mutations and homology-directed repair[20], we deleted both copies of the miR-2954 locus ($Z^{KO}Z^{KO}$) in male-embryo-derived PGCs[17,21]. These edited PGCs were injected into surrogate embryos lacking endogenous germ cells[16,21], producing a gonadal chimaeric rooster (generation 0 (G0)). The G0 male matured and successfully mated with six wild-type hens, generating outcross generation 1 (OC G1) offspring: hemizygous females ($Z^{KO}W$) and heterozygous males ($Z^{KO}Z$). OC G1 birds showed no deleterious phenotypes, reached sexual maturity and produced viable offspring (Fig. 1b). Mating OC G1 males and females produced second-generation (G2) embryos (ZW, $Z^{KO}W$, $Z^{KO}Z$ and $Z^{KO}Z^{KO}$). Additionally, mating an OC G1 male with wild-type females yielded a second outcross generation (OC G2), which was interbred to generate third-generation (G3) embryos, enabling the confirmation of G2 phenotypes (Fig. 1b).

Our KO strategy allowed phenotypic comparisons among outbred sibling birds of different genotypes, thus minimizing genetic background variability. Repeated outcrossing effectively ruled out autosomal off-target edits because any such mutations would need to be closely linked to the Z-chromosomal miR-2954 deletion. To exclude this possibility, we performed deep long-read genome resequencing of five homozygous KO males and four wild-type controls using adaptive sampling[22], selectively enriching Z-chromosomal sequences to approximately 10-fold to 32-fold coverage (Extended Data Fig. 2a,b). These analyses revealed only the intended 36-bp miR-2954 deletion in KO birds, with no other structural variations, confirming the absence of off-target edits (Extended Data Fig. 2c).

Requirement of miR-2954 for male survival

To assess the phenotypic consequences of miR-2954 KO, we examined the viability of 297 G2 embryos at embryonic (E) days E3, E4–5 and E7–13, evaluating morphology and heartbeat under a stereomicroscope (Supplementary Table 2 and Supplementary Data 1 and 2). Wild-type female (ZW), hemizygous female ($Z^{KO}W$) and heterozygous male ($Z^{KO}Z$) embryos had statistically indistinguishable survival rates of approximately 79–85% ($P > 0.15$; two-sided χ^2 test) (Fig. 1c,d). These observations demonstrate that miR-2954 is dispensable for female development, which is consistent with its very low expression in this sex[3], and haplosufficient in males. By contrast, all $Z^{KO}Z^{KO}$ males died before E7, demonstrating significantly higher lethality ($P < 10^{-9}$; two-sided χ^2 test) (Fig. 1c,d). Analysis of 45 more embryos from generation 3 (G3) at E14 confirmed these results, again showing 100% lethality in $Z^{KO}Z^{KO}$ males, whereas all other genotypes exhibited high survival rates (approximately 86–100% for all other genotypes; Fig. 1e). Notably, the restriction of lethality to homozygous KOs further confirms the absence of off-target edits on autosomes or distant Z-chromosomal regions because these would also affect heterozygous or hemizygous embryos.

Altogether, this study revealed that miR-2954 is male-essential, making it the only known sex-specific essential miRNA identified across taxa so far.

Derepression of Z-linked genes in KOs

Although miRNAs can repress translation, they predominantly regulate target mRNA abundance by guiding Argonaute proteins to complementary sites in the 3′ untranslated region (UTR) through short (6–8 nt) seed sequences, causing mRNA degradation[23–25]. To investigate the molecular basis of the male-lethal phenotype in miR-2954 homozygous KO embryos, we systematically compared genome-wide transcript abundance between KO and wild-type embryos. For this, we generated RNA sequencing (RNA-seq) data for male and female embryos, as well as for the head, heart and the rest of the body of E3 and E5 males across all KO and wild-type genotypes (Extended Data Fig. 3a).

On the basis of these RNA-seq data, we first assessed gene expression changes between male homozygous ($Z^{KO}Z^{KO}$) KO and wild-type (ZZ) embryos. To link potential expression changes to miR-2954 loss, we predicted its potential targets using TargetScan[26], screening for complementary 6–8 nt seed sequences in the 3′ UTR (Supplementary Table 3). Predicted Z-linked and autosomal miR-2954 targets showed significantly greater increases in transcript abundance in KOs than in non-target genes (Fig. 2a (left and middle) and Supplementary Table 4), consistent with miRNA-mediated repression and miR-2954 removal in KO embryos. However, the predicted Z-linked targets displayed markedly higher upregulation (median $\log_2$[fold change] ($\log_2$[FC]): whole embryo, 0.41; head, 0.45; heart, 0.48; body, 0.48) than autosomal targets (median $\log_2$[FC]: whole embryo, 0.01; head, 0; heart, −0.01; body, 0.01), which were minimally affected (Fig. 2a (left) and Extended Data Fig. 3c).

Using high-quality ribosome profiling (or ribosome sequencing (Ribo-seq)) data (Extended Data Fig. 4), which measures protein synthesis rates at high resolution[27], and matched RNA-seq data, we confirmed that the pronounced upregulation of Z-linked targets in KO embryos extends to the translational level. Specifically, Z-target genes displayed similarly elevated expression levels in this expression layer compared to other genes (Extended Data Fig. 5a), indicating that translational buffering does not mitigate miR-2954 loss in KO males, and Z-linked protein abundance is indeed elevated. Moreover, the comparable extent of upregulation at both the transcriptome and translatome layers suggests that miR-2954 primarily reduces target mRNA stability rather than repress translation, consistent with the predominant mode of action of miRNAs in most contexts[23–25].

To identify the genes that drive these patterns, we used DESeq2 (ref. 28) to compare the transcript levels between $Z^{KO}Z^{KO}$ and wild-type (ZZ) embryos. Approximately 50% of the 375 predicted Z-linked target genes were significantly differentially expressed across tissues (Fig. 2a (right)), whereas only approximately 3–5% of the 3,383 predicted autosomal targets showed expression differences, a proportion only slightly higher than that for autosomal non-target genes (approximately 2–5%) (Fig. 2a). Among the differentially expressed genes, the vast majority of Z-linked targets (approximately 99–100%) were upregulated compared with approximately 38–54% of differentially expressed autosomal targets ($P < 10^{-15}$; two-sided χ^2 test). Supporting the role of miR-2954, which is broadly expressed throughout development[3,14], in repressing Z-linked transcript abundance, we found that approximately 54% of Z-linked targets showed consistent upregulation across tissues in KO embryos, in stark contrast to predicted autosomal targets (Fig. 2b).

We also observed that 26–35 predicted non-target Z-linked genes were differentially expressed between $Z^{KO}Z^{KO}$ and wild-type ZZ embryos, with approximately 71–92% of these being upregulated, levels comparable to those of predicted Z-linked targets (approximately 99–100%), and significantly higher than those of autosomal non-targets (approximately 38–54%) (all P values $< 10^{-6}$; two-sided χ^2 test) (Fig. 2a). These findings probably reflect false negatives in the Z-target predictions, potentially owing to non-canonical binding sites (for example, within gene bodies), incomplete UTR annotations in the chicken genome or indirect effects stemming, for example, from the repression of transcription factors that regulate other Z-linked genes. Thus, miR-2954 may directly or indirectly regulate more Z-linked genes than currently predicted. Notably, this could include *XPA*, the host gene of miR-2954, which is not a predicted target but is significantly upregulated in KOs ($\log_2$[FC] of approximately 0.6–1.2) relative to controls (Supplementary Table 4).

We next examined gene expression changes in male heterozygous ($Z^{KO}Z$) KO embryos. Similar to the pattern observed in homozygous KOs ($Z^{KO}Z^{KO}$), $Z^{KO}Z$ embryos showed predominant upregulation of the predicted Z-linked target genes (Fig. 2c and Extended Data Fig. 6). However, the magnitude of upregulation was significantly lower, consistent with the presence of one intact copy of the miR-2954 locus, enabling partial repression of its targets and explaining the absence of a lethal phenotype in these embryos. Small RNA-seq data analysis of E5 males confirmed that miR-2954 expression in $Z^{KO}Z$ embryos was reduced by 15–48% compared to the wild type and, as expected, completely absent in $Z^{KO}Z^{KO}$ embryos (Fig. 3a). Global miRNA profiling revealed high miR-2954 expression in males across tissues and developmental stages[3,14], with no changes in other miRNAs upon miR-2954 KO (Fig. 3b,c and Supplementary Table 5), reinforcing the specific role of miR-2954 in the observed regulatory and phenotypic effects.

In female hemizygous ($Z^{KO}W$) embryos, we also observed upregulation of predicted Z-linked target genes compared to wild-type controls, although the effect was substantially weaker than that in male heterozygotes ($Z^{KO}Z$) (Fig. 2c). This pattern is consistent with the 7-fold to 15-fold lower expression of miR-2954 in females[3,14] (Fig. 3c; see above) and the lack of any observable deleterious phenotype in $Z^{KO}W$ embryos.

To understand why miR-2954 preferentially targets Z-linked genes, as reflected in the disproportionately high upregulation of these targets in KO embryos, we examined features known to influence miRNA-mediated repression. Repression efficacy is primarily determined by seed–UTR complementarity, with 8-mer binding sites exerting the strongest effects and several target sites within a 3′ UTR amplifying repression[23,26]. Other factors, including seed-pairing stability, local sequence context and site position, were integrated into TargetScan context+ score[26]. Consistent with these principles, Z-linked targets were significantly enriched for 8-mer sites and several seed matches (6–8 nt) relative to autosomal targets (Fig. 4a), and they also exhibited significantly lower context+ scores, indicating a greater susceptibility to repression by miR-2954 (Fig. 4b).

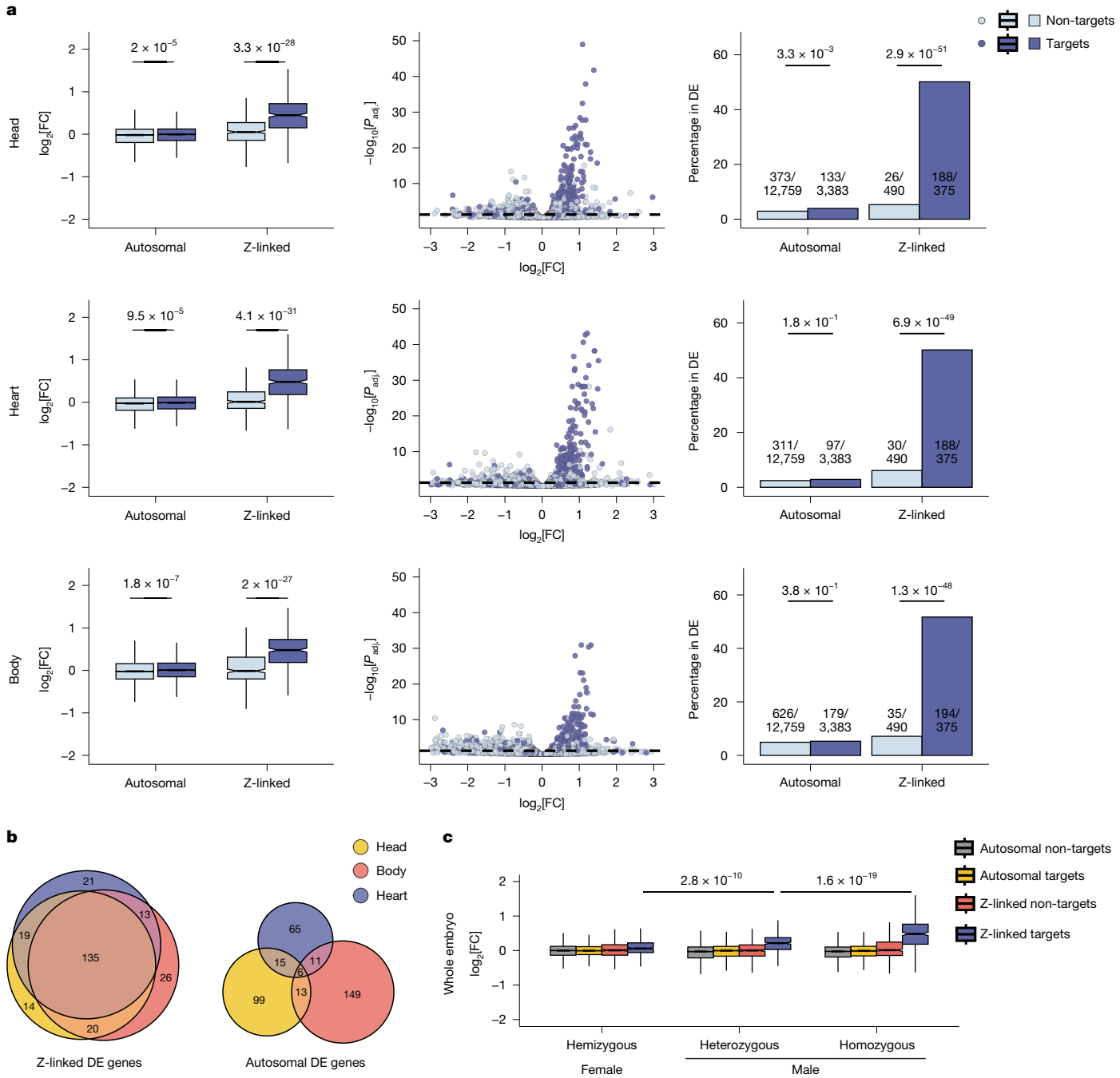


Fig. 2 | Impact of miR-2954 KO on gene expression in different tissues. **a**, Left, $\log_2$[FC] in gene expression between $Z^{KO}Z^{KO}$ and ZZ genotypes for autosomal (n = 16,142) and Z-linked (n = 865) protein-coding genes; P values from two-sided Wilcoxon rank-sum tests are shown above. The $\log_2$[FC] estimates were on the basis of transcriptomes of E3 and E5 embryos from head, heart and body tissues (n = 3 biological replicates per tissue, genotype and developmental stage) using a statistical model that accounts for embryonic age. Middle, volcano plots showing the $\log_2$[FC] and $-\log_{10}$ of Benjamini–Hochberg $P_{adj.}$ values for predicted miR-2954 target and non-target protein-coding genes compared with the $Z^{KO}Z^{KO}$ and ZZ genotypes. Right, proportions of autosomal and Z-linked target and non-target genes among differentially expressed (DE) genes (Benjamini–Hochberg $P_{adj.} < 0.05$) compared with the $Z^{KO}Z^{KO}$ and ZZ genotypes. The distribution of the predicted targets and non-targets of miR-2954 was compared using a χ^2 test for each group. **b**, Overlap of Z-linked and autosomal differentially expressed genes across tissues. **c**, The $\log_2$[FC] in gene expression of autosomal and Z-linked target and non-target protein-coding genes in female hemizygous ($Z^{KO}W$), male heterozygous ($Z^{KO}Z$) and male homozygous ($Z^{KO}Z^{KO}$) genotypes compared with the corresponding wild-type controls. P values from a two-sided Wilcoxon signed-rank test assessed differences in $\log_2$[FC] of Z-linked targets (n = 375) between $Z^{KO}W$ and $Z^{KO}Z$ and between $Z^{KO}Z$ and $Z^{KO}Z^{KO}$ genotypes. The $\log_2$[FC] estimates were on the basis of the transcriptome of the whole embryos at E2. All box plots show the median, 25th–75th percentiles and whiskers extending to 1.5× the interquartile range (IQR).

To validate the male-specific essentiality of miR-2954 and confirm its mechanism of action, we performed an orthogonal in vivo knockdown using an miR-2954-specific inhibitor (that is, complementary chemically modified oligonucleotides that sequester miR-2954) and a delivery reagent (Methods). Consecutive injections at E2.5 and E4 led to significantly increased mortality in male knockdown embryos by E12, compared to both female knockdown embryos (P = 0.0005; two-sided χ^2 test) and negative controls (P = 0.0006; two-sided χ^2 test) (Extended Data

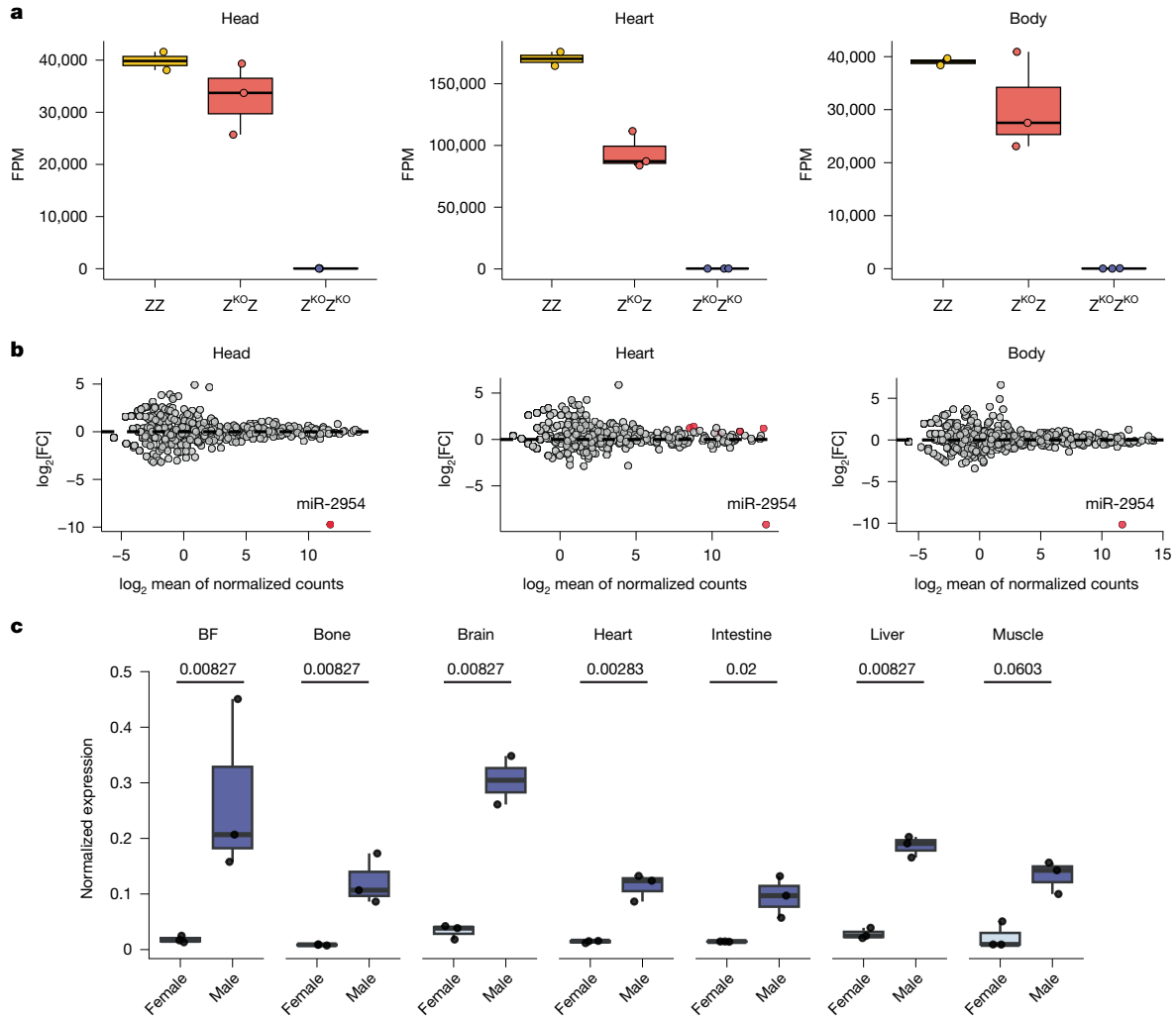


Fig. 3 | Expression patterns of miR-2954 and other miRNAs across tissues in control and KO embryos. a, Distribution of expression levels (fragments per million mapped reads (FPM)) of miR-2954 in ZZ ($n = 2$), $Z^{KO}Z$ ($n = 3$) and $Z^{KO}Z^{KO}$ ($n = 3$) genotypes across head, heart and body at E5. Individual data points are overlaid with jitter. **b**, MA plot showing mean expression and $\log_2$[FC] of mature miRNAs ($n = 674$) when comparing $Z^{KO}Z^{KO}$ and ZZ genotypes across tissues. miRNAs with significant expression changes (Benjamini–Hochberg $P_{adj.}$ value < 0.01) are shown in red. For miR-2954, Benjamini–Hochberg $P_{adj.}$ values were less than 1×10^{-38} for all tissues. **c**, Normalized expression ($2^{-\Delta CT}$) of miR-2954 across the bursa of Fabricius (BF), bone, brain, heart, intestine, liver and muscle in male and female chicken embryos at E12 on the basis of reverse transcription–quantitative polymerase chain reaction (RT–qPCR) ($n = 3$). Benjamini–Hochberg $P_{adj.}$ values from two-sided *t*-tests are shown above. Individual data points were overlaid with jitter. All box plots show the median, 25th–75th percentiles and whiskers extending to 1.5× the IQR.

