HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2022 December 01.

Published in final edited form as:

Science. 2021 August 06; 373(6555): 662–673. doi:10.1126/science.aba4991.

A noncoding RNA modulator potentiates phenylalanine metabolism in mice

Yajuan Li[1,#], **Zhi Tan**[2,3,#], **Yaohua Zhang**[1,#], **Zhao Zhang**[4], **Qingsong Hu**[1,5], **Ke Liang**[1], **Jun Yao**[1], **Youqiong Ye**[4], **Yi-Chuan Li**[4], **Chunlai Li**[1,6], **Lan Liao**[7], **Jianming Xu**[8], **Zhen Xing**[1,9], **Yinghong Pan**[10,11], **Sujash S. Chatterjee**[10], **Tina K. Nguyen**[1], **Heidi Hsiao**[1], **Sergey D. Egranov**[1], **Nagireddy Putluri**[8], **Cristian Coarfa**[8], **David H. Hawke**[12], **Preethi H. Gunaratne**[10], **Kuang-Lei Tsai**[4], **Leng Han**[13], **Mien-Chie Hung**[14,15], **George A. Calin**[16], **Fares Namour**[17,18], **Jean-Louis Guéant**[17,18], **Ania C. Muntau**[19], **Nenad Blau**[20], **V. Reid Sutton**[21], **Manuel Schiff**[22,23], **François Feillet**[18,24,*], **Shuxing Zhang**[2,25,*], **Chunru Lin**[1,25,*], **Liuqing Yang**[1,25,26,*]

[1]Department of Molecular and Cellular Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030, USA

[2]Intelligent Molecular Discovery Laboratory, Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX, 77054, USA

[3]Current address: Center for Drug Discovery, Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX, 77030, USA

[4]Department of Biochemistry and Molecular Biology, The University of Texas Health Science Center at Houston McGovern Medical School, Houston, TX, 77030, USA

[5]Current address: The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui 230001, PR China

[6]Current address: Department of Liver Surgery, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, PR China

[7]Genetically Engineered Mouse Core, Advanced Technology Cores, Baylor College of Medicine, Houston, TX, 77030, USA

[8]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA

[9]Current address: Sanofi U.S., Cambridge, MA, 02139, USA

[*]To whom correspondence should be addressed: f.feillet@chru-nancy.fr, shuzhang@mdanderson.org, clin2@mdanderson.org and lyang7@mdanderson.org.

[#]Contribute equally

Author contributions: L.Q.Y. and C.R.L. conceived the project and designed the experiments. Y.J.L. executed the primary studies. Y.J.L. developed genetic mouse models and related experiments with assistance of Q.S.H. and Y.H.Z. T.K.N., X.Z., L.C.L., H.H., L.L. and J.M.X. Z.T. and S.Z. performed the proteomics and modeling studies. J.Y., Y.Q.Y., Z.Z. and L.H. performed bioinformatics analysis. The histological staining and corresponding analysis were performed by Y.J.L. with assistance of K.L. D.H. executed mass spectrometry analysis. CLIP assay were performed by L.Q.Y., S.S.C., Y.H.P. and P.H.G. Metabolic profiling were performed and analyzed by N. P. and analyzed by C.C. Clinical specimens were ascertained and processed by F.N., J.L.G., F.F. with assistance of M.S. an L.Q.Y. Protein purification were performed by Y.J.L. with assistance of Y.C.L. and K.L.T., S.Z. M.C.H., G.A.C., M.S., A.C.M., N.B., V.R.S. and F.F. contributed to experimental design and data interpretation. S.D.E. assisted with manuscript drafting. L.Q.Y. and C.R.L. wrote the manuscript.

Competing interests: The RNA mimics, including *HULC*, biotinylated *HULC