Supplementary information, Data S1 Materials and Methods

Plasmid construction

All PCR reactions were carried out using KOD-Plus-Neo polymerase (TOYOBO) or Q5 Hot Start High-Fidelity DNA Polymerase (NEB). DNA encoding a single copy of the suppressor tRNA expression was assembled by fusing the human U6 promoter (hU6) with the Bst-tRNA gene [1]. Restriction site SpeI was introduced at the 5' end using PCR primers (T1-S) and (T1-A). NheI and EcoRI sites were introduced at the 3' end using PCR primers (T2-S) (T2-A). The tRNA cassette was created by fusing the two PCR products using primers (T1-S) and (T2-A), and subcloned into the pEASY-Blunt Cloning Vector (YT120). The resultant vector was double digested separately with NheI + EcoRI and SpeI + EcoRI to generate multimerized tRNA cassettes containing 2(YT121), 4 (YT122), 6 (YT123) or 8 (YT124) copies. The C-terminal FLAG tagged AzF-RS [1] with EF1α promoter and SV40 polyA was constructed by PCR using primers (AS2-S), (AS2-A), (AS1-S), (AS1-A), (AS3-S) and (AS3-A), and subcloned into pEASY-Blunt Cloning vector. The resultant vector (YT125) and tRNA containing vectors YT120, YT121, YT122, YT123 and YT124 were double digested with NheI + EcoRI to generate vectors containing RS$(tRNA)_n$ (YT126, YT127, YT128, YT129, YT130). The fluorescent EGFP reporter with an amber stop codon (TAG) introduced at the Y39 position was generated using primers (R1-S) and (R1-A) (YT131). The fluorescent EGFP-mCherry reporter was constructed by fusing the wild-type EGFP to mCherry linked by the amber stop codon (TAG). EGFP was amplified by PCR using primers R2-S and R2-A, and subcloned into pCDNA3.1/myc-His using BamH1 and EcoR1 restriction sites (YT132). mCherry was amplified by PCR using primers R3-S and R3-A, and subcloned in YT132 using EcoR1 and Xho1 restriction sites. The resultant vectors were double digested and ligated (YT133). EGFP was amplified by PCR using primers R4-S and R4-A, and subcloned into pLVX-mCherry-N1 lentiviral vector using XhoI and EcoRI restriction sites (YT134). In transgenic fish AzF-RS expression is driven by a ubiquitous promoter (ubi), the multiple tRNA cassette with human U6 promoters and

the reporter fluorescent proteins with the sCMV promoter. Primer sequences are listed at the end of ***Materials and Methods***.

Mammalian HEK293 cell culture and transfection

HEK293T cells (ATCC) were maintained in medium containing DMEM (Gibco), 10% fetal bovine serum (Biowest) at 37°C in a 5% CO_2 atmosphere. Cells were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. For transfection in 12-well plates, 0.8 μg of RS(tRNA)n, 0.8 μg of EGFPamb or EGFP-TAG-mCherry were mixed with Lipofectamine 2000 and added to cells at 70–80% confluence according to the manufacturer's protocol. Four hours after transfection, cells were fed with equal volume of DMEM supplemented with 10% fetal bovine serum. The unnatural amino acid *p*-azido-L-phenylanaline (AzF) (Chem-Impex Ineternational) was supplied at a final concentration of 1 mM 4 h post-transfection. Cells were harvested 24-48 h post-transfection and imaged or analyzed.