Fig. 7a,b). The qPCR analysis of E5 male hearts showed significant upregulation of five out of eight target genes (Benjamini–Hochberg-adjusted P ($P_{adj.}$) < 0.1) and host gene *XPA*, whereas non-target genes remained unaffected (Extended Data Fig. 7c). On average, target genes showed a substantially higher fold change (mean $\log_2$[FC] = 1.06) than non-target genes (mean $\log_2$[FC] = 0.17; $P = 0.011$; paired *t*-test). These results further confirmed that the male-lethal phenotype stems from miR-2954 loss rather than off-target effects. Moreover, the significant upregulation of *XPA* in knockdown animals ($P_{adj.} = 0.03$), as also seen in KO embryos, indicates that it is indirectly regulated by miR-2954, probably through a transcriptional activator that is itself a direct miR-2954 target.

Overall, these results demonstrate that miR-2954 specifically targets Z-linked genes and that, therefore, its removal leads to their upregulation, with lethal consequences in homozygous male KO embryos.

Targeting of dosage-sensitive Z genes

To understand why the upregulation of Z-linked miR-2954 targets is lethal in $Z^{KO}Z^{KO}$ embryos, we analysed these genes in more detail. We focused on 'experimentally validated' targets that were predicted to be Z-linked genes and were significantly differentially expressed in at least one tissue of homozygous male KOs, suggesting that they represent actual targets of miR-2954 (Supplementary Table 3). This set comprised 249 genes (248 upregulated), representing approximately 66% of all 375 predicted Z-linked targets and approximately 29% of the 865 Z-linked protein-coding genes.

We next investigated whether the experimentally validated Z-linked targets were dosage-sensitive, as suggested by the deleterious effects of their upregulation in homozygous male KOs. Ohnologues, which are gene duplicates retained from two rounds of whole-genome duplication in vertebrate ancestors[29,30], are often dosage-sensitive because of their enrichment in developmental pathways and protein complexes[31]. We found that 49% (98 out of 200) of all Z-linked ohnologues were among the validated targets, representing a significant enrichment compared to other Z-linked genes (Fig. 4c).

Given that not all dosage-sensitive genes are ohnologues, and not all ohnologues are necessarily dosage-sensitive, we assessed dosage sensitivity across the entire Z chromosome using a resource that estimates

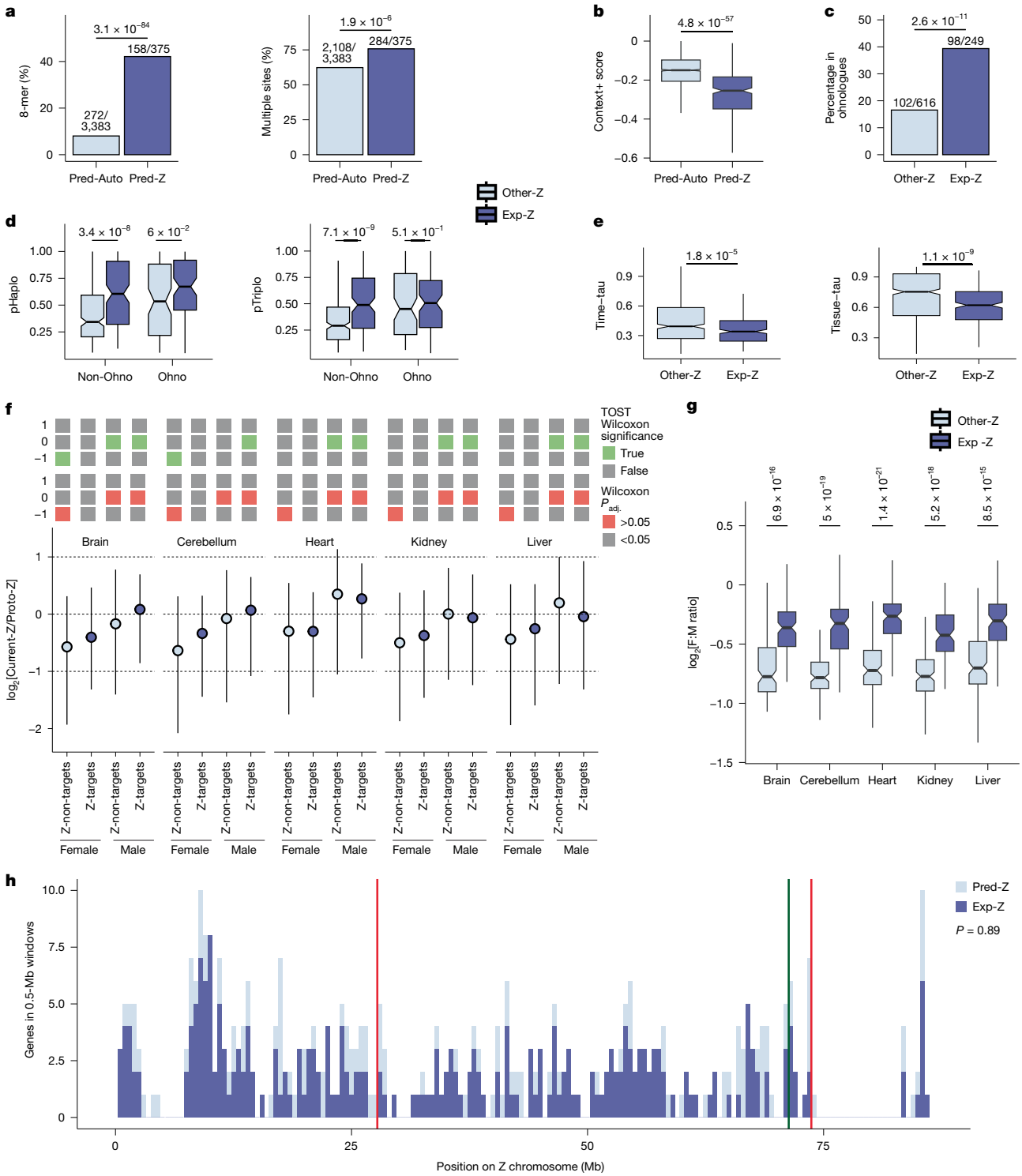


Fig. 4 | Experimental miR-2954 targets. a, Proportions of predicted autosomal (Pred-Auto) and Z-linked (Pred-Z) miR-2954 targets with 8-mer binding sites (left) or several sites (right); *P* values from two-sided χ^2 tests. **b**, The context+ scores of Pred-Auto ($n = 1{,}892$) and Pred-Z ($n = 321$); *P* value from two-sided Wilcoxon rank-sum test. **c**, Proportions of experimentally validated Z-linked targets (Exp-Z) and other Z-linked protein-coding genes (Other-Z) among chicken ohnologues; *P* value from two-sided χ^2 test. **d**, Probabilities of haploinsufficiency (pHaplo) and triplosensitivity (pTriplo) between Exp-Z ($n = 210$) and Other-Z ($n = 320$) genes, split by ohnologue status; *P* values from two-sided Wilcoxon rank-sum tests. **e**, Tissue and developmental tau scores (0, broad; 1, specific) between Exp-Z ($n = 248$) and Other-Z ($n = 461$) genes; *P* values from two-sided Wilcoxon rank-sum tests. **f**, Median and IQRs of the ratios of current versus proto-Z (ancestral) expression for Z-linked non-targets ($n = 193$) and Exp-Z genes ($n = 201$) for males and females on the $\log_2$ scale. Reference lines indicate the ratios of −1 (half-ancestral expression), 0 (equal expression) and 1 (twofold ancestral expression). Statistical significance is shown as two one-sided test (TOST) Wilcoxon equivalence test (green, within ±0.5 of reference; grey, not significant) and two-sided one-sample Wilcoxon test (red, significant deviation; grey, not significant). **g**, The $\log_2$ female-to-male expression ratios for Exp-Z and Other-Z with fragments per kilobase million (FPKM) greater than one in the brain ($n = 236$), cerebellum ($n = 220$), heart ($n = 225$), kidney ($n = 229$) and liver ($n = 237$); *P* values from two-sided Wilcoxon rank-sum tests. **h**, Distribution of Exp-Z and Pred-Z genes in 0.5-Mb windows along the Z chromosome. The locations of MHM regions 1 and 2 (MHM1 and MHM2, red lines) and miR-2954 (green line) are indicated; *P* value from a two-sample Kolmogorov–Smirnov test. All box plots show the median, 25th–75th percentiles and whiskers extending to 1.5× the IQR.

triplosensitivity (overexpression intolerance) and haploinsufficiency (deletion intolerance) for all human autosomal protein-coding genes[32], transferred to the chicken genome by means of corresponding Z-linked 1:1 orthologues (Supplementary Table 3). As expected, Z-linked ohnologues showed significantly higher triplosensitivity and haploinsufficiency scores than other Z-linked genes (Fig. 4d). Notably, non-ohnologue Z-linked experimental targets also exhibited significantly higher scores than other non-ohnologues, indicating that miR-2954 broadly regulates dosage-sensitive genes beyond the ohnologue subset. The strong upregulation of these genes in the $Z^{KO}Z^{KO}$ embryos probably underlies the observed lethality.

To further explore the lethality resulting from upregulated Z-linked miR-2954 targets in $Z^{KO}Z^{KO}$ embryos, we examined their spatiotemporal expression patterns. Expression pleiotropy, the breadth of gene expression across tissues and developmental stages, is a key predictor of mutational sensitivity, given that broader expression typically implies stronger functional constraint[33]. Using a developmental transcriptome resource from our laboratory[34] (Supplementary Table 6), we found that Z-linked targets exhibited significantly broader spatiotemporal expression than other Z-linked genes (Fig. 4e), reinforcing their essential roles and helping explain the severe phenotype in homozygous male KOs.

Consistent with these findings, functional analyses showed that Z-linked target genes are involved in critical developmental roles, participating in key biological pathways, such as JAK–STAT[35], PI3K–Akt[36], Rap1 (ref. 37), stem cell pluripotency signalling, growth hormone synthesis, and essential processes, including nervous system development, cell proliferation, apoptosis regulation, Golgi organization and cell adhesion (Supplementary Table 7). Notable examples include dosage-sensitive ohnologues, such as *PTCH1*, which encodes the Sonic Hedgehog receptor[38], and the transcription factor *KLF9* (ref. 39). The targets also include broadly expressed genes with essential cellular and developmental functions, such as *RAD23B*[40], and more specialized genes required for organ-specific development, such as *TAL2* for the brain[41] and *ALPK2* for the heart[42].

Given the substantial number of indirectly regulated or unpredicted direct targets (see above), we sought to comprehensively assess the effect of miR-2954 on Z-linked gene regulation. Using our previous RNA-seq dataset[34] (Supplementary Table 6), we identified 576 Z-linked genes expressed during chicken development and analysed their expression in homozygous male KOs, expanding beyond the initially predicted targets. Of these, 311 genes were significantly upregulated in at least one tissue (Benjamini–Hochberg $P_{adj.} < 0.1$; $\log_2[FC] > 0.2$), suggesting direct or indirect regulation by miR-2954. Notably, 64 of these genes were not computationally predicted as targets yet exhibited significantly higher triplosensitivity and haploinsufficiency scores than non-predicted non-differentially expressed genes (Extended Data Fig. 8 and Supplementary Table 3), indicating that they probably include true indirect or previously unrecognized direct targets. Combined with the 247 upregulated predicted targets consistently upregulated across tissues, approximately 54% (311 of 576) of developmentally expressed Z-linked genes were affected by miR-2954 loss, highlighting its broad regulatory scope and helping explain the severity of the KO phenotype.

Consistent with the broad spatiotemporal expression and diverse developmental roles of miR-2954 targets, $Z^{KO}Z^{KO}$ embryos exhibited various abnormalities before E7, including delayed growth, small eyes and brains and stunted axial elongation, with many embryos fully degraded by later stages (Supplementary Data 1 and 2). These phenotypes probably reflect the cumulative effects of upregulated dosage-sensitive Z-linked developmental genes, resulting in lethality by E5, a critical phase of increased blood circulation and rapid growth and differentiation of key structures, such as the allantois and vasculature. Disruptions in cardiovascular development, metabolic regulation, organogenesis and/or cell differentiation probably contribute to this lethality. Notably, the timing and variability of the miR-2954 KO phenotype resemble those in *Xist* KO mice, where loss of this lncRNA, which is essential for secondary dosage compensation in placental mammals, also causes lethality in a similar developmental window[43] (E5–E12).

Mechanisms of avian dosage compensation

The essential and highly specific role of miR-2954 in repressing dosage-sensitive Z-linked genes in males, along with the overall only partial transcriptional upregulation of Z-linked genes in females, whereas males retained ancestral expression levels[5,9], suggests an overall model for the evolution of dosage compensation in birds. We suggest that W chromosome gene loss during sex chromosome differentiation exerted selective pressure on females, driving transcriptional and potentially translational[6] upregulation of dosage-sensitive genes on a single Z, thereby restoring ancestral (proto-Z) expression levels. This upregulation also probably occurred in males, where two Z-linked gene copies caused transcript overabundance, prompting the evolution of a compensatory mechanism: miR-2954-mediated degradation of the excess transcripts. In this model, female gene upregulation and male-specific repression by miR-2954 would align the expression levels of dosage-sensitive Z-linked genes between sexes, yielding more balanced outputs than those observed for other Z-linked genes.

To test this model of dosage compensation evolution, we first investigated whether dosage-sensitive Z-linked genes targeted by miR-2954, that is, those upregulated in males upon miR-2954 KO, have also become upregulated in females over evolutionary time. We addressed this by comparing their current expression levels in female birds with their inferred ancestral levels before sex chromosome differentiation using an RNA-seq dataset for adult organs[34] (Methods). Ancestral (proto-Z) expression levels were estimated on the basis of the expression of autosomal 1:1 orthologues in mouse, an outgroup species—a method previously shown to reliably approximate ancestral gene expression patterns[4,5,9,44].

This analysis confirmed that Z-linked miR-2954 target genes have indeed become significantly upregulated in females over evolutionary time across several organs, with current-to-ancestral expression ratios exceeding 0.5 ($\log_2$ ratio greater than −1), unlike other Z-linked genes (Fig. 4f). In males, the expression levels of both targets and non-targets remained close to the ancestral levels ($\log_2$ ratio ≈ 0), consistent with the presence of two Z copies and miR-2954-mediated repression, which was absent in the KO model. Notably, analyses of brain Ribo-seq data showed even stronger upregulation of Z-linked targets in females, with expression levels statistically indistinguishable from the ancestral values (Extended Data Fig. 5b). These findings indicate that the combined transcriptional and translational upregulation in females can fully restore ancestral (proto-Z) expression levels, similar to the scenario we previously reported for X-linked genes in therian mammals[6].

We next tested the final prediction of our dosage compensation model: Z-linked target genes, owing to the interplay of compensatory mechanisms, should exhibit more similar expression levels between the sexes than other Z-linked genes. Using a comprehensive RNA-seq dataset covering several organs and developmental stages[34], we found that female-to-male expression ratios were indeed significantly higher for Z-linked targets across all samples (Fig. 4g), indicating a more sex-balanced expression of these genes. Notably, Ribo-seq data from the adult brain and embryonic head showed even higher female-to-male ratios than RNA-seq data, with expression levels in the adult brain statistically indistinguishable between sexes for Z targets ($\log_2$ ratio = 0; Extended Data Fig. 5c). These findings indicate that translational upregulation in females, reflected by higher translational efficiencies of Z-linked mRNAs in this sex (Extended Data Fig. 5d; see also ref. 45), together with transcriptional

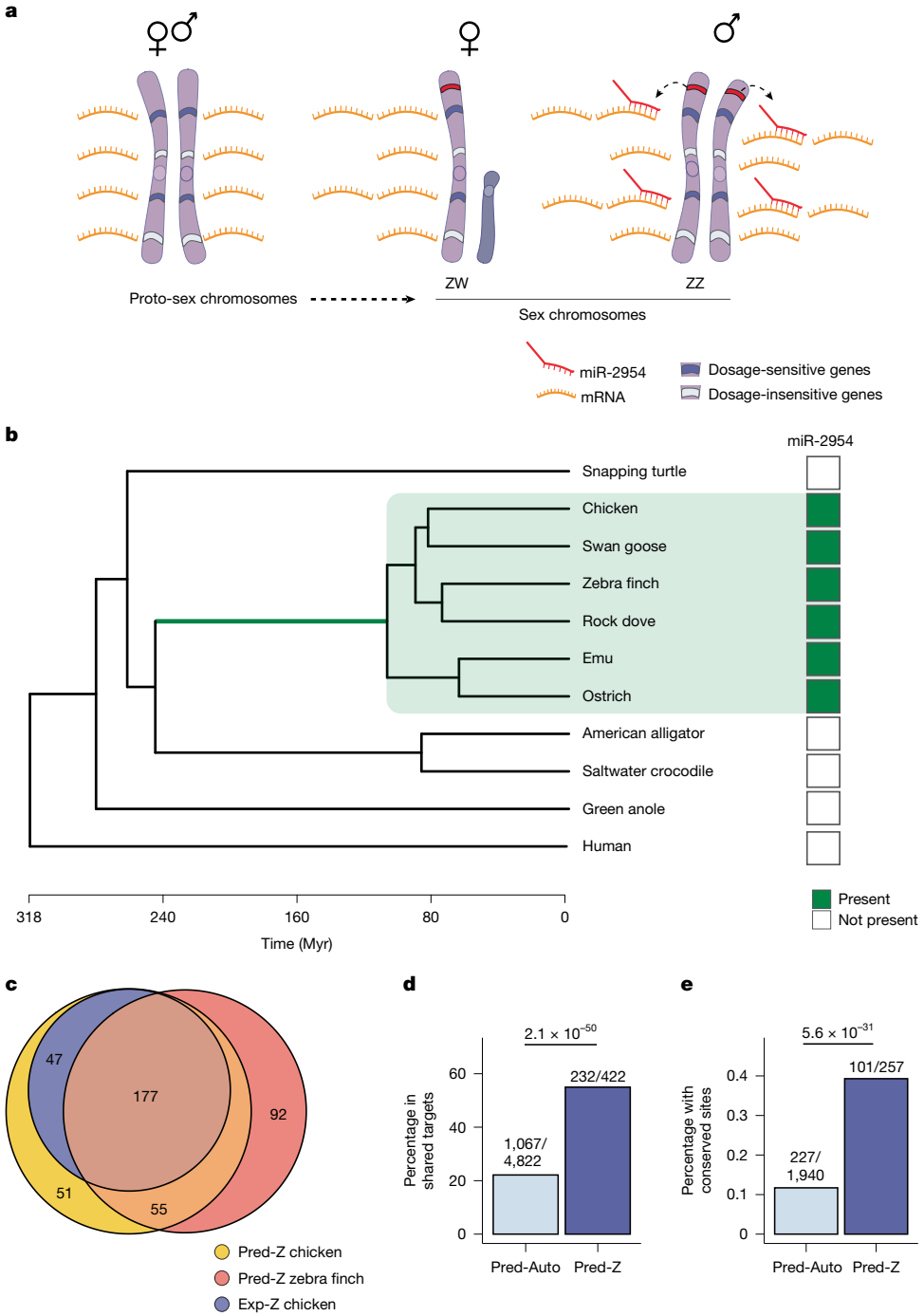


Fig. 5 | Evolution of avian dosage compensation and the emergence and conservation of miR-2954 and its targets. a, Illustration of the upregulation of dosage-sensitive Z-linked genes in birds following W gene loss and the concurrent evolution of miR-2954 to mitigate excess transcript accumulation in males. **b**, Phylogenetic distribution of miR-2954. The tree includes representative avian and other amniote species, with branch lengths scaled to represent the evolutionary time (million years). miR-2954 originated in the avian stem lineage (green branch). **c**, Overlap between miR-2954 targets identified by experimental data in chicken (Exp-Z) and predicted Z-linked targets in chicken and zebra finch (Pred-Z). The Venn diagram shows the number of shared and unique targets across datasets. **d**, Proportions of miR-2954 target genes on autosomes (Pred-A) and Z chromosome (Pred-Z) in chicken that are shared with zebra finch. The two-sided χ^2 test P value is shown above. **e**, Proportions of Pred-A and Pred-Z genes with at least one target site conserved between chicken and zebra finch. The two-sided χ^2 test P value for the comparison is indicated above.

upregulation in both sexes and miR-2954-mediated mRNA degradation in males, can lead to full protein-level dosage compensation of Z targets.

Finally, we examined the genomic distribution of Z-linked targets along the Z chromosome and found them to be relatively evenly dispersed (Fig. 4h), with no notable clustering around the two previously identified male hypermethylated (MHM) regions specific to galliform birds (chicken and turkey), which have been proposed as localized dosage-compensated regions in this lineage[46,47].

Altogether, our findings strongly support a model for the evolution of dosage compensation in the avian ZW system, in which W chromosome gene loss in females drove the transcriptional upregulation of dosage-sensitive Z-linked genes in both sexes, accompanied by translational upregulation in females. In males, this overexpression was

secondarily balanced by the evolution of a targeted miR-2954-mediated transcript degradation mechanism (Fig. 5a).

Evolution of miR-2954-mediated dosage compensation

To investigate the origin and evolution of this avian dosage compensation system, we first screened for the presence of the miR-2954 locus across amniotes (Methods). We found that miR-2954 is present in all surveyed bird genomes but is absent from all non-avian species, including crocodiles, the closest living relatives of birds, which carry *XPA*, the miR-2954 host gene in birds (Fig. 5b and Extended Data Fig. 9). Remarkably, both the mature (22 nt) and precursor (68 nt) sequences of miR-2954 are identical in 351 of the 363 bird species examined[48]. The remaining 12 species include 2 with a single nucleotide substitution in the precursor region (outside the mature sequence) and 10 with missing sequence data for this locus (Supplementary Data 3).

To further evaluate the conservation of miR-2954-mediated dosage compensation, we compared its features between chicken and zebra finch, representatives of the two major avian lineages Galliformes/Anseriformes and Neoaves, which diverged approximately 91 million years ago[49] and mark the deepest split within Neognathae. Approximately 78% of experimentally validated Z-linked targets in chicken are also computationally predicted targets of miR-2954 in zebra finch, a species with previously documented male-biased miR-2954 expression[50] (Fig. 5c). Moreover, Z-linked predicted targets in zebra finch are significantly more likely than autosomal ones to overlap with chicken targets (Fig. 5d) and show a higher proportion of conserved binding sites (Fig. 5e). Alongside our earlier finding that Z-linked targets, especially ohnologues, are preferentially upregulated following miR-2954 knockdown in a male zebra finch cell line[3], these results strongly support a conserved functional role for miR-2954 across species.

Further analyses revealed that miR-2954 target genes are also enriched on the Z chromosome of the ostrich, a flightless ratite representing Palaeognathae, the sister lineage to all other extant birds (Neognathae), which exhibits less pronounced sex chromosome differentiation[51] (Extended Data Fig. 10a). By contrast, the autosomal orthologues of Z-linked genes in non-avian reptiles and mammals (such as crocodiles and humans) show no enrichment. Furthermore, target site densities are similar between human Z orthologues, other human autosomal genes and chicken autosomal genes but significantly higher for chicken Z-linked genes (Extended Data Fig. 10b). These findings indicate that miR-2954 target sites began accumulating specifically on the Z chromosome in the common ancestor of modern birds, following the emergence of miR-2954.

Altogether, our evolutionary analyses indicate that the avian sex chromosome dosage compensation system, including the secondary miR-2954-mediated mechanism, emerged in the avian stem lineage between approximately 108 and 245 million years ago[49] (Fig. 5b), coinciding with the origin of the ZW sex chromosomes, and has been selectively maintained across extant bird species.

Discussion

In this study, using a chicken KO model, we uncovered the function of miR-2954 and the evolutionary mechanisms underlying avian dosage compensation. We propose a scenario in which the decay of the W chromosome led to transcriptional and translational upregulation of dosage-sensitive Z-linked genes in (ZW) females and secondarily to transcriptional upregulation in (ZZ) males. To counteract the resulting transcript overabundances in males, a highly targeted miR-2954-mediated degradation mechanism evolved. During the co-evolution of miR-2954 and its targets, this miRNA acquired high and widespread expression in males, along with the emergence of strong and often several binding sites in the regulated genes, enabling effective suppression and restoring their transcript abundances to ancestral levels. Our findings indicate that the miR-2954 mechanism is conserved across birds, consistent with its crucial role in avian males. Indeed, in chicken, complete miR-2954 KO leads to early embryonic lethality, presumably because of the upregulation of numerous dosage-sensitive genes with key developmental functions that are normally repressed by this miRNA.

The avian ZW dosage compensation system unveiled here parallels that of the therian mammalian XY system[1,4,6] and aligns with Ohno's original hypothesis[52]. In both systems, gene expression upregulation was triggered by the degeneration of sex-specific chromosomes—the W in female birds and the Y in male mammals. Unlike in the XY systems of the anole lizard[9] or fruitflies[53], where upregulation is restricted to the heterogametic sex, birds and mammals exhibit upregulation in both sexes. This non-sex-specific upregulation necessitated the evolution of secondary silencing mechanisms: in female mammals through *XIST* and *RSX* lncRNAs[1,8,43] and in male birds through miR-2954. Although both systems involve non-coding RNAs essential for the viability of the heterogametic sex, they differ markedly; birds use targeted post-transcriptional repression of dosage-sensitive Z-linked genes, whereas mammals use broad transcriptional inactivation of the X chromosome. Future studies may clarify why birds and mammals evolved such distinct solutions, possibly reflecting differences in the distribution of dosage-sensitive genes on their respective ancestral autosomes.

Altogether, our study has unveiled a crucial and previously unrecognized role for a miRNA in dosage compensation. Thus, rather than acting as a 'sculptor' of the transcriptome[15], like many miRNAs that have important but more subtle regulatory effects on gene expression[23], miR-2954 has evolved into a key mediator of a targeted post-transcriptional silencing network that has ensured the survival of males in the wake of avian sex chromosome differentiation.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-09256-9.

1. Graves, J. A. Evolution of vertebrate sex chromosomes and dosage compensation. *Nat. Rev. Genet.* **17**, 33–46 (2016).
2. Cortez, D. et al. Origins and functional evolution of Y chromosomes across mammals. *Nature* **508**, 488–493 (2014).
3. Warnefors, M. et al. Sex-biased microRNA expression in mammals and birds reveals underlying regulatory mechanisms and a role in dosage compensation. *Genome Res.* **27**, 1961–1973 (2017).
4. Necsulea, A. & Kaessmann, H. Evolutionary dynamics of coding and non-coding transcriptomes. *Nat. Rev. Genet.* **15**, 734–748 (2014).
5. Julien, P. et al. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biol.* **10**, e1001328 (2012).
6. Wang, Z. Y. et al. Transcriptome and translatome co-evolution in mammals. *Nature* **588**, 642–647 (2020).
7. Grant, J. et al. Rsx is a metatherian RNA with Xist-like properties in X-chromosome inactivation. *Nature* **487**, 254–258 (2012).
8. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–137 (1996).
9. Marin, R. et al. Convergent origination of a *Drosophila*-like dosage compensation mechanism in a reptile lineage. *Genome Res.* **27**, 1974–1987 (2017).
10. Ellegren, H. et al. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. *BMC Biol.* **5**, 40 (2007).
11. Itoh, Y. et al. Dosage compensation is less effective in birds than in mammals. *J. Biol.* **6**, 2 (2007).
12. Rodriguez-Montes, L. et al. Sex-biased gene expression across mammalian organ development and evolution. *Science* **382**, eadf1046 (2023).
13. Zimmer, F., Harrison, P. W., Dessimoz, C. & Mank, J. E. Compensation of dosage-sensitive genes on the chicken Z chromosome. *Genome Biol. Evol.* **8**, 1233–1242 (2016).
14. Zhao, D. et al. Somatic sex identity is cell autonomous in the chicken. *Nature* **464**, 237–242 (2010).
15. van de Lavoir, M. C. et al. Germline transmission of genetically modified primordial germ cells. *Nature* **441**, 766–769 (2006).
16. Ballantyne, M. et al. Direct allele introgression into pure chicken breeds using sire dam surrogate (SDS) mating. *Nat. Commun.* **12**, 659 (2021).

17. Idoko-Akoh, A., Taylor, L., Sang, H. M. & McGrew, M. J. High fidelity CRISPR/Cas9 increases precise monoallelic and biallelic editing events in primordial germ cells. *Sci. Rep.* **8**, 15126 (2018).
18. Woodcock, M. E., Idoko-Akoh, A. & McGrew, M. J. Gene editing in birds takes flight. *Mamm. Genome* **28**, 315–323 (2017).
19. Kleinstiver, B. P. et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
20. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
21. Idoko-Akoh, A. & McGrew, M. J. Generation of genome-edited chicken through targeting of primordial germ cells. *Methods Mol. Biol.* **2631**, 419–441 (2023).
22. Martin, S. et al. Nanopore adaptive sampling: a tool for enrichment of low abundance species in metagenomic samples. *Genome Biol.* **23**, 11 (2022).
23. Bartel, D. P. Metazoan microRNAs. *Cell* **173**, 20–51 (2018).
24. Hausser, J. & Zavolan, M. Identification and consequences of miRNA-target interactions—beyond repression of gene expression. *Nat. Rev. Genet.* **15**, 599–612 (2014).
25. Shang, R., Lee, S., Senavirathne, G. & Lai, E. C. MicroRNAs in action: biogenesis, function and regulation. *Nat. Rev. Genet.* **24**, 816–833 (2023).
26. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, e05005 (2015).
27. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
28. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
29. Wolfe, K. Robustness—it's not where you think it is. *Nat. Genet.* **25**, 3–4 (2000).
30. Ohno, S. *Evolution by Gene Duplication* (Allen & Unwin, 1970).
31. McLysaght, A. et al. Ohnologs are overrepresented in pathogenic copy number mutations. *Proc. Natl Acad. Sci. USA* **111**, 361–366 (2014).
32. Collins, R. L. et al. A cross-disorder dosage sensitivity map of the human genome. *Cell* **185**, 3041–3055.e25 (2022).
33. Carroll, S. B. Evolution at two levels: on genes and form. *PLoS Biol.* **3**, 1159–1166 (2005).
34. Cardoso-Moreira, M. et al. Gene expression across mammalian organ development. *Nature* **571**, 505–509 (2019).
35. Hou, S. X., Zheng, Z., Chen, X. & Perrimon, N. The Jak/STAT pathway in model organisms: emerging roles in cell movement. *Dev. Cell* **3**, 765–778 (2002).
36. Yu, J. S. & Cui, W. Proliferation, survival and metabolism: the role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination. *Development* **143**, 3050–3060 (2016).
37. Bos, J. L., de Rooij, J. & Reedquist, K. A. Rap1 signalling: adhering to new models. *Nat. Rev. Mol. Cell Biol.* **2**, 369–377 (2001).
38. Jiang, J. & Hui, C. C. Hedgehog signaling in development and cancer. *Dev. Cell* **15**, 801–812 (2008).
39. Drepanos, L. et al. Loss of Kruppel-like factor 9 deregulates both physiological gene expression and development. *Sci. Rep.* **13**, 12239 (2023).
40. Ng, J. M. et al. Developmental defects and male sterility in mice lacking the ubiquitin-like DNA repair gene mHR23B. *Mol. Cell Biol.* **22**, 1233–1245 (2002).
41. Bucher, K. et al. The T cell oncogene *Tal2* is necessary for normal development of the mouse brain. *Dev. Biol.* **227**, 533–544 (2000).
42. Hofsteen, P. et al. ALPK2 promotes cardiogenesis in zebrafish and human pluripotent stem cells. *iScience* **2**, 88–100 (2018).
43. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11**, 156–166 (1997).
44. Murat, F. et al. The molecular evolution of spermatogenesis across mammals. *Nature* **613**, 308–316 (2023).
45. Papanicolaou, N. et al. Multi-layered dosage compensation of the avian Z chromosome. Preprint at *bioRxiv* https://doi.org/10.1101/2024.08.20.608780 (2024).
46. Itoh, Y., Kampf, K. & Arnold, A. P. Possible differences in the two Z chromosomes in male chickens and evolution of MHM sequences in Galliformes. *Chromosoma* **120**, 587–598 (2011).
47. Sun, D., Maney, D. L., Layman, T. S., Chatterjee, P. & Yi, S. V. Regional epigenetic differentiation of the Z chromosome between sexes in a female heterogametic system. *Genome Res.* **29**, 1673–1684 (2019).
48. Stiller, J. et al. Complexity of avian evolution revealed by family-level genomes. *Nature* **629**, 851–860 (2024).
49. Kumar, S. et al. TimeTree 5: an expanded resource for species divergence times. *Mol. Biol. Evol.* **39**, msac174 (2022).
50. Luo, G. Z. et al. Genome-wide annotation and analysis of zebra finch microRNA repertoire reveal sex-biased expression. *BMC Genomics* **13**, 727 (2012).
51. Zhou, Q. et al. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* **346**, 1246338 (2014).
52. Ohno, S. *Sex Chromosomes and Sex-Linked Genes* (Springer-Verlag, 1967).
53. Conrad, T. & Akhtar, A. Dosage compensation in *Drosophila melanogaster*: epigenetic fine-tuning of chromosome-wide transcription. *Nat. Rev. Genet.* **13**, 123–134 (2012).

Methods

Ethics information

All animal procedures were conducted in compliance with national and international ethical guidelines and regulations. Mouse experiments were approved by the local animal welfare authorities at Heidelberg University Interfaculty Biomedical Research Facility (T-64/17). Chicken experiments were conducted under UK Home Office licence PP9565661 and approved by the Roslin Institute Animal Welfare and Ethical Review Board Committee and the Linköping Council for Ethical Licensing of Animal Experiments (288-2019). Mice (*Mus musculus*; strain CD-1; RjOrl:SWISS; RRID: MGI:5603077) were purchased from Janvier Labs and euthanized by means of cervical dislocation. All chicken (Hy-Line Brown; *Gallus gallus*) management, maintenance and embryo manipulation followed the relevant regulatory guidelines.