Optimization of tRNA/aaRS expression constructs for genomic integration

Before generation of a transgenic animal, we began by creating a gene construct to efficiently express the orthogonal tRNA/aaRS pair for mouse genomic integration. The amber suppressor $tRNA_{CUA}$ derived from *Bacillus stearothermophilus* $tRNA^{Tyr}$ (*Bst*-tRNA) [1] and the AzFRS, an *Escherichia coli* TyrRS variant (Y37L/D182S/F183M/L186A) specific for the Uaa *p*-azido-L-phenylalanine (AzF) [2], were chosen. This $tRNA_{CUA}$/AzF-RS pair has been demonstrated previously in mammalian cells [3-5], amphibian *Xenopus laevis* oocytes [6] and primary neurons of rats [7] to efficiently incorporate AzF in response to an amber stop codon TAG. To express the tRNA/AzF-RS genes of bacterial origin in mouse, we drove the tRNA expression by the type-3 pol III promoter U6 [8], and the AzF-RS gene by the human elongation factor 1 alpha (EF1α) promoter. The AzF-RS gene was fused with various copies of tRNA in one plasmid ($RS/tRNA_n$), because previous experiments in mammalian cells [4] and *C. elegans* [9] suggest that tRNA expression is limiting. By using the EGFP with an amber stop codon at the permissive Y39 site (EGFPamb) as the reporter, we tested the amber codon suppression of designed genes in mammalian cells [8]. HEK293T cells

were cotransfected with the EGFPamb reporter plasmid and the RS/$tRNA_n$ plasmid with the suppressor tRNA gene cassette multimerized at 2×, 4×, 6×, and 8× copies per plasmid (Figure 1A). Suppression of the Y39amb codon in eGFP results in full-length eGFP, whose fluorescence signal reports the incorporation of AzF at the TAG codon. In the presence of 1 mM AzF in the growth media, eGFP fluorescence intensity of cells, as measured by fluorescence microscopy and FACS, increased in comparison to the absence of AzF, indicating the incorporation of AzF into eGFP. In addition, the number of green fluorescent cells increased when the tRNA expression cassette was raised from 1× to 4×, but plateaued out from 4× to 8× (Data not shown due to lack of space). We therefore decided to use the RS$(tRNA)_{4x}$ gene cassette for transgenesis.

Pronuclear injection of zygotes

Mouse zygotes were obtained by superovulation of C57BL/6J females (SLAC Shanghai) mating with the males of the same strain. The zygotes were cultured in KSOM embryo culture medium (Millipore) before injection. Linearized DNA (5 ng/μl) was injected into the nuclei of one-cell stage embryos through the injection needle. Injections were performed using an Eppendorf transferMan NK2 micromanipulator. Injected zygotes were transferred into pseudopregnant female mice immediately after injection or after overnight culture in KSOM medium. Mice were housed in standard cages in a specific pathogen-free facility on a 12 h light/dark cycle with ad libitum access to food and water. All animal experiments conformed to the regulations drafted by the Association for Assessment and Accreditation of Laboratory Animal Care in Shanghai and were approved by the East China Normal University Center for Animal Research.

Amplifying the target RS$(tRNA)_{4x}$ by PCR

Toe or tail clips were subjected to standard DNA extraction procedure. The extracted DNA was amplified using the primers (Azf-S1 and Azf-A1) around the target sites to produce amplifications at sizes of 490 bp. The PCR products were resolved on a 1.5% agarose gel using electrophoresis.

Detection of gene copy number by qPCR

Toe or tail clips were subjected to standard DNA extraction procedure, and treated with 0.5 U/μl RNaseA (Takara) for 30min at 37°C. qPCR was performed against the housekeeping gene Gapdh using following primers based on the protocol [10]. For exogenous gene RS(tRNA)$_{4x,}$ Azf-S1 and Azf-A2 primers were used. The final concentration of RS(tRNA)$_{4x}$ primers were 400 nM each. Gene Expression Assay (MX3005P; Agilent Stratagene) was used for assaying Gapdh (endogenous control) in separate reactions. This proprietary assay resulted in a 223 bp product, detected with a SYBR Green probe. All reactions were 40 cycles using standard cycling conditions (initial 30s at 95°C and 40 cycles of 30 s denaturation at 95°C and 30 s annealing at 58°C and extension at 72°C, finally 1 cycle of 30 s at 95°C, 30 s at 58°C, 30 s at 95°C). The number of RS(tRNA)$_{4x}$ transgene copies in 100 ng DNA from RS(tRNA)$_{4x}$ transgenic mice or from wild-type mice was calculated from their respective CT (cycle threshold) using the linear equation from the respective plasmid standard curve. RS(tRNA)$_{4x}$ copy number per animal was calculated by dividing the number of RS(tRNA)$_{4x}$ to the number of Gapdh.