Isolating, sexing and culturing PGCs

Genome editing in chickens involves the derivation and culturing of PGCs, performing genome editing on these cells and the subsequent injection of the edited cells into surrogate hosts depleted of their native PGCs[21]. Following the injection of the genetically edited PGCs into the gonads of sterile surrogate hosts, the resulting offspring will inherit the genetic modifications introduced into the PGCs[21] (Fig. 1b and Extended Data Fig. 1). To establish miR-2954 KO lines, ten PGC lines were derived from the blood of Hy-Line Brown chicken embryos at Hamburger–Hamilton stage 16 (E2.5) and cultured according to previously described methods[21]. The sex of the PGC lines was determined according to previous studies[21,54] on the basis of two sets of primers for one W-chromosome-specific gene and one autosomal gene (the control), respectively; the latter serves as a control for polymerase chain reaction (PCR) success (Supplementary Table 1). We cultured four male PGC lines and subsequently randomly selected one line for the KO experiment. This PGC line was cultured for 22 days in total before transfection.

Design of sgRNA and homology-directed repair template

Inducing double-stranded breaks at specific genomic loci, followed by homology-directed repair using a template, introduces precise nucleotide substitutions[20]. Using CHOPCHOP v.2 (ref. 55), we designed and tested five custom sgRNAs (Supplementary Table 1) to target the miR-2954 (MIR2954) locus (Gene ID: 100498678), located within the second intron of the DNA damage recognition and repair factor gene, *XPA* (ENSGALG00010009534), on the forward strand of chromosome Z (location: NC_052572.1: 71305174-71305241; reference genome: bGalGal1.mat.broiler.GRCg7b (GCF_016699485.2)). Additionally, we designed one single-stranded DNA oligonucleotide (ssODN) sequence as a repair template to exploit the homology-directed repair pathway. The ssODN repair template consisted of Ultramer DNA Oligonucleotides, custom-synthesized by Integrated DNA Technologies. The ssODN template contained homology arms flanking miR-2954, designed specifically to introduce a 36-bp deletion encompassing the entire mature miR-2954 sequence and part of its flanking pre-miRNA sequence. Additionally, we incorporated an EcoRI restriction endonuclease site (5′-GAATTC-3′) into this ssODN (Supplementary Table 1). These modifications effectively knock out miR-2954 and allow PCR-based genotyping for successful deletion events in both PGCs and the derived chickens (Extended Data Fig. 1 and Supplementary Table 1).

Genotyping

We designed PCR primers to amplify a 550-bp region within the second intron of the *XPA* gene, encompassing the targeted deletion site (Supplementary Table 1) using Primer-BLAST[56]. EcoRI restriction endonuclease enzyme specifically recognizes and cuts DNA at the restriction site (5′-GAATTC-3′). Following EcoRI digestion of this PCR product and subsequent gel electrophoresis, we expected to observe a single 550-bp band in wild-type individuals (ZZ and ZW) owing to the absence of the EcoRI restriction site, three bands (550, 298 and 221 bp) in heterozygote KO individuals ($Z^{KO}Z$) owing to digestion of half of the product and two bands (298 and 221 bp) in homozygote males ($Z^{KO}Z^{KO}$) and hemizygote females ($Z^{KO}W$) owing to complete EcoRI restriction site digestion. This differential PCR band pattern served as a molecular signature for genotyping the individuals. PCR was performed using Phusion High-Fidelity PCR Master Mix with GC Buffer from New England Biolabs, in accordance with the manufacturer's guidelines. The reaction mixture was prepared using 1.25 µl of 10 µM forward primer, 1.25 µl of 10 µM reverse primer, 0.75-µl dimethyl sulfoxide, 12.5 µl of 2X Phusion Master Mix and approximately 100 ng of DNA in 1 µl of water. The thermal cycling conditions were set as follows: an initial denaturation at 98 °C for 60 s, followed by 35 cycles of 98 °C for 10 s, 62 °C for 20 s and 72 °C for 20 s, concluding with a final extension at 72 °C for 10 min. To perform genotyping, we first extracted DNA from approximately 10,000 PGCs or embryonic tissues using DNeasy Blood & Tissue Kits from QIAGEN, according to the manufacturer's protocol. We then conducted PCRs as described above and subjected the PCR products to EcoRI digestion using EcoRI-HF and rCutSmart buffer from New England Biolabs, following the manufacturer's guidelines. Each reaction consisted of 5 µl of the PCR product, 1 µl of EcoRI-HF, 1 µl of rCutSmart buffer and 8 µl of water. The reactions were incubated at 37 °C for 30 min, followed by a 5-min heat inactivation at 65 °C. Alternatively, the genotypes of several samples were analysed on the basis of the size of the undigested PCR products using the Agilent Fragment Analyzer system. In this approach, a 550-bp band represented ZZ and ZW, a 520-bp band represented $Z^{KO}Z^{KO}$ and $Z^{KO}W$ and two bands (550 and 520 bp) in $Z^{KO}Z$.

PGC transfection, selection and clonal expansion

We used a high-fidelity Cas9 variant (SpCas9-HF1), which significantly reduces off-target effects compared to wild-type Cas9 (ref. 19). For the expression of SpCas9-HF1 and sgRNAs in PGCs, we used the HF-PX459 (V2) expression vector, which also bears puromycin resistance as an antibiotic selection gene[17] (Addgene plasmid 118632). We cloned all five sgRNAs individually into the plasmids according to previous descriptions[17,20] and then tested the effectiveness of three of these plasmids harbouring sgRNAs 1–3. We transfected 1.5 µg of the vector and 0.5 µg of ssODNs into approximately 100,000 Hy-Line Brown PGCs using Lipofectamine 2000 transfection reagent (Thermo Fisher Scientific). After 24 h in culture, the cells were treated with 0.6 µg ml^{-1} of puromycin for 48 h for the selection of successfully transfected cells. We cultured these cells for around 2 weeks and then genotyped them for the presence of deletions through EcoRI digestion of the PCR product. Using gRNA3, we observed a strong PCR band at 550 bp and two faint bands at approximately 300 and 220 bp. This pattern suggested the incorporation of the ssODN template in a subset of transfected PGCs. Accordingly, these PGCs were sorted using the BD FACSAria III Cell Sorter (BD Biosciences) into a 96-well plate at a rate of one cell per well to identify the clonal populations with the deletion of miR-2954. After 3 weeks of culturing, we screened the genotypes of 42 clonal PGC populations that survived and propagated. We identified four $Z^{KO}Z$ and two $Z^{KO}Z^{KO}$ clonal populations among them (6 of 42 clones were targeted). Subsequently, we cryopreserved the homozygote and heterozygote populations following established protocols[21] and used one of the $Z^{KO}Z^{KO}$ populations for confirmation of the deletion and injection to surrogate hosts to generate the KO animals. To confirm the deletion of miR-2954, we performed PCR on the DNA obtained from the PGC line before transfection. The selected clonal $Z^{KO}Z^{KO}$ PGC population and the resulting PCR products were sequenced by Eurofins Genomics using their Sanger sequencing services (TubeSeq Service). Analysis of the sequences confirmed the deletion of miR-2954 and integration of the EcoRI site in accordance with the design of the provided ssODN repair template (Extended Data Fig. 1).

Generation of the G0 rooster

$Z^{KO}Z^{KO}$ PGCs were injected into surrogate host embryos using our previously described method[21]. In brief, we thawed the cryopreserved clonal $Z^{KO}Z^{KO}$ PGCs 7 days before the intended injection date and propagated them to a density of approximately 150,000 cells per well in a 24-well tissue culture plate. These cultured PGCs were pelleted by means of standard centrifugation and then resuspended in the PGC culture medium to achieve a concentration of 5,000 cells per microlitre. To this suspension, we added 0.1 µl of the chemical compound AP20187 (B/B) (25 mM) per 5 µl of PGC suspension. Approximately 1 µl of this mixture was aspirated into a microcapillary injection tube and injected into each iCaspase9 sterile embryo[16] at Hamburger–Hamilton stages 15 and 16. AP20187 (B/B), present in the injected PGC mixture, induces the dimerization of the FK506-binding protein, leading to the activation of the attached caspase-9 protein and the induced apoptotic cell death of the endogenous PGCs in the iCaspase9 sterile embryos, thereby allowing the colonization of gonads by the injected $Z^{KO}Z^{KO}$ PGCs[16]. Injecting the clonal PGCs into 20 iCaspase9 sterile embryos resulted in hatching of 7 G0 chicks comprising 1 male and 6 females.

Generation of miR-2954 KO chickens

We maintained the male G0 and raised it to sexual maturity. This G0 was then paired with six Hy-Line Brown hens (same breed), producing $Z^{KO}Z$ and $Z^{KO}W$ individuals (OC G1). We then raised five male and six female OC G1 individuals to sexual maturity. One of these males was mated with the OC G1 females to generate second-generation (G2) embryos ($Z^{KO}Z$, $Z^{KO}Z^{KO}$, ZW or $Z^{KO}W$) that were used for viability studies and tissue collection for gene expression analyses. A second OC G1 male, not involved in generating G2 individuals, was mated with six Hy-Line Brown females. This pairing produced OC G2 individuals for the genotypes ZZ, $Z^{KO}Z$, ZW and $Z^{KO}W$. Finally, upon reaching sexual maturity, a OC G2 $Z^{KO}Z$ rooster was mated with six OC G2 $Z^{KO}W$ hens to produce G3 embryos ($Z^{KO}Z$, $Z^{KO}Z^{KO}$, ZW or $Z^{KO}W$). These G3 embryos were then used to confirm the phenotypes observed in the G2 generation (Fig. 1a).

Selection and processing of chicken embryos and tissues for RNA-seq analysis

Upon completing the genotyping and sexing of G2 embryos, we selected 36 embryos for RNA-seq. This selection included 18 E2 embryos (9 males and 9 females) (Hamburger–Hamilton stage 12), 9 E3 males (Hamburger–Hamilton stages 18 and 19) and 9 E5 males (Hamburger–Hamilton stages 24 and 25). For the E2 cohort, RNA extraction was performed on whole embryos after the removal of extra-embryonic membranes. This cohort included nine female embryos of various genotypes (three ZW, three $Z^{KO}W$ and three pure Hy-Line Brown ZW embryos (female embryos from the original stock), as a control for maternal effects on gene expression), and nine male embryos (three $Z^{KO}Z$, three $Z^{KO}Z^{KO}$ and three ZZ genotypes). Given the low expression of miR-2954 in females and their survival, we then focused on gene expression in males. For the E3 and E5 cohorts, we investigated tissue-specific gene expression by dissecting the head, heart and rest of the body (referred to as the body) from each male embryo under a stereomicroscope, with all dissections performed in ice-cold PBS. Each tissue type from each embryo was represented by three replicates derived from three individuals. We note that all ZZ are pure Hy-Line Brown, and all other genotypes (ZW, $Z^{KO}W$, $Z^{KO}Z$ and $Z^{KO}Z^{KO}$) are G2.

RNA extraction and sequencing

A total of 72 samples from E2, E3 and E5 embryos were used for the generation of RNA-seq libraries. We extracted total RNA from whole embryos or dissected tissues using the AllPrep DNA/RNA/miRNA Universal Kit (QIAGEN), following the manufacturer's protocols. The RNA quality was assessed using the Fragment Analyzer system (Agilent), and all RNA quality numbers were equal to 10, indicating a lack of degradation. The RNA-seq libraries were prepared from 400 ng of RNA per sample using the NEBNext Ultra II RNA Library Prep Kit for Illumina sequencing on an Illumina NextSeq 2000 system, using NextSeq 2000 P3 Reagents (100 cycles), with samples multiplexed in two sets of 36.

Additionally, we generated small RNA libraries using RNA derived from the same E5 male samples (which were also used to generate RNA-seq libraries). This included the generation of small RNA libraries for RNA derived from ZZ (two replicates), $Z^{KO}Z$ (three replicates) and $Z^{KO}Z^{KO}$ (three replicates) for each tissue type (head, body and heart, respectively). These libraries were prepared using the NEBNext Small RNA Library Prep Set for Illumina and were sequenced on an Illumina NextSeq 550 system using NextSeq 500/550 High Output Kit v.2.5 (75 cycles), with samples multiplexed in two sets of 12.

Estimation of gene expression levels

The chicken reference genome (bGalGal1.mat.broiler.GRCg7b; GCA_016699485.1) and corresponding gene transfer format (GTF) annotation file were obtained from Ensembl[57] (release 109). Raw reads from each library were aligned to the reference genome using STAR aligner v.2.7.2b (ref. 58). This alignment process involved generating STAR indices, aligning reads to the reference genome in an annotation-aware manner and quantifying the number of reads mapped to each gene using the quantMode GeneCounts option in STAR. The median uniquely mapped reads number across all samples was 34,703,339. The resulting gene count matrices, along with a metadata file containing sample information and the GTF file, were used to create a RangedSummarizedExperiment object. This object was imported into DESeq2 v.1.24.0 (ref. 58) for downstream analysis. Gene expression data were normalized using variance-stabilizing transformation (VST) through the vsn package v.3.52.0 in R v.4.1 (ref. 59) implemented in the DESeq2 package. Subsequently, principal component analysis (PCA) was conducted as implemented in the DESeq2 package to examine sample relationships and identify potential outliers. The PCA results revealed a clear clustering of samples (including biological replicates) for the respective tissues and ages without outliers, supporting the high quality of the expression data (Extended Data Fig. 3a).

Raw short RNA-seq data were preprocessed using a custom Bash script. Adaptor sequences were trimmed and reads were size-selected using Cutadapt v.4.4. The parameters set a maximum error rate of 0.25, targeted the adaptor sequence AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC with a minimum overlap of 6 nucleotides and allowed no indels while selecting for read lengths between 19 and 26 nucleotides. After trimming and size selection, the reads were aligned to the chicken reference genome using STAR following the ENCODE miRNA-seq pipeline[60] (www.encodeproject.org/microrna/microrna-seq-encode4/) (May 2017). This alignment process included mapping to the miRNA subset of the chicken GTF gene annotation and quantifying the number of aligned reads in STAR. The median of the number of uniquely mapped reads across all samples was 432,069.

MiRNA target prediction in chicken, zebra finch, ostrich, crocodile and human

To identify potential targets of miR-2954, we used TargetScan[26], which detects 6mer, 7mer-1a, 7mer-m8 and 8mer-1a target sites in the 3′ UTRs of mRNA transcripts, aligning them with the miRNA seed sequence. We obtained 3′ UTR sequences for all splice variants of genes within both the chicken (bGalGal1.mat.broiler.GRCg7b), zebra finch (bTaeGut1_v1.p), Australian saltwater crocodile (CroPor_comp1), African ostrich (ASM69896v1) and human (GRCh38.p14) genomes using BioMart (ref. 61). Subsequent identification of target sites was performed using TargetScan v.7.0 for each species. A gene was categorized as a predicted target if it contained any of these target site types within its UTRs. We then counted the total number of target sites for each predicted target gene in chicken. To calculate context+ scores, we performed a separate target prediction step specifically for chicken using TargetScan v.6.0,

along with its associated Perl script (targetscan_60_context_scores.pl). We used the same chicken 3′ UTR sequences in this step as in the initial TargetScan analysis. Finally, we calculated the median context+ score for all target sites within each gene, considering only those with a context+ score of less than 0 (Supplementary Table 3).

Identification of conserved target sites in chicken and zebra finch

To identify conserved target sites, we selected the longest annotated 3′ UTR for each gene in both the chicken (*G. gallus*) and zebra finch (*Taeniopygia guttata*) genomes. These 3′ UTR sequences were then aligned using Clustal Omega. We subsequently used TargetScan v.7.0 to predict conserved target sites within the aligned sequences (Supplementary Table 3).

Differential gene expression analysis

The 3′ UTR is specific to protein-coding genes, and miRNA targets are predicted on the basis of the presence of target sites within their 3′ UTRs. Consequently, we limited the DESeq2 dataset to protein-coding genes (as identified in the GTF annotation). Differential expression analysis was conducted using DESeq2. Differentially expressed genes were identified using a threshold of less than 0.05 for $P_{adj.}$ according to the Benjamini–Hochberg method[62]. The effect of genotype on gene expression in E2 whole embryos was independently analysed in male and female embryos (model: gene expression as a function of genotype). For each tissue (head, heart and body), gene expression analysis was performed collectively across ages using a model that included both genotype and embryonic age as variables (model: gene expression ≈ genotype + embryonic age). The log fold changes and differentially expressed genes were determined for each genotype contrast (Supplementary Table 4).

Differential miRNA expression analysis

Differential expression analysis was conducted using DESeq2. Differentially expressed miRNAs were identified using a threshold of less than 0.05 for $P_{adj.}$ according to the Benjamini–Hochberg method[62]. The effect of genotype on miRNA expression in E5 head, heart and body was independently analysed in each tissue (model: gene expression as a function of genotype) (Supplementary Table 5).

Comparison of pure Hy-Line Brown females with ZW G2

Although all chickens used in the gene expression analysis were of the Hy-Line Brown breed, the G2 animals, comprising genotypes ZW, $Z^{KO}W$, $Z^{KO}Z$ and $Z^{KO}Z^{KO}$, originated from different parents compared with the ZZ genotype, which was derived from the pure Hy-Line Brown breed (the original stock). To ensure the rigour of all expression comparisons, we aimed to confirm that the G2 ZW and pure Hy-Line Brown ZW had similar gene expression profiles (ZZ embryos cannot be derived from the G2 (hemizygous/heterozygous KO) parents), thereby eliminating potential confounding factors, such as maternal effects on gene expression. Accordingly, we conducted different expression analyses between pure Hy-Line Brown ZW and G2 ZW chickens and compared the fold changes across different gene categories. This analysis confirmed that gene expression patterns are statistically indistinguishable between Hy-Line Brown and G2 and therefore do not confound our results (Extended Data Fig. 3b).

Identifying ohnologues

The list of chicken ohnologues was retrieved from the OHNOLOGS v.2 database[63], available at http://ohnologs.curie.fr/ ('relaxed' dataset). These ohnologues were identified using gene IDs from the galGal4 assembly (Ensembl release 80), which is incompatible with the gene IDs of the chicken genome assembly used in our study (GRCG7b). To resolve this, we retrieved the unspliced DNA sequences of these ohnologue gene IDs from the GRCg6a assembly (Ensembl release 106) through BioMart. Subsequently, these sequences were aligned to the unspliced DNA sequences of protein-coding genes from the GRCG7b assembly using BLASTn (BLAST+ 2.4)[64], with the settings -perc_identity 95 and -evalue 0.001. We sorted the results by bit scores to identify the best hits between the two gene sets. Cross-referencing protein names for matched gene IDs confirmed a high accuracy (88.6% exact matches) of this ID conversion method (Supplementary Table 3).

Dosage sensitivity scores

Dosage sensitivity scores for human genes, including haploinsufficiency (pHaplo) and triplosensitivity (pTriplo), were sourced from a previous study[32]. These scores were then assigned to chicken genes on the basis of their 1:1 orthology relationship (retrieved using BioMart) (Supplementary Table 3).

Assessment of time and tissue specificity

To evaluate the time and tissue specificity of chicken genes, we calculated time and tissue specificity indexes on the basis of the tau metric[65] using a developmental time-series RNA-seq dataset[34] (Supplementary Table 6). As in previous studies[34], for the tissue specificity index, the tau metric was applied to the maximum expression of the gene observed during development in each organ, whereas for the time specificity index, the tau metric was applied to the expression of the gene at different time points instead of organs. In both cases, indexes range from 0 (indicating broad expression) to 1 (indicating restricted expression).

Identification of developmentally expressed genes and female-to-male expression level ratios

Gene expression ratios between the sexes were analysed using a published RNA-seq time-series dataset[34,66,67]. We obtained raw read (FASTQ) files for various chicken organs (blastoderm, brain, cerebellum, gonads, heart, kidney and liver) across different embryonic stages (E0, E4.5 and E6 for gonads and E10, E12, E14 and E17) and post-hatch periods (P0, P7, P35, P70 and P155). Reads were aligned to the bGalGal1.mat.broiler.GRCg7b reference genome, with read counts generated as detailed in the 'Estimation of gene expression levels' section. We then calculated the FPKM values for each gene using the fpkm function in DESeq2 and determined the median expression values for all embryonic and post-hatch samples (Supplementary Table 6). An FPKM threshold greater than 1, on the basis of the median for each group, was applied to filter out non-expressed and lowly expressed genes in both sexes. To identify developmentally expressed genes, we selected genes with FPKM greater than 1 in at least one tissue and time point (Supplementary Table 3).

Assessment of Z to proto-Z expression levels

For this analysis, RNA-seq data (log_2-transformed reads per kilobase of transcript per million reads mapped values from ref. 34) from brain, cerebellum, heart, kidney and liver from adult male and female chicken (P155), and the corresponding stage in mice (P63) was used. Akin to previous studies[5,6,9], ancestral expression levels of Z-linked genes (proto-Z genes) were estimated by calculating the median expression levels of the corresponding expressed autosomal 1:1 orthologues in an outgroup species with non-ZW sex chromosomes (in this case, mouse). In a similar way, ancestral expression levels of autosomal genes (proto-autosomal genes) were estimated by calculating the median expression levels of corresponding 1:1 orthologues that are autosomal in the same outgroup species with non-ZW sex chromosomes.

To obtain the current-Z to proto-Z expression ratios, we first normalized the current expression levels of Z-linked genes by the median current expression level of all 1:1 orthologous genes that are autosomal in the outgroup species. We then normalized the ancestral expression levels of each proto-Z-linked gene (computed as described above) by the median ancestral expression level of all proto-autosomes in the outgroup species. We then computed the ratio of these two values for each gene, resulting in the current-Z to proto-Z ratios.

Finally, we compared the current-Z to proto-Z ratios for Z-linked miR-2954 targets and Z-linked miR-2954 non-targets. As Z-linked targets, we used the experimental miR-2954 targets; as non-targets, we used Z-linked genes that are neither experimental miR-2954 targets nor predicted miR-2954 targets. In both cases, we made sure that autosomal miR-2954 targets were excluded when normalizing the expression of current-Z and proto-Z genes by current-autosomal and proto-autosomal genes. Statistically significant deviations of the medians of these ratios from key reference values (for example, 0.5 ($\log_2$ ratio of −1), 1 ($\log_2$ ratio of 0) and 2 ($\log_2$ ratio of 1)) were assessed using one-sample Wilcoxon signed-rank tests. *P* values were corrected for multiple testing using the Bonferroni procedure[68], with $P_{adj.} < 0.05$ indicating significance. Statistical equivalence to these same reference values was assessed using Wilcoxon TOST (two one-sided test) equivalence tests. This approach tests whether the medians fall within a predefined equivalence margin around each reference value, meaning the expression ratios are neither significantly above nor significantly below the specified bounds. In this analysis, the equivalence bounds were set as the reference value ± 0.5. *P* values were corrected using the Benjamini–Hochberg procedure, with $P_{adj.} < 0.05$ on both one-sided tests required for significant equivalence.

Location of genes along the Z chromosome

To visualize the location of target genes on the Z chromosome, we counted the number of protein-coding genes in windows of 0.5 Mb on the basis of gene annotations of Ensembl[57] (v.111). To indicate the location of the MHM regions, we used the regions defined by Sun et al.[47]. We lifted these regions from Galgal5.0 to the bGalGal1.mat.broiler.GRCg7b genome assembly by extracting flanking sequences from and aligning them to the new genome with BLAT.

Sequence conservation

The sequence of the miR-2954 locus was retrieved from the National Center for Biotechnology Information (NCBI) and blasted against the reference genomes of the target species (Extended Data Fig. 9 and Supplementary Data 3) using BLASTn[64].

RNA isolation, reverse transcription and RT–qPCR for miR-2954

Total RNA, including miRNA, was isolated from seven tissues (bursa of Fabricius, leg bone, brain, heart, intestine, liver and pectoral muscle) of six individual E12 chicken embryos (three male and three female; Lohmann breed) using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. Approximately 50 mg of each tissue was homogenized in 500 µl of TRIzol using a TissueLyser LT (QIAGEN) at 40 Hz for 1–2 min. RNA quality was assessed by visualizing the 28S and 18S ribosomal RNA (rRNA) bands on a denaturing agarose gel and further quantified using a NanoDrop spectrophotometer.

Reverse transcription was performed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) in accordance with the manufacturer's instructions. In each 15-µl reaction, 1,000 ng of total RNA was reverse transcribed using stem-loop reverse transcription primers specific for gga-miR-2954 (Assay ID: 243071_mat; Applied Biosystems) and U6 small nuclear RNA (snRNA) (Assay ID: 001973; Applied Biosystems). The cycling conditions were as follows: 16 °C for 30 min, 42 °C for 30 min and 85 °C for 5 min, followed by holding at 4 °C.

qPCR was performed using TaqMan MicroRNA Assay for gga-miR-2954 (Assay ID: 243071_mat) and U6 snRNA (Assay ID: 001973) as the endogenous control for normalization on a QuantStudio 6 Real-Time PCR System (Applied Biosystems) following the manufacturer's protocol. Each sample was run in triplicate. Each 10-µl reaction mixture contained 0.66 µl of complementary DNA (cDNA), 0.5 µl of TaqMan MicroRNA Assay and 5 µl of TaqMan Fast Advanced Master Mix (catalogue no. 4444557). The cycling conditions were as follows: 95 °C for 20 s, followed by 40 cycles of 95 °C for 1 s and 60 °C for 20 s.

The cycle threshold values were normalized using the $2^{\Delta CT}$ method, where ΔCT is the difference between the target gene and the endogenous control (U6 snRNA) (Supplementary Table 8).

miR-2954 knockdown and RT–qPCR

miR-2954 knockdown was achieved by injecting mirVana miRNA inhibitor specific to miR-2954 (Thermo Fisher Scientific; catalogue no. 4464088) or mirVana miRNA Inhibitor, Negative Control #1 (Thermo Fisher Scientific; catalogue no. 4464076) into chick embryos at two different embryonic stages. Lyophilized miRNA inhibitors (250 nmol; high-performance liquid chromatography; in vivo ready) were resuspended in nuclease-free water to prepare a stock solution with a final concentration of 2.5 mg ml^{-1}. The miRNA inhibitor solutions were then complexed with Invivofectamine 3.0 Reagent (Thermo Fisher Scientific; catalogue no. IVF3001). The Invivofectamine 3.0–miRNA duplex mixtures were incubated for 30 min at 50 °C and subsequently diluted with PBS (pH 7.4) according to the manufacturer's instructions.

A total of 240 fertilized eggs were obtained from Lohmann Sverige AB and placed in an incubator at 37.5 °C with 50% humidity. At E2.5, a small window was created in the eggshell above the embryo using an engraving machine. Using a fine glass needle, 2 µl of the Invivofectamine 3.0–miRNA duplex mixture (containing a final concentration of 0.63 mg ml^{-1} of the inhibitor) was injected into the dorsal aorta. Following successful injections in 170 knockdown and 28 negative control embryos, the eggs were sealed with tape and returned to the incubator at 37.5 °C with 50% humidity. A second injection was performed at E4 in surviving embryos using the same procedure but with 3 µl of the Invivofectamine 3.0–miRNA duplex mixture. Embryo viability was evaluated at E12 by observing blood flow after removal of the chorioallantoic membrane. A subset of embryos was frozen at E5 for subsequent gene expression analysis.

For gene expression analysis, 20 embryos (12 knockdown and eight control) were injected as described above and snap-frozen at E5, 1 day after the second injection, for subsequent RNA extraction and quantification (Supplementary Table 2 and Extended Data Fig. 5).

To determine the impact of miR-2954 knockdown on the expression of target and non-target genes, including *XPA*, we performed molecular sexing, dissected the heart tissue and isolated total RNA. Three controls and five knockdown embryos were used for gene expression analysis. RNA extraction was performed using TRIzol reagent, following the protocol outlined in the previous section.

Using the NCBI Primer-BLAST tool, we designed forward and reverse primers for eight target genes, eight non-target genes, the *XPA* gene and the reference gene *GAPDH* (Supplementary Table 1). Then, 1,000 ng of RNA from each sample was reverse transcribed into cDNA using the First Strand cDNA Synthesis Kit (Thermo Fisher Scientific; catalogue no. K1612) and oligo(dT) primers, according to the manufacturer's instructions. QPCR was carried out on a QuantStudio 6 Real-Time PCR System (Applied Biosystems) using SYBR Green Universal Master Mix (catalogue no. 4309155). The thermal cycling profile consisted of an initial denaturation step at 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s (denaturation) and 55 °C for 60 s (annealing and extension). Each PCR was run in triplicate. A final melting curve analysis was performed to confirm the specificity of the PCR products. Data were analysed using the delta–delta cycle threshold method. The cycle threshold values were normalized to GAPDH, and $\log_2$ fold changes between miR-2954-KD and control were generated (Supplementary Table 8).

Generation of Ribo-seq data

To compare transcriptome versus translatome patterns, we used a recently developed Ribo-seq procedure[69], on the basis of previously established methods[27,70], optimized for generating high-quality data from low-input frozen tissue samples, including small embryonic specimens. Using this method, we generated Ribo-seq and matched RNA-seq

data for a total of eight adult chicken and mouse brain (forebrain/cerebrum) samples, as well as chicken embryonic head samples (Supplementary Table 9). These data were further complemented by our previously published Ribo-seq dataset[6], which cover three additional adult chicken and mouse brain samples. Detailed protocols for the new Ribo-seq and matched RNA-seq experiments are provided below, followed by a description of the methods used to analyse these data.

Ribo-seq footprint generation

Frozen tissues were lysed in 150 µl of ice-cold lysis buffer (20 mM Tris (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 1% (v/v) Triton X-100, 1 mM dithiothreitol, 0.4 U ml^{-1} RiboLock and 100 µg ml^{-1} of cycloheximide) using a micropestle. Lysates were clarified by centrifugation at 20,000*g* for 7 min at 4 °C. For nuclease digestion, 450 U RNase I (Ambion) and 3.75 U TURBO DNase I (Thermo Fisher Scientific) were added, and samples were incubated at 25 °C for 45 min with gentle agitation. Digestion was stopped by the addition of 0.5-µl SUPERase·In RNase Inhibitor (Ambion).

To purify ribosome-protected fragments, lysates were overlaid on 700 µl of 30% sucrose cushion in 13 × 51 mm centrifuge tubes (Beckman Coulter). Samples were centrifuged at 100,000 rpm for 1 h at 4 °C using an S100-AT6 rotor (Ultracentrifuge Sorvall Discovery M120 SE). The supernatant was discarded, and the pellet was resuspended in 700 µl of 10 mM Tris (pH 7.0). To extract RNA, 40 µl of 20% sodium dodecyl sulfate and 750 µl of 65 °C acid phenol:chloroform were added, followed by incubation at 65 °C for 10 min with agitation. After centrifugation at a maximum speed for 4 min, the aqueous phase was transferred to a fresh tube containing 700 µl of acid phenol:chloroform, incubated at room temperature with intermittent vortexing and centrifuged for 4 min. Next, 600-µl chloroform was added, vortexed and centrifuged for 4 min. RNA was precipitated overnight at −70 °C in the presence of 600-µl isopropanol, 66.7 µl of 3 M sodium acetate (pH 5.5) and 2-µl GlycoBlue (Thermo Fisher Scientific). RNA was pelleted by centrifugation for 40 min at maximum speed, washed with 80% ethanol and resuspended in 12.5 µl of 10 mM Tris (pH 7.0).

The extracted RNA was separated on a 15% denaturing urea polyacrylamide gel (Thermo Fisher Scientific) and stained with SYBR Gold (Thermo Fisher Scientific). Fragments of 27–33 nt were excised and disrupted using gel breaker tubes. RNA was extracted in 0.5 ml of 10 mM Tris (pH 7.0) for 10 min at 70 °C with agitation. Gel debris was removed by centrifugation in Spin-X filter tubes (Corning) for 2 min at maximum speed. RNA was precipitated overnight at −70 °C in the presence of 1 volume isopropanol, 0.1 volume 3 M sodium acetate (pH 5.5) and 2-µl GlycoBlue (Thermo Fisher Scientific). RNA was pelleted by centrifugation for 40 min at maximum speed and washed with 80% ethanol.

Ribo-seq library preparation and sequencing

Ribo-seq library preparation was performed as described in ref. 69 with several modifications. In brief, ribosome footprints were dephosphorylated and ligated to a pre-adenylated 3′ linker (L1), followed by enzymatic removal of unligated linkers. Footprint–linker complexes were captured on streptavidin beads, phosphorylated and ligated to a 5′ linker (L2). Reverse transcription was performed on bead-bound templates, and the resulting cDNA libraries were amplified by PCR. To improve depletion of unligated L1, we modified the digestion step by incubating samples sequentially at 30 °C for 60 min and 37 °C for 60 min with deadenylase and RecJf. Libraries were PCR-amplified using eight cycles of amplification. A modified version of the previously published Cas9-mediated Ribocutter tool[71] was used to deplete rRNA from the Ribo-seq libraries. The sgRNAs were designed to target the most abundant contaminants of previously sequenced libraries derived from chicken or mouse telencephalon. To enhance the efficiency of rRNA removal, a lower library concentration (6 nM) was used as input for Cas9-mediated depletion and extended the Cas9 treatment to 4.5 h. An additional 6% polyacrylamide gel electrophoresis Tris–borate–EDTA gel step was introduced to remove preferentially amplified adaptor dimers following a seven-cycle PCR reamplification. All further steps were performed according to the original protocol. The libraries were resuspended and quality controlled using Qubit (Thermo Fisher Scientific) and Fragment Analyzer (Agilent Technologies) platforms. Sequencing was performed on both the Illumina NextSeq 2000 and Illumina NextSeq 550 systems, using NextSeq 2000 P4 Reagents (50 cycles) and NextSeq 550 High Output Reagents (75 cycles). The samples were multiplexed into one set of six and another set of one. All further steps were performed in accordance with the original protocol.

RNA library preparation and sequencing

To generate matched RNA-seq libraries prepared from the same lysates, total RNA was extracted from dissected tissues using the RNeasy Micro Kit (QIAGEN), following the manufacturer's instructions. The RNA quality was assessed using the Fragment Analyzer system (Agilent), and RNA quality numbers ranged from 7.7 to 10, indicating minimal degradation. The RNA-seq libraries were prepared using the SMART-Seq Total RNA High Input kit with (Mammalian) RiboGone (Takara Bio). The concentration and quality of the libraries were determined using Qubit (Thermo Fisher Scientific) and Bioanalyzer (Agilent Technologies) platforms. Illumina sequencing was performed on an Illumina NextSeq 2000 system using NextSeq 2000 P2 Reagents (100 cycles), with samples multiplexed in one set of six.

Read mapping and processing

Raw sequencing reads with Illumina 3′ adaptor and low-quality bases (Phred score below 20) were trimmed using cutadapt v.4.6 (parameters: --adapter=AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC --minimum-length=6 -q 20). For Ribo-seq libraries, unique molecular identifiers (UMIs) were extracted using UMI-tools v1.1.4 (ref. 72) (parameters: --bc-pattern = ^(?P<umi_1>.{5}).+(?P<umi_2>.{5})$ --extract-method=regex OR --bc-pattern = ^(?P<umi_1>.{10}).+(?P<umi_2>.{10})$ --extract-method=regex), and leading nucleotides were removed with cutadapt v.4.6 (parameters: -u 6). RNA-seq libraries did not contain UMIs, and leading nucleotides were removed with cutadapt v.4.6 (parameters: -u 3). Trimmed reads were consecutively mapped to the index libraries of species-specific (chicken or mouse) contaminating RNAs obtained from RNAcentral[73] (rRNAs, mitochondrial RNAs and transfer RNAs) using Bowtie 2 v.2.5.1 (ref. 74) (parameters: --phred33 -L 20 -N 1 -t --no-unal). Aligned reads were discarded, and only those within the defined length ranges (26–34 nt for Ribo-seq and 20–50 nt for RNA-seq) were kept for downstream analysis. As expected[27], Ribo-seq read lengths peaked at 28–30 nt and predominately mapped to coding DNA sequences (CDSs) (Extended Data Fig. 4a,b). To mitigate bias in the mapping of RNA-seq reads in exon–exon junctions owing to length discrepancies between both methods, RNA-seq reads were cut to 29 nt. Reads were then aligned to the reference genomes (bGalGal1.mat.broiler.GRCg7b; GCA_016699485.1 OR GRCm39; GCA_000001635.9, Ensembl release 113; ref. 75) using STAR aligner v.2.7.11a (ref. 58) (parameters: --alignEndsType EndToEnd --outSAMattributes All --outSAMtype BAM SortedByCoordinate --outMultimapperOrder Random). As previously described[27], peptidyl-site offsets were estimated per read length, and Ribo-seq reads were calibrated accordingly.