Western blot analysis of AzF-RS

Tissues were dissected and rapidly frozen in liquid nitrogen. Samples were homogenized with a homogenizer (IKA) in 400 μL sample buffer containing lysis buffer (50 mM Tris-base pH 7.4), 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, and Complete Mini Protease Inhibitor cocktail; Roche). Cultured cells from a 6 cm dish reaching 90~100% confluence were harvest and lysed in lysis buffer comprising 50 mM Tris-base (pH 7.4), 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, and Complete Mini Protease Inhibitor cocktail (Roche). Protein concentration was measured using the BCA Protein Assay Kit (Pierce). 100 μg of protein from each tissue sample was loaded onto a gradient SDS-PAGE. The gel was run at 60 V for 1 h then at 80 V for 1.5 h. Separated proteins were transferred to a PVDF membrane (Millipore) at 100 V for 1.5 h in a room temperature. The PVDF membrane was blocked for one hour with 5% w/v BSA (Sigma-Aldrich), PBS at 25°C with gentle shaking. The blocked membrane was incubated with a rabbit anti-FLAG polyclonal antibody (MFCD02262912,

Sigma-Aldrich, 1:2500 dilution, Primary Antibody Dilution Buffer; Beyotime) overnight at 4°C and then washed in PBST. The blot was then incubated with secondary antibody IRDye 800 CW Goat anti-Rabbit IgG (H+C) (LI-COR, 1:5000 dilution) for one hour at room temperature. After washing with PBST, the blot was exposed to Odyssey (LI-COR) to detect signal. Protein loading was monitored by anti-GAPDH antibody (EPR16891, Abcam, 1:3000 dilution) or anti-β-actin antibody (A5441, Abcam, 1:5000 dilution).

Isolation and culture of primary cells from transgenic RS(tRNA)$_{4\times}$ mice.

Fibroblasts: adult mouse fibroblast cultures were obtained according to the protocol established in the laboratory of Dr. Evan Eichler (University of Washington, https://genome.ucsc.edu/ENCODE/protocols/cell/mouse/Fibroblast_Stam_protocol.pdf) from mice that were less than six months old (ear punch or tail clip). Mouse embryonic fibroblasts (MEF): isolation and culture of MEFs from mouse embryos were performed according to the protocol [11]. Neural stem cell lines culture and differentiation. Neural stem cell lines (n = 3) were derived from the subventricular zone (SVZ) of adult C57BL/6 and transgenic RS(tRNA)$_{4\times}$ mice, respectively based on the protocol [12]. Cells were seeded in T25 flask (Nunc) at 5×10^4 cells per milliliter and maintained as neurospheres in full neurosphere medium, consisting of DMEM/F12 (Gibco. Invitrogen), 0.6% glucose, 3 mM $NaHCO_3$, 5 mM HEPES, 2 mM L-glutamine (Sigma), supplemented with hormone mix: 0.1 mg/ml apo-transferrin, 25 μg/ml insulin, 60 μM putrescene, 30 mM sodium selenite, 20 nM progesterone, 1% BSA (Sigma), 20 ng/ml EGF (BD Biosciences) and 10 ng/ml FGF-2 (Roche Diagnostics). Primary cultures were passaged every 6 days. For the induction of neural stem cell differentiation, dissociated stem cells were plated at 2×10^5 cells per well in basal neurosphere medium (full media without EGF and FGF-2) on poly-L-lysine coated 6-well plates (Nunc). After 24 h, cells were transfected with the lentivirus carrying the EGFP-amb-mCherry reporter. Cells were differentiated for following 7 days under differentiation conditions, with or without 1 mM AzF. Bone marrow cells were extracted and cultured based on the protocol [13]. Femur and tibia of transgenic RS(tRNA)$_{4\times}$ mice (5-6 weeks old; n=3) were cut out to obtain clean bones,