Triplet periodicity

To assess whether our Ribo-seq libraries showed patterns of true translation, we analysed the triplet periodicity using raw reads mapped to the complete CDS regions of protein-coding genes. To ensure robust analysis, we focused on protein-coding genes annotated as canonical in Ensembl (release 113). The number of reads mapped to the three reading frames was normalized by the total number of reads within the CDS. As shown in Extended Data Fig. 4c,d, in contrast to RNA-seq reads, our Ribo-seq data predominantly mapped to the first nucleotide

of the codon showing continuous and significant triplet periodicity across the CDS.

Estimation of gene expression levels

Transcript abundances were estimated in FPKM. Only uniquely mapped RNA-seq reads and de-duplicated uniquely mapped Ribo-seq reads within the CDS regions were considered. On the basis of our triplet periodicity analysis, we further restricted the analysis to read lengths that exhibited significant triplet periodicity. Moreover, only the CDS region from the +4th to the −3rd codon was used to avoid inflated counts owing to random translation initiation and ribosome enrichment at the stop codon[70]. For each gene, the longest isoform was used as a representative. Gene count matrices were then loaded into R v.4.4.0, and gene expression levels were estimated using the rpkm function of edgeR v.4.2.0, which accounts for both CDS length and library depth. The FPKM values were $\log_2$-transformed.

Assessment of reproducibility

To assess the reproducibility of our Ribo-seq data, we calculated the Spearman's correlation coefficient (ρ) between the read counts of canonical protein-coding genes in two biological chicken female brain replicates. The high correlation ($\rho = 0.98$) demonstrates strong biological reproducibility (Extended Data Fig. 4e).

Comparison of $Z^{KO}Z^{KO}$ and ZZ genotypes

Processed FPKM values were used to calculate $\log_2$[FC] in gene expression between $Z^{KO}Z^{KO}$ and ZZ genotypes for both layers (transcriptome and translatome) in head tissue. Genes with FPKM values greater than 1 in both genotypes were kept to exclude non-expressed or lowly expressed genes. To enable direct comparisons between layers, the FPKM values were normalized using the median expression of autosomal non-target transcriptome or translatome genes, respectively.

Assessment of Z to proto-Z translation levels

To estimate the ancestral translatome levels of Z-linked genes, we combined our newly generated Ribo-seq data with our previously published dataset[6]. The analysis followed the same approach as the RNA-seq analysis described earlier (Assessment of Z to proto-Z expression levels).

Female-to-male expression-level ratios

The processed FPKM values were used to calculate female-to-male ratios for two tissues (fetal head and adult brain) for both layers (transcriptome and translatome). Genes with FPKM values greater than 1 were kept to filter out non-expressed and lowly expressed genes. To allow for comparisons between layers, the FPKM values were normalized using either the median of autosomal transcriptome or translatome expression. The ratios were then compared to the key reference values, as described earlier (Assessment of Z to proto-Z expression levels).

Translation efficiency estimation

The $\log_2$-transformed FPKM values at the translatome (ribosome-protected fragment) and transcriptome (RNA) were used to calculate translation efficiency across samples as:

$$TE = \log_2(RPF_{FPKM}) - \log_2(RNA_{FPKM})$$

where RPF is the ribosome-protected fragment, and TE is the translation efficiency. Further, to highlight the differences between male and female translation efficiencies, the female-to-male-translation-efficiency ratios were calculated as:

$$TE_{F\text{-}to\text{-}M} = \log_2(TE_{female}) - \log_2(TE_{male})$$

Finally, the ratios were normalized using the median of autosomal female-to-male-translation-efficiency ratios.

Long-read genome sequencing of miR-2954 KOs and controls

For long-read sequencing, we selected five miR-2954 KO individuals and four non-edited controls. Genomic DNA was isolated through DNeasy Blood & Tissue Kit (QIAGEN). Library preparation was performed using the Rapid Barcoding Kit (SQK-RBK114-24) or the Native Barcoding Kit (SQK-NBD114-24) (Oxford Nanopore Technologies). Sequencing was conducted on PromethION R10.4.1 flow cells with adaptive sampling[76] to specifically enrich for Z-chromosomal reads.

For basecalling, we used the high-accuracy model (dna_r10.4.1_e8.2_400bps_sup@v5.0.0) implemented in dorado-0.9.0. Reads were then aligned to the GRCg7b chicken genome assembly using minimap2 (ref. 77) (v.2.27-r1193) with the long-read high-quality preset (-x lr:hq). The per-sample coverage depth across chromosome Z was calculated using SAMtools 1.20 (ref. 78) depth command.

We next searched for structural variants using Sniffles2 (v.2.2)[79], configuring the tool to call small indels and putative structural variants on chromosome Z with the parameters --minsvlen 5 and --minsupport 0. This analysis identified six candidate variants shared among the five KO samples but absent in controls (Supplementary Table 10). Manual inspection of these variants in Integrative Genomics Viewer revealed that only one site showed a consistent difference between KO and control samples: the intended Cas9 target, which produced a 32-bp deletion starting at chr. Z: 71,305,198. Notably, this deletion was also the only variant displaying zero coverage exclusively in KO individuals, confirming the successful miR-2954 KO.

Reporting summary

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

Data availability

Raw sequencing data (mRNA-seq, small RNA-seq, DNA-seq and Ribo-seq) were deposited in the NCBI Sequence Read Archive under BioProject accession no. PRJNA1079296. Processed data for mRNA-seq and small RNA-seq are available in the Gene Expression Omnibus repository under the same BioProject accession. Dosage sensitivity scores are available at Zenodo (https://doi.org/10.5281/zenodo.6347672)[80]. The list of ohnologues is available from the Ohnologs database (http://ohnologs.curie.fr). Gene phylogenetic ages are available from GenTree (http://gentree.ioz.ac.cn/). Genome assemblies, 3′ UTR sequences and 1:1 orthologues were retrieved from Ensembl releases 109–113 (https://www.ensembl.org/info/website/archives/index.html) and BioMart (https://www.ensembl.org/info/data/biomart/index.html). The following genome assemblies were used: chicken (*G. gallus*) bGalGal1.mat.broiler.GRCg7b, GRCg6a and galGal4; zebra finch (*T. guttata*) bTaeGut1_v1.p; crocodile (*Crocodylus porosus*) CroPor_comp1; ostrich (*Struthio camelus*) ASM69896v1; human (*Homo sapiens*) GRCh38.p14; and mouse (*M. musculus*) GRCm39.

Code availability

Custom scripts used to generate the results reported in the paper and processed data are available at GitHub (https://github.com/amirshahr/MIR2954).

54. He, L. et al. Simple, sensitive and robust chicken specific sexing assays, compliant with large scale analysis. *PLoS ONE* **14**, e0213033 (2019).
55. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272–W276 (2016).
56. Ye, J. et al. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinf.* **13**, 134 (2012).
57. Martin, F. J. et al. Ensembl 2023. *Nucleic Acids Res.* **51**, D933–D941 (2023).
58. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

59. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
60. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
61. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
62. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Methodol.* **57**, 289–300 (1995).
63. Singh, P. P. & Isambert, H. OHNOLOGS v2: a comprehensive resource for the genes retained from whole genome duplication in vertebrates. *Nucleic Acids Res.* **48**, D724–D730 (2020).
64. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
65. Yanai, I. et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650–659 (2005).
66. Ayers, K. L. et al. RNA sequencing reveals sexually dimorphic gene expression before gonadal differentiation in chicken and allows comprehensive annotation of the W-chromosome. *Genome Biol.* **14**, R26 (2013).
67. Ayers, K. L. et al. Identification of candidate gonadal sex differentiation genes in the chicken embryo using RNA-seq. *BMC Genomics* **16**, 704 (2015).
68. Bonferroni, C. E. *Teoria Statistica Delle Classi e Calcolo Delle Probabilità* (Seeber, 1936).
69. Koubek, J. et al. A simple, fast and cost-efficient protocol for ultra-sensitive ribosome profiling. Preprint at *bioRxiv* https://doi.org/10.1101/2025.04.09.647970 (2025).
70. McGlincy, N. J. & Ingolia, N. T. Transcriptome-wide measurement of translation by ribosome profiling. *Methods* **126**, 112–129 (2017).
71. Wilkins, O. G. & Ule, J. Ribocutter: Cas9-mediated rRNA depletion from multiplexed Ribo-seq libraries. Preprint at *bioRxiv* https://doi.org/10.1101/2021.07.14.451473 (2021).
72. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
73. Consortium, R. N. RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Res.* **49**, D212–D220 (2021).
74. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
75. Dyer, S. C. et al. Ensembl 2025. *Nucleic Acids Res.* **53**, D948–D957 (2025).
76. Loose, M., Malla, S. & Stout, M. Real-time selective sequencing using nanopore technology. *Nat. Methods* **13**, 751–754 (2016).
77. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
78. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
79. Smolka, M. et al. Detection of mosaic and population-level structural variants with Sniffles2. *Nat. Biotechnol.* **42**, 1571–1580 (2024).
80. Collins, R. et al. A cross-disorder dosage sensitivity map of the human genome. Zenodo https://doi.org/10.5281/zenodo.6347672 (2022).

Acknowlegements We thank all members of the Kaessmann group for discussions; D. Ibberson and B. Berki from the Deep Sequencing Core Facility of Heidelberg University for sequencing support; D. Meunier, K. Hogan, F. Thomson, N. Russell and the Roslin poultry team in the National Avian Research Facility for supporting the chicken in vivo work; M. Clinton and L. McTeir for useful discussions and support; A. Wallberg for long-read sequencing analyses; and T. Punga for supporting the knockdown experiments. This study was supported by grants from the European Research Council (grant agreement nos. 101019268 and 615253) and the NOMIS Foundation (no. 2024-A-0035) to H.K., the Swedish Research Council (VR) (2017-06218) and the Swedish Research Council for Sustainable Development (Formas) (2021-00513) to A.F., Formas (2023-01396) to S.Y.T. and by an add-on fellowship of the Joachim Herz Stiftung to L.R.-M. M.C.-M. was supported by The Francis Crick Institute, which receives its core funding from Cancer Research UK (grant CC2185 to M.C.-M.), the UK Medical Research Council (grant CC2185 to M.C.-M.) and the Wellcome Trust (grant CC2185 to M.C.-M.). This study was also supported by the Institute Strategic Grant Funding from the BBSRC (BBS/E/RL/230001C and BBS/E/RL/230002A) to the Roslin Institute. The National Avian Research Facility (NARF) was partially funded through the BBSRC Core Capabilities Grant BB/CCG2270/1. J.K. and G.K. were supported by the German Israeli Foundation for Scientific Research and Development (I-1564-417.13/2023 awarded to G.K.). We thank Lohmann Sverige for providing fertile eggs used in this study.

Author contributions A.F., M.J.M. and H.K. conceived and organized the study. A.F. and H.K. wrote the paper with input from all authors. A.F. and S.Y.T. performed most of the experiments and bioinformatics analyses. L.R.-M. performed the evolutionary and spatiotemporal expression analyses. E.S. performed the embryo survival analysis. M.J., M.B., A.I.-A., L.T., A.S., C.G., Y.X., M.D. and L.L. contributed to experimental procedures. C.-J.R. performed the data generation and analysis of long-read sequencing. C.M. performed the evolutionary conservation analyses of miR-2954. D.F., J.K. and G.K. developed the Ribo-seq protocol. D.F. generated and P.L. analysed the Ribo-seq data. N.T., E.S., M.J. and M.C.-M. provided key feedback and discussions.

Competing interests The authors declare no competing interests.

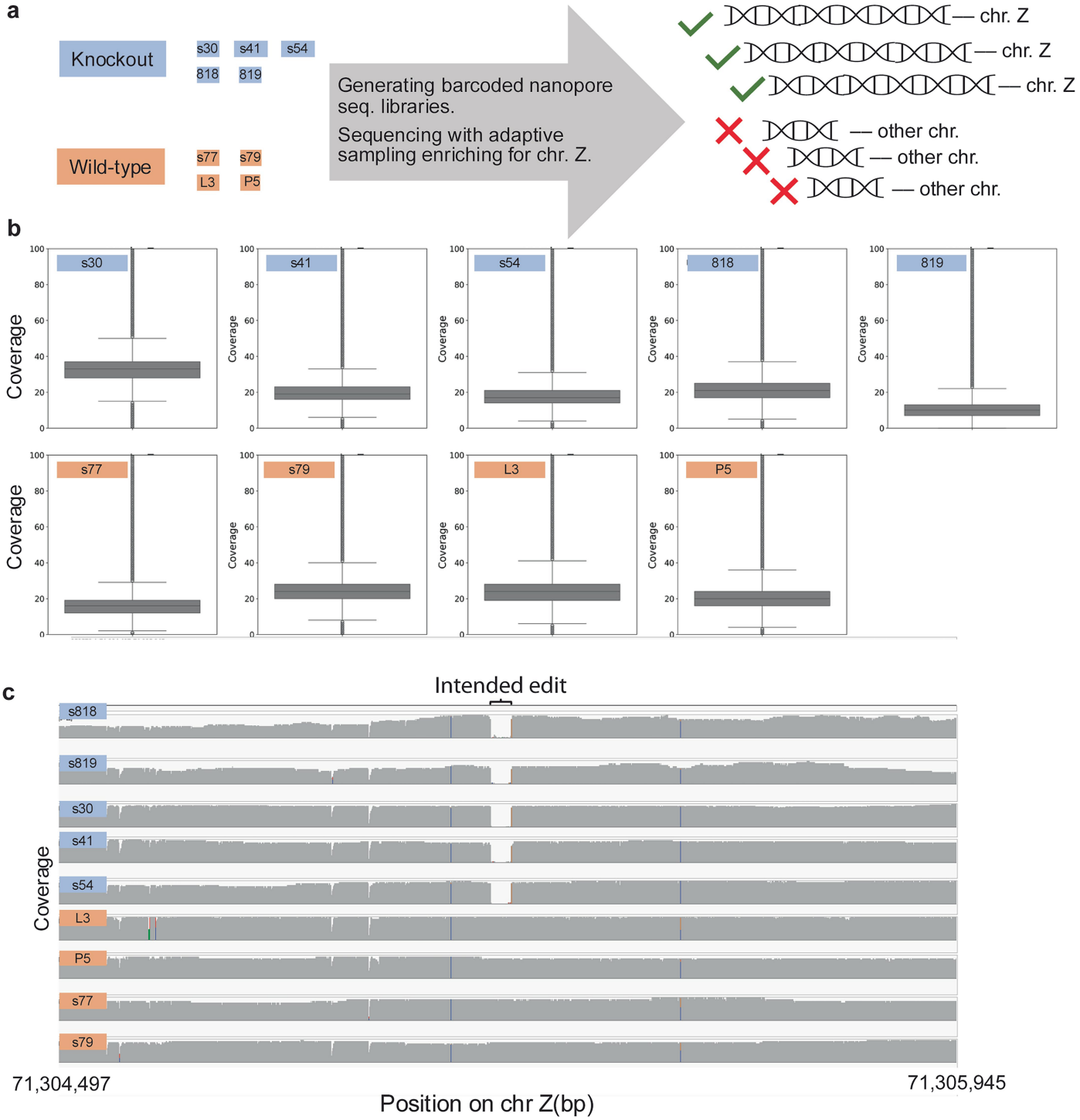
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-09256-9.

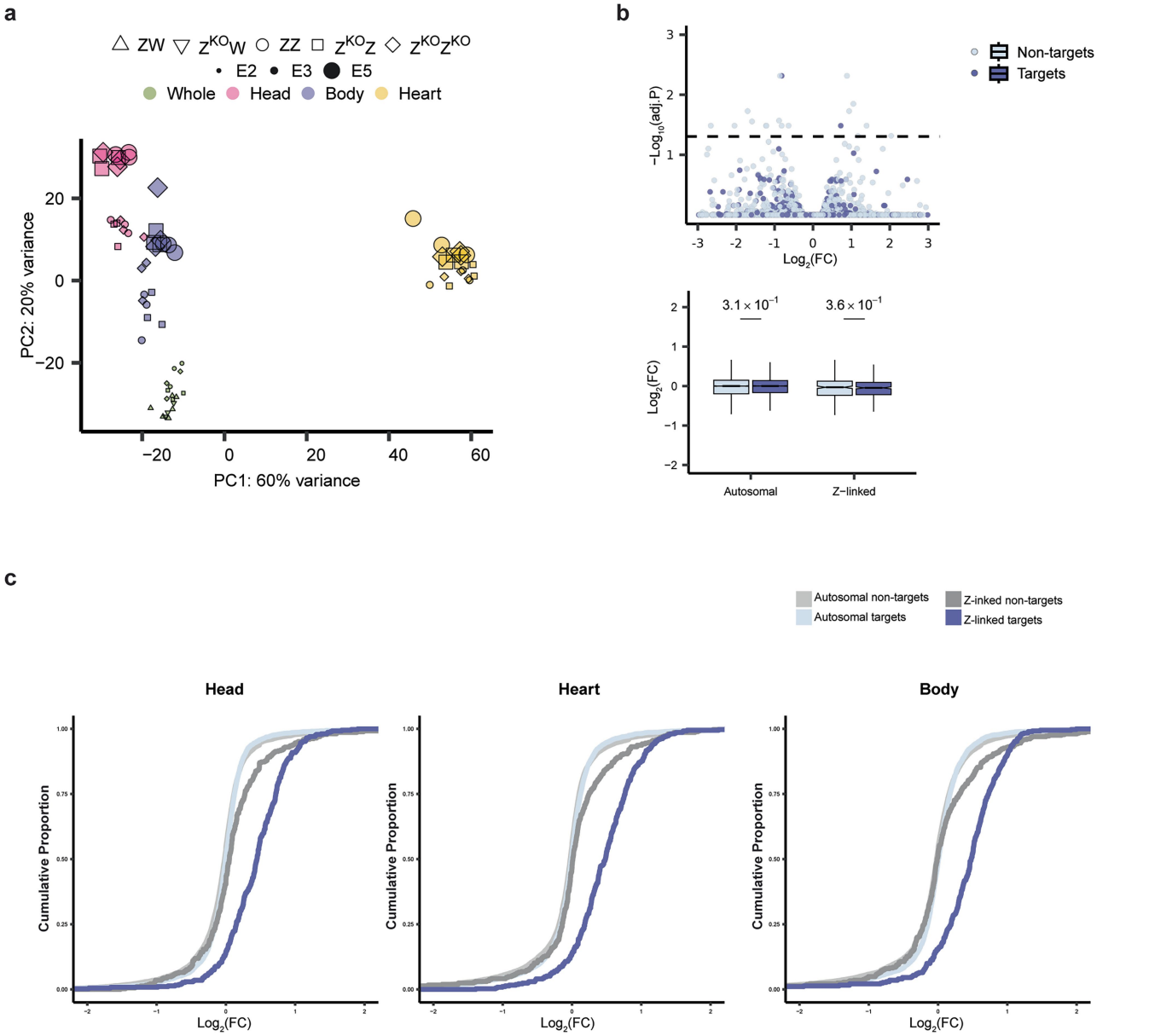
Correspondence and requests for materials should be addressed to Amir Fallahshahroudi, Mike J. McGrew or Henrik Kaessmann.