flushed with PBS+2% FCS and bone marrow cells cultured in DMEM containing 20% fetal calf serum, 20 ng/ml IL3, 50 ng/ml IL6 and 50 ng/ml stem cell factor for 24 hours. Cells were transduced using retronectin coated plates, lentiviral supernatant and polybrene (6 μg/ml) and cultured in the presence or absence of 1 mM Azf for 5 days prior to FACS purification of high GFP expressing cells and fluorescence imaging.

Flow Cytometry and FACS

Cells were trypsinized in 0.25% trypsin (Gibco) and triturated to dissociate aggregates and taken up in full medium before centrifugation. Cells were centrifuged at 800 rpm for 5 min, washed in PBS and 5% FBS, and resuspended in DMEM for analysis on FACSCalibur Flow Cytometer (BD Biosciences) equipped with Cellquest software (BD Biosciences). Bone marrow derived cells were sorted for high GFP expression using an Aria II (BD Biosciences) prior to microscopy.

Lentivirus production

Polyethylenimine (Polysciences, PA) was used to transfect HEK293T cells. For the lentivirus packaging, 3×10^6 cells were plated on a 10 cm plate the day before the transfection. The lentiviral transfer vector [10 μg] and the packaging plasmids (psPAX2 [10 μg], pMD2.G [10 μg] were evenly mixed and dissolved in 910 μL Optimal minimum essential media (Invitrogen). Polyethylenimine (90 μL) was added subsequently. The mixture was vortexed gently. After 20-30 min of incubation at room temperature, the transfectant was evenly dropped onto the cells. After 6-8 h of post-transfection, the medium was exchanged for fresh medium. Lentiviral particles were collected 72 h post transfection. For the preparation of concentrated lentivirus, the above procedure was scaled up to 10 plates. Virus-containing media was centrifuged at 1000 rpm for 5 min to remove cell debris and then filtered through 0.45 μm filter (Life sciences). The filtered virus-containing media were ultracentrifuged at 25000 rpm for 3 h (4°C). The virus-containing white pellet was dissolved with phosphate-buffered saline (PBS containing Ca^{2+} and Mg^{2+}; Sigma) by gentle vortexing. The concentrated viruses were stored at 80°C in aliquots. Functional titers were estimated around 7×10^8 transducing units per milliliter by enhanced green

fluorescent protein expression, which was measured by flow cytometry (FACSCalibur, Flow Cytometer, BD) analysis with limiting dilution in HEK293T cells.

RNA-seq

RNA-seq was carried out from three biological replicates for wild type and transgenic mice. The liver tissue was homogenized in liquid nitrogen and total RNA extracted using RNAiso Plus (TaKaRa). RNA quality was assessed using a bioanalyzer (BerryGenomics, RNA integrity number (RIN) > 9). Library generation and sequencing was carried out at BerryGenomics (http://www.berrygenomics.com/) in Beijing. We used TopHat2 to align reads to the mouse genome (build mm10). Reads per kilobase of transcript per million mapped reads and differential expression values were calculated using CuffLinks and CuffDiff. Data analysis was performed using R. Gene ontology analysis was performed using R package clusterProfiler. Statistically significant changes ($P < 0.05$). These data will be uploaded to the Gene Expression Omnibus database.

Blood biochemical index determination

Mouse blood was collected by retro-orbital bleeding and stored at 4°C for 2 h. Thereafter, the supernatants of the samples were collected after centrifuging at 12000 rpm for 5 min at 4°C. Biochemical parameters of the mouse serum were measured in ADICON Clinical Laboratories (INC, Shanghai).

Histological analysis of tissues.

Tissues from wild-type and RS(tRNA)$_{4x}$ mice were individually dissected, 4% paraformaldehyde-fixed at 4°C for overnight. All tissues were Ethanol and xylene gradient dehydrated and paraffin-embedded (LEICA EG1150C) prior to sectioning. Five-micron transverse sections of tissues were prepared on a rotary microtome (Leica RM2235), a water bath (Leica HI1210) for paraffin sections, and a flattening table (Leica HI1220) for histopathology. H&E staining was performed according to established staining protocols.

Fish care and transgenic fish

Zebrafish colonies (nacre strain) were maintained using standard methods. The animal facility obtained an agreement from the French Ministère de l'Agriculture

(n°C75-05-12), and the protocols were approved by the Ministère de l'Education Nationale de l'Enseignement Supérieur et de la Recherche (00477.02). To maintain a healthy colony, a 14 h light–10 h dark cycle was adopted, and a water temperature of 28°C was maintained, with a maximal density of five fish per litre. Water filtration depended on the aquatic habitat stand-alone fish housing and was operated automatically (Aquatic Habitat Inc, FL, USA). The fish were fed twice a day with live 2-day-old paramecia. The coding sequence of AzFRS and tRNA expression cassette were introduced in zebrafish by standard transgenic methods using a Tol2 strategy[14]. Details of the cloning and transgenesis procedures are available upon request.

Unnatural amino acid (Uaa) delivery in zebrafish in vivo

AzF stock solution was prepared in mineral water (Volvic), diluted directly in embryo water bath during gastrulation, and GFP fluorescence analyzed 24 hpf.

Fluorescence imaging

Transgenic embryos or larvae were anesthetized in tricaine solution and embedded in low-melt agarose (0.8%). Spinning-disk images were acquired using either a 10× (Plan Apochromat 10×, NA=0.45, dry) or a 25× (LCI Plan Neofluar 25×, NA=0.8, oil immersion) objectives on a Zeiss Axio Observer Z1, Yokogawa W1 scanning unit equipped with a Hamamatsu Flash 4.0 camera. HEK293T and primary cells isolated from transgenic mice were visualized with a Hamamatsu digital camera (Model C8484) on an Olympus microscope (BX51WI), or a CCD Digital Camera (OLYMPUS DP71) on an Olympus microscope (OLYMPUS IX71). Bone marrow cells were imaged using DeltaVision confocal microscope (GE Healthcare Life Sciences).

Statistical Analysis

All statistical analyses were performed using GraphPad Prism version 5. Data are expressed as means ± S.E. Means of two groups were compared using Student's t test (unpaired, two-tailed), with $P < 0.05$ was considered to be statistically significant. All the experiments were performed at least three times with similar results.

Primers used in this study:

Primer name	Sequences (5'-3')
T1-S	GTCACTAGTGAGGGCCTATTTCCCATGAT
T1-A	CGGTGTTTCGTCCTTTCCACA
T2-S	TGGAAAGGACGAAACACCGAGCGCTCCGGTTTTTCTGTG
T2-A	GATGAATTCAGAATCAATGCCATCCCCCAGCTAGCCAAAAAAAGGACTGGAGGGG
AS1-S	ATCCATTCGATTAGTGAACG
AS1-A	TCACGACACCTGAAATGGAA
AS2-S	CCATTTCAGGTGTCGTGAGCCACCATGGCAAGCAGTAAC
AS2-A	TTACTTGTCGTCATCGTCTT
AS3-S	GACGATGACGACAAGTAAATCTAGATAACTGATCATAA
AS3-A	GATGAATTCATCGGCTAGC AGCCAAAACACTGCACGAACTAAGATACATTGATGAGTTT
R1-S	ATGCCACCTAGGGCAAGCTGACCCTGAAGTTCATC
R1-A	TCAGCTTGCCCTAGGTGGCATCGCCCTCGCCCTCG
R2-S	ACTGGATCCGCCACCATGGTGAGCAAGGG
R2-A	GCTGAATTCCTACTTGTACAGCTCGTCCA
R3-S	CATGAATTCGCCACCATGGTGAGCAAGGG
R3-A	TCTCGAGTTAGGCGTAGTCGGGGACGTCGTAGGGGTATCCCATCTTGTACAGCTCGTCCA
R4-S	AGTCTCGAGGCCACCATGGTGAGCAAGGG
R4-A	GCAGAATTCCCTACTTGTACAGCTCGTCCA
Azf-S1	CACTTCTGGTATCGACCTGA
Azf-A1	CAGTTCAGAATCGACCAGTG
Gapdh-S	GCCTGGAGAAACCTGTATGT
Gapdh-A	GTGGAAGAGTGGGAGTTGCT
Azf-A2	TCAGGAAGCGGTAAACGTCG

Supplementary References

1. Ye S, Kohrer C, Huber T *et al.* Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis *J Biol Chem* 2008; **283**:1525-1533.
2. Chin JW, Cropp TA, Anderson JC *et al.* An expanded eukaryotic genetic code *Science* 2003; **301**:964-967.
3. Ye S, Zaitseva E, Caltabiano G *et al.* Tracking G-protein-coupled receptor activation using genetically encoded infrared probes *Nature* 2010; **464**:1386-1389.
4. Coin I, Katritch V, Sun T *et al.* Genetically Encoded Chemical Probes in Cells Reveal the Binding Path of Urocortin-I to CRF Class B GPCR *Cell* 2013; **155**:1258-1269.
5. Rannversson H, Andersen J, Sorensen L *et al.* Genetically encoded photocrosslinkers locate the high-affinity binding site of antidepressant drugs in the human serotonin transporter *Nature communications* 2016; **7**:11261.
6. Ye S, Riou M, Carvalho S, and Paoletti P Expanding the genetic code in Xenopus laevis oocytes *ChemBioChem* 2013; **14**:230-235.
7. Zhu S, Riou M, Yao CA *et al.* Genetically encoding a light switch in an ionotropic glutamate receptor reveals subunit-specific interfaces *Proc Natl Acad Sci U S A* 2014; **111**:6081-6086.
8. Wang W, Takimoto JK, Louie GV *et al.* Genetically encoding unnatural amino acids for cellular and neuronal studies *Nat Neurosci* 2007; **10**:1063-1072.
9. Parrish AR, She X, Xiang Z *et al.* Expanding the Genetic Code of Caenorhabditis elegans Using Bacterial Aminoacyl-tRNA Synthetase/tRNA Pairs *ACS Chem Biol* 2012; **7**:1292-1302.
10. Joshi M, Keith Pittman H, Haisch C, and Verbanac K Real-time PCR to determine transgene copy number and to quantitate the biolocalization of adoptively transferred cells from EGFP-transgenic mice *BioTechniques* 2008; **45**:247-258.
11. Xu J Preparation, culture, and immortalization of mouse embryonic fibroblasts

Curr Protoc Mol Biol 2005; **Chapter 28**:Unit 28 21.

12. Brewer GJ, and Torricelli JR Isolation and culture of adult neurons and neurospheres *Nat Protoc* 2007; **2**:1490-1498.
13. Holst J, Szymczak-Workman AL, Vignali KM *et al.* Generation of T-cell receptor retrogenic mice *Nat Protoc* 2006; **1**:406-417.
14. Suster ML, Kikuta H, Urasaki A, Asakawa K, and Kawakami K Transgenesis in zebrafish with the tol2 transposon system *Methods Mol Biol* 2009; **561**:41-63.