Peer review information *Nature* thanks Hubert Schwabl, Craig Smith and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

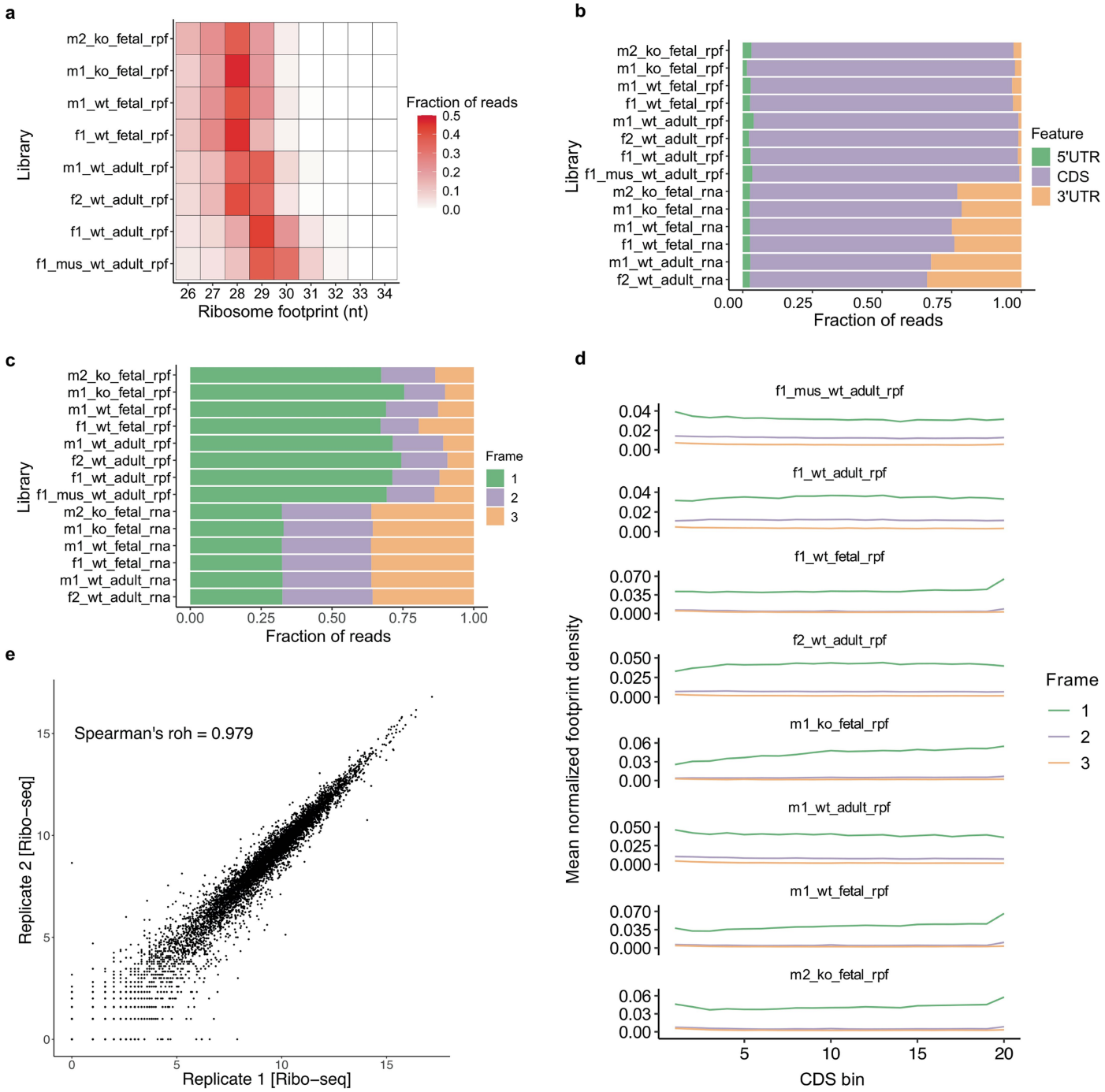
Extended Data Fig. 1 | Overview of the knockout of the miR-2954 locus. Top: overview of the *XPA* host gene showing the the miR-2954 locus, located in the second or third (depending on the isoform) intron of this gene. Bottom and middle panels: alignments and overview of the genomic reference sequence around miR-2954 locus, the induced deletion, and the single-stranded DNA oligonucleotide (ssODN) repair template used to leverage the homology-directed repair (HDR) pathway (Methods). A sequence track highlights the positions of the pre- and mature miR-2954 sequences. The ssODN repair template aligns with the post-editing PGC clone sequence, as confirmed by Sanger sequencing (clone #36, used for generating KO chickens), which includes a 36 bp deletion adjacent to an EcoRI restriction site. The accompanying chromatogram verifies the deletion, illustrating the consistency between the edited and expected sequences. We note that miR-1583, also shown in the figure, is not listed in miRGeneDB and is therefore not considered a confidently annotated microRNA.



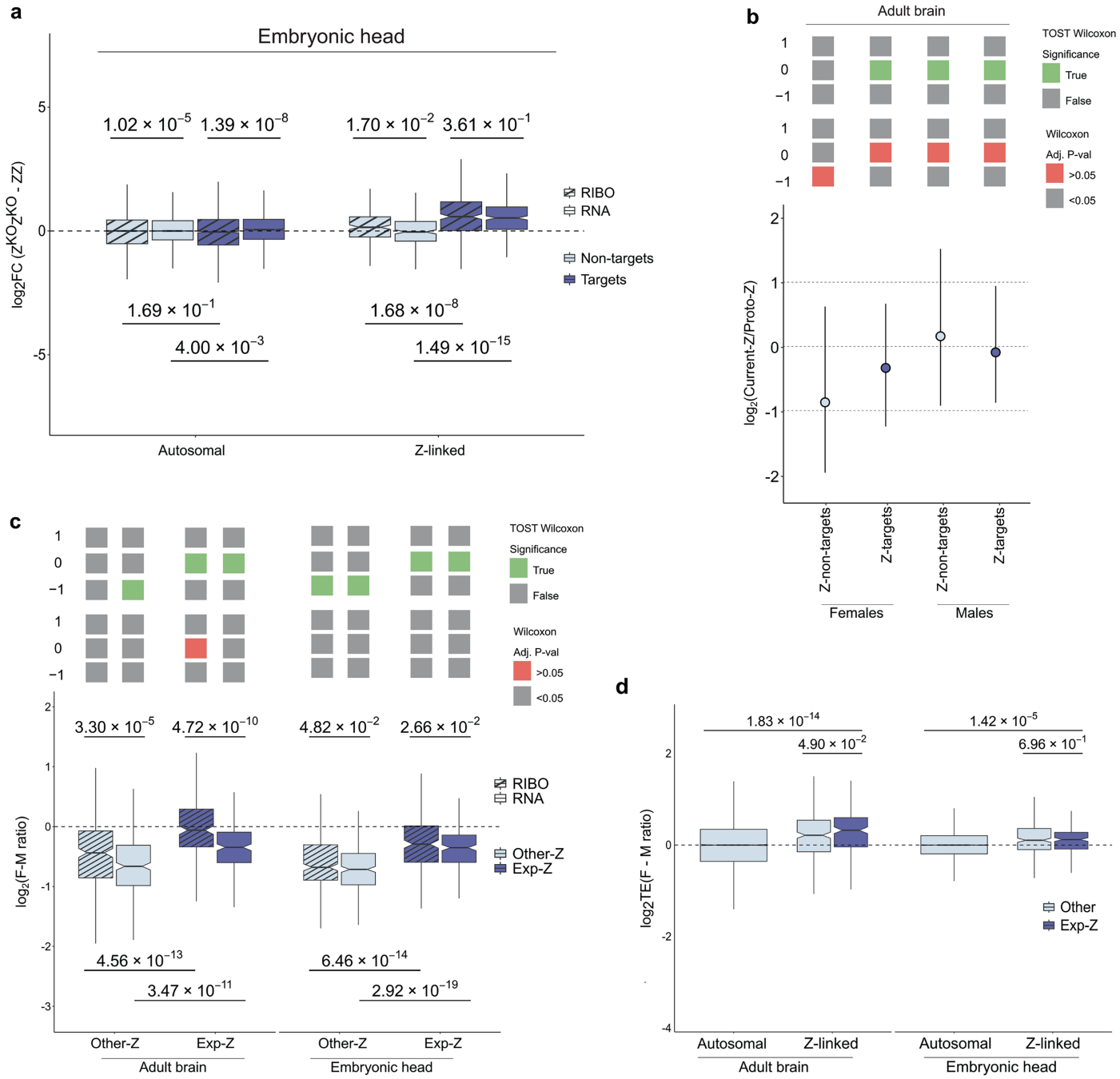
Extended Data Fig. 2 | Targeted long-read Oxford Nanopore sequencing of chromosome Z in Knock-out (KO) and control individuals. a, DNA from five miR-2954 KO individuals and four control individuals was sequenced using adaptive sampling to enrich coverage for chromosome Z. **b**, Distributions of depth of coverage on chromosome Z among the nine sequenced samples. Box plots show the median, 25th–75th percentiles, and whiskers extending to 1.5× the IQR. **c**, All KO samples carried a 32 bp homozygous deletion over the targeted region, while the control samples had normal coverage in this region. No other genetic variants consistently distinguishing KO from control samples were detected on chromosome Z.



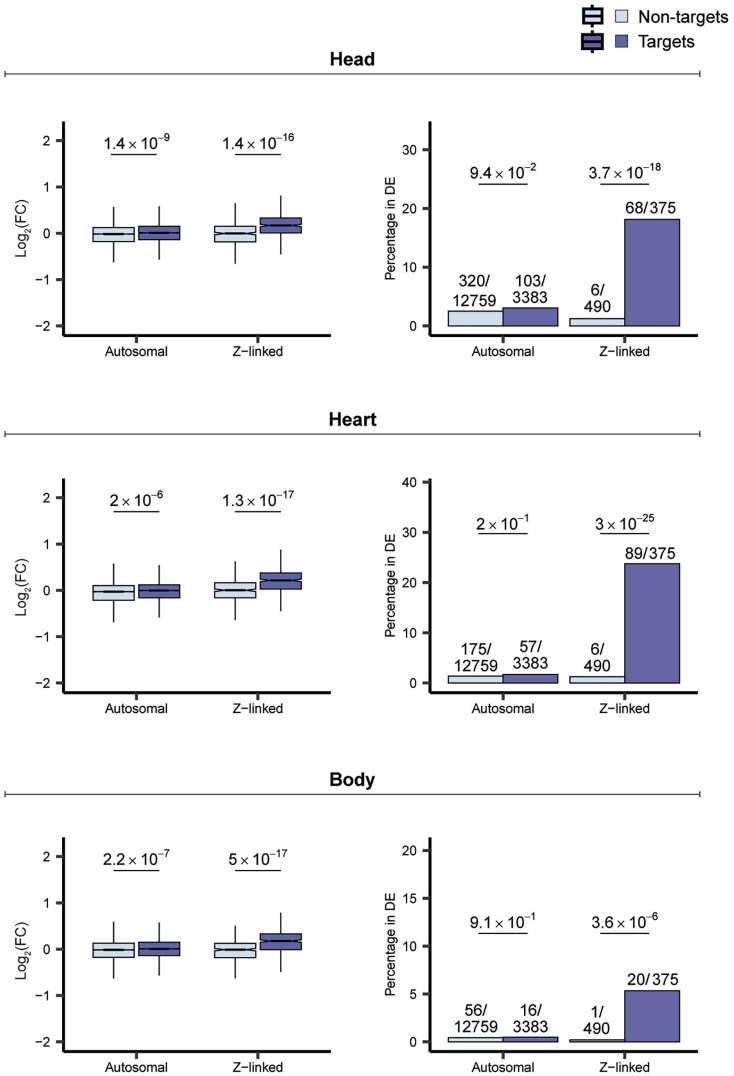
Extended Data Fig. 3 | Quality control of RNA sequencing data. a, Principal component analysis (PCA) of mRNA expression profiles across samples. The percentage of variance explained by the first two principal components (PC1 and PC2) is shown. **b**, Comparison of E2 whole embryos from "pure" Hy-Line (HL) females (original stock, $n = 3$) and wild-type ZW females from the G2 generation ($n = 3$). Top: volcano plot showing $\log_2$-fold changes ($\log_2$FC) and $-\log_{10}$ of Benjamini-Hochberg-adjusted *P*-values for predicted miR-2954 target and non-target protein-coding genes. Bottom: Log_2FC values in gene expression between pure Hy-Line and G2 ZW females for autosomal (n = 16,142) and Z-linked (n = 865) protein-coding genes; *P*-values from two-sided Wilcoxon rank-sum tests are shown above. Box plots show median, 25th–75th percentiles, and whiskers extending to 1.5× the IQR. **c**, Cumulative distribution of $\log_2$FC values in gene expression between $Z^{KO}Z^{KO}$ and ZZ genotypes for autosomal and Z-linked protein-coding genes across head, body, and heart tissues.



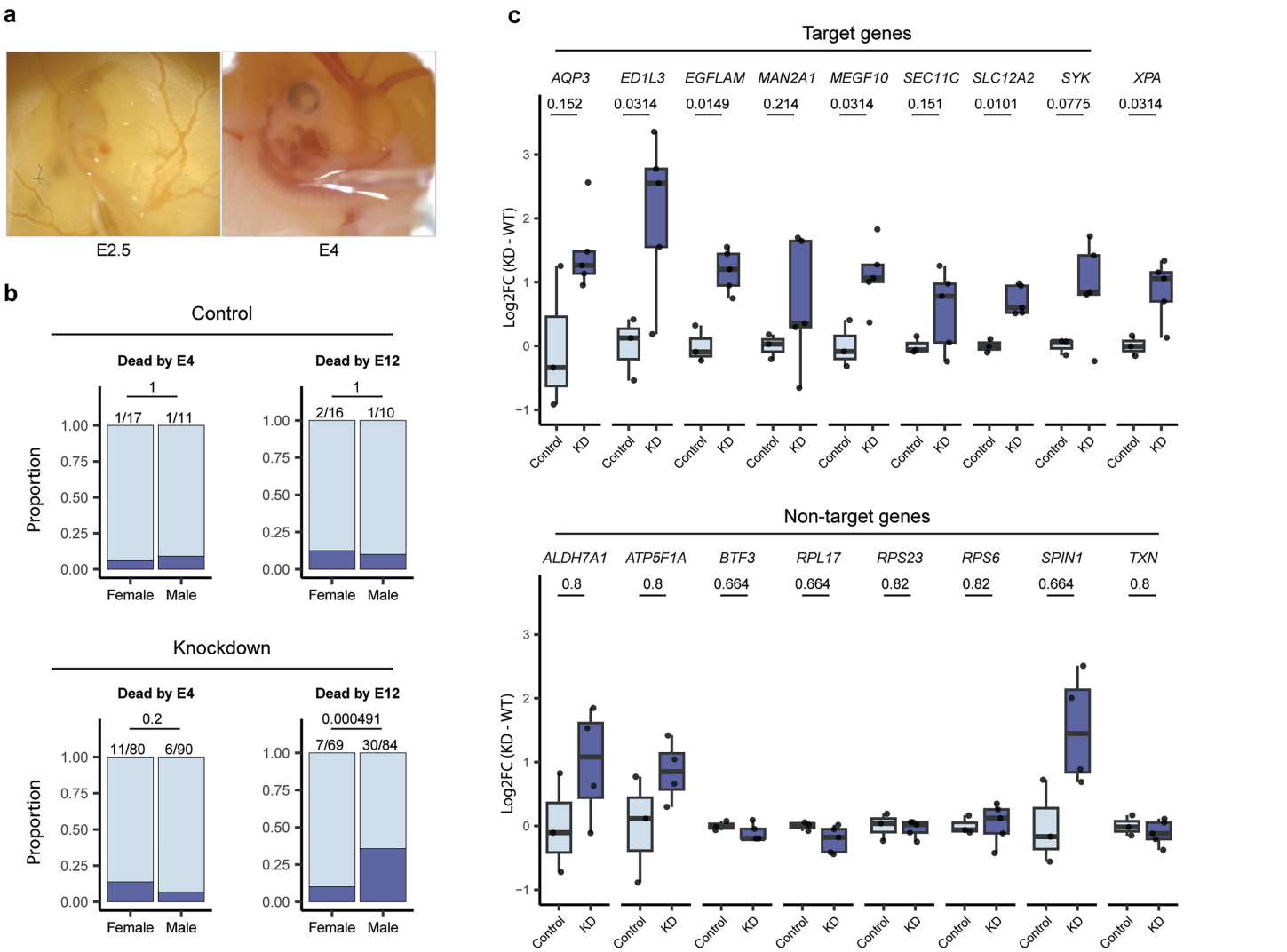
Extended Data Fig. 4 | Quality control of Ribo-seq libraries. a, Distribution of ribosome footprint length across Ribo-seq libraries (nt, nucleotides). **b**, Ribo-seq and RNA-seq read fractions mapped to 5′-untranslated regions (5′-UTRs), coding sequences (CDSs) and 3′ untranslated regions (3′-UTRs). **c**, Distribution of Ribo-seq and RNA-seq reads across reading frames in the CDS of canonical protein-coding genes. **d**, Mean normalized footprint density along the CDS of canonical protein-coding genes for the Ribo-seq data. For each library, only read lengths with strong triplet periodicity were included. Each CDS was divided into 20 equal-length bins, with each bin representing the proportion of reads that mapped to that segment. **e**, Distribution of read counts (values > 1) for canonical protein-coding genes between two biological replicates of chicken brain Ribo-seq libraries, with the Spearman's correlation coefficient (ρ) indicated.



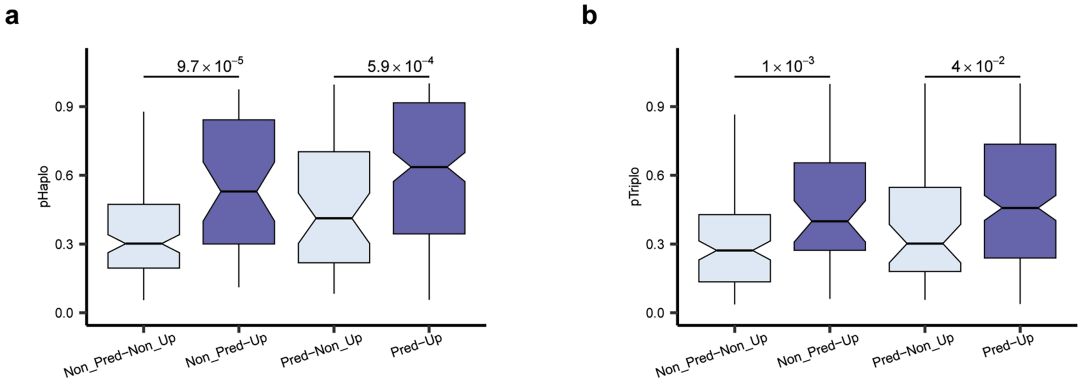
Extended Data Fig. 5 | Transcriptome and translatome comparison across different tissues and genotypes. a, Log_2-fold changes (Log_2FC) between $Z^{KO}Z^{KO}$ and ZZ genotypes for both autosomal (Non-targets: $n = 7279$ (RIBO), $n = 7566$ (RNA); Targets: $n = 2238$ (RIBO), $n = 2309$ (RNA)) and Z-linked protein-coding genes (Non-targets: $n = 197$ (RIBO), $n = 220$ (RNA); Targets: $n = 264$ (RIBO), $n = 292$ (RNA)). Two-sided Wilcoxon test *P*-values between data types or target groups are indicated above and below. Log_2FC estimates are either based on transcriptome (RNA) or translatome data (RIBO, indicated by dashed lines) of E3 embryo heads. Genes with FPKM > 1 were used and values were normalized using either the median of autosomal non-target transcriptome or translatome expression. **b**, Median and interquartile ranges of ratios of current versus proto-Z (ancestral) translation for Z-linked non-targets ($n = 174$) and Exp-Z ($n = 188$) genes for males and females (log_2). Reference lines indicate ratios of −1 (half ancestral expression), 0 (equal expression), and 1 (twofold ancestral expression). Statistical significance is indicated above as TOST (Two One-Sided Tests) Wilcoxon equivalence test (green indicates significance within equivalence bounds (reference value ± 0.5, Methods), gray indicates non-significance) and two-sided one-sample Wilcoxon test (red equals significant deviation from reference value, gray equals non-significant deviation). **c**, Female-to-male expression level ratios (log_2) for Exp-Z and Other-Z genes in adult brain (Exp-Z: $n = 203$ (RIBO), $n = 215$ (RNA); Other-Z: $n = 270$ (RIBO), $n = 297$ (RNA)) and embryonic head (Exp-Z: $n = 201$ (RIBO), $n = 212$ (RNA); Other-Z: $n = 257$ (RIBO), $n = 282$ (RNA)). Two-sided Wilcoxon test *P*-values between data types or between target groups are indicated above and below. Ratios were calculated using either transcriptome (RNA) or translatome data (RIBO, indicated by dashed lines). Genes with FPKM > 1 were used and values were normalized using either the median of autosomal transcriptome or translatome expression. Statistical significance was assessed as in b). **d**, Female to male translational efficiency (TE) ratios (log_2) for Exp-Z, Other-Z and Autosomal genes in adult brain (Exp-Z: n = 194; Other-Z: n = 259; Autosomal: n = 9843) and embryonic head (Exp-Z: n = 199; Other-Z: n = 246; Autosomal: n = 9276). Two-sided Wilcoxon test *P*-values between target groups are indicated above. Genes with FPKM > 1 were used and values were normalized using the median of autosomal TE ratios. All box plots show the median, 25th–75th percentiles, and whiskers extending to 1.5× the IQR.



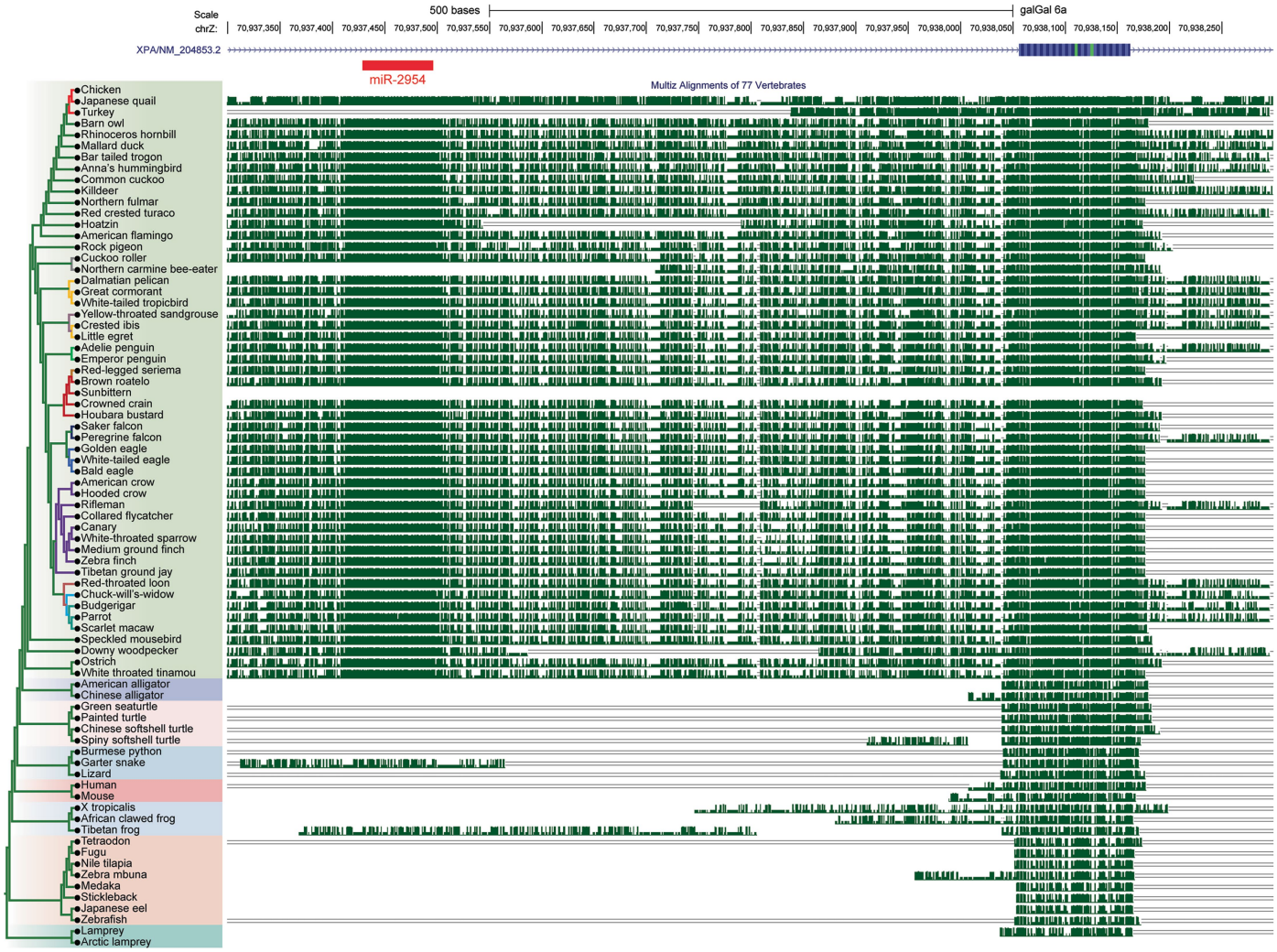
Extended Data Fig. 6 | Impact of miR-2954 knockout on gene expression in heterozygous KO males. Left column: $\log_2$-fold changes (Log_2FC) in gene expression between $Z^{KO}Z$ and ZZ genotypes for autosomal ($n = 16{,}142$) and Z-linked ($n = 865$) protein-coding genes; *P*-values from a two-sided Wilcoxon rank-sum test are indicated above. Log_2FC estimates are based on transcriptomes of E3 and E5 embryos across head, heart, and rest of the body tissues ($n = 3$ biological replicates per tissue, genotype, and developmental stage). Box plots show the median, 25th–75th percentiles, and whiskers extending to 1.5× the IQR. Right column: proportions of autosomal and Z-linked target and non-target genes among the differentially expressed (DE) genes (Benjamini-Hochberg adjusted $P < 0.05$) when comparing $Z^{KO}Z$ and ZZ genotypes. Two-sided χ^2 test *P*-values are shown above.



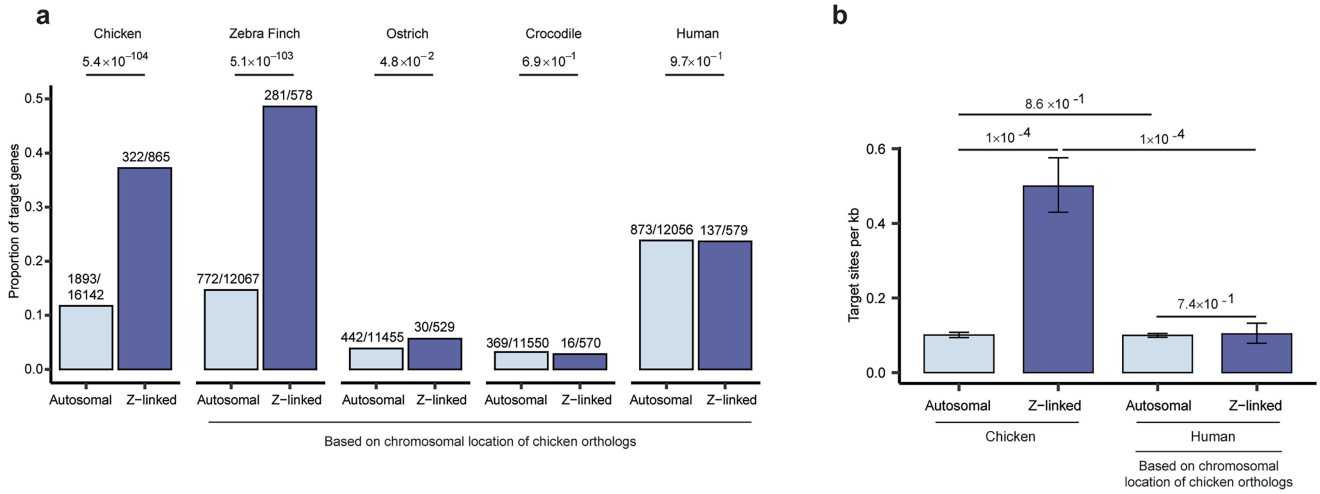
Extended Data Fig. 7 | Effects of miR-2954 knockdown on survival and gene expression. a, Sequential injections of the mirVana inhibitor complexed with Invivofectamine 3.0 Reagent into embryos at E2.5 (left) and E4 (right) using a microcapillary glass needle. **b**, Survival proportions of miR-2954 knockdown (KD) embryos categorized by treatment, sex, and embryonic day (E) of development. Numbers above the bars indicate the numbers of dead vs. total number of embryos analyzed per subgroup at each timepoint; *P*-values are from two-sided χ^2 tests. E12 survival represents the subset of embryos that survived at E4 and received a second injection. **c**, Log_2-fold change (KD vs. Control) of miR-2954 target genes and the *XPA* gene (top), and non-target genes (bottom) in heart tissue of male chicken embryos at E5 for control ($n = 3$) and miR-2954 KD ($n = 5$), measured by RT-qPCR. *P*-values are from two-sided t-tests with Benjamini-Hochberg correction. All box plots show the median, 25th–75th percentiles, and whiskers extending to 1.5× the IQR. Individual data points are overlaid with jitter.



Extended Data Fig. 8 | Dosage sensitivity assessments in the extended set of 576 genes expressed during chicken development. a, Probabilities of haploinsufficiency (pHaplo). **b**, Probabilities of triplosensitivity (pTriplo). In both panels, comparisons are shown between upregulated (Up) and non-upregulated (Non_Up), developmentally expressed Z-linked (n = 576) genes for both predicted (Pred) and non-predicted (Non_Pred) targets of miR-2954. Two-sided Wilcoxon test *P-values* are shown above. All box plots show the median, 25th–75th percentiles, and whiskers extending to 1.5× the IQR.



Extended Data Fig. 9 | Alignment of the miR-2954 locus and segment of the *XPA* host gene sequence across vertebrates and conservation pattern. MultiZ alignment of the genomic region containing miR-2954 across 77 vertebrate species, highlighting sequence conservation across species and phylogenetic relationships. On the left, a phylogenetic tree shows the evolutionary connections among species. The red bar marks the location of miR-2954. The main alignment consists of green bars, where each row represents a species' sequence; green regions indicate conserved nucleotides, while white gaps signify divergence or deletions.



Extended Data Fig. 10 | Genes with miR-2954 target sites in birds and outgroup species. a, Proportion of autosomal and Z-linked genes with predicted miR-2954 target sites (7–8mer matches) among all protein-coding genes in chicken, and among 1:1 chicken orthologs in zebra finch, ostrich, crocodile, and human. Distributions of predicted targets and non-targets were compared using two-sided χ^2 tests. The number of predicted targets and the total number of genes in each category are shown above the bars. **b**, Mean number of miR-2954 target sites per kilobase of 3′ UTR in chicken and human. A total of 11,789 autosomal and 571 Z-linked genes with annotated 3′ UTRs in both chicken and human were used to calculate the number of target sites per kilobase of 3′ UTR. Error bars indicate 95% confidence intervals from permutation tests (n = 10,000), with empirical *P-values* from two-sided pairwise permutation tests shown above the bars. In both panels, genes were classified as autosomal or Z-linked based on the chromosomal location of their 1:1 chicken orthologs.

nature portfolio

Corresponding author(s): Amir Fallahshahroudi, Mike McGrew, Henrik Kaessmann

Last updated by author(s): May 25, 2025

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

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	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	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)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	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	No software was used.
Data analysis	CHOPCHOP v2 and Primer-BLAST were used for designing quide RNAs and primers. Analyses were performed using STAR aligner (v2.7.2b and v2.7.11a), DESeq2 (v1.24.0), R (v4.1 and v4.4.0), vsn (v3.52.0), Cutadapt (v4.4 and v4.6), TargetScan (v7.0 and v6.0), UMI-tools (v1.1.4), Bowtie2 (v2.5.1), minimap2 (v2.27-r1193), Sniffles2 (v2.2), samtools (v1.2), BLAST+ (v2.4), edgeR (v4.2.0), Clustal Omega, Primer-BLAST, IGV, and a modified Ribocutter pipeline (2021 preprint version). Gene annotations and ortholog mappings were retrieved using BioMart (Ensembl releases 106, 109, and 113). Custom codes developed in the study are available at https://github.com/amirshahr/MIR2954.

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw sequencing data (mRNA-seq, small RNA-seq, DNA-seq, and Ribo-seq) have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1079296 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1079296) . Processed data for mRNA-seq and small RNA-seq are available in the Gene Expression Omnibus (GEO) repository under the same BioProject accession. Dosage sensitivity scores are available from Zenodo (zenodo.org/

records/6347673). The list of ohnologs is available from the OHNOLOGS database (ohnologs.curie.fr). Gene phylogenetic ages are available from GenTree (gentree.ioz.ac.cn). Genome assemblies, 3′ UTR sequences, and 1:1 orthologs were retrieved from Ensembl releases 109–113 (ensembl.org) and BioMart (ensembl.org/biomart/martview/64183b5d28d21d3d657ef39fbb9ad5b2). The following genome assemblies were used: chicken (Gallus gallus) bGalGal1.mat.broiler.GRCg7b, GRCg6a, and galGal4; zebra finch (Taeniopygia guttata) bTaeGut1_v1.p; crocodile (Crocodylus porosus) CroPor_comp1; ostrich (Struthio camelus) ASM69896v1; human (Homo sapiens) GRCh38.p14; and mouse (Mus musculus) GRCm39.

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

Life sciences study design

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

Sample size	No statistical methods were used to determine sample size. Sample size was based on the number of individuals available (see Supplementary Table 2).
Data exclusions	One batch of eggs was not used due to incubator malfunction.
Replication	Each tissue type from each embryo was represented by 3 replicates derived from 3 individuals, respectively. All attempts at replication were successful.
Randomization	Individuals collected from heterozygous parents were ensured a random distribution of the dissected embryos in terms of genotype and sex.
Blinding	The collection of samples and estimation of phenotypes were performed blindly, with the experimenter unaware of the individuals' genotype and sex.

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
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The primordial germ cells were derived from chicken embryo at embryonic day 2.5.
Authentication	The cell-lined that was used in this study was authenticated based on it appearance and its ability to propagate in the host's gonads.
Mycoplasma contamination	No test for Mycoplasma contamination was performed on the cell lines. The chicken lines used to derive PGCs had been routinely tested and were free of avian mycoplasma.
Commonly misidentified lines (See ICLAC register)	None

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Both male and female adults and embryos of domestic chickens (Gallus gallus domesticus, Hy-Line and Red Junglefowl breeds) were used in this study. Adult chickens, aged between 6 and 24 months, were used for breeding purposes and for generation of riboseq libraries. Chicken embryos were collected at embryonic days 3, 4, 5, 7, 12, 13, and 14. In addition, one adult female mouse brain (Mus musculus, strain: CD-1, RjOrl:SWISS; RRID:MGI:5603077) was used.
Wild animals	No wild animals were used in the study.
Field-collected samples	No field collected samples were used in the study.
Ethics oversight	All animal procedures were conducted in compliance with national and international ethical guidelines and regulations. Mouse experiments were approved by the local animal welfare authorities at Heidelberg University Interfaculty Biomedical Research Facility (T-64/17). Chicken experiments were conducted under UK Home Office license PP9565661 and approved by the Roslin Institute Animal Welfare and Ethical Review Board Committee, and Linköping Council for Ethical Licensing of Animal Experiments (288-2019). Mice (Mus musculus, strain: CD-1, RjOrl:SWISS, RRID:MGI:5603077) were purchased from Janvier Labs (France) and euthanized by cervical dislocation. All chicken (Lohmann white; Gallus gallus) management, maintenance, and embryo manipulation followed the relevant regulatory guidelines.

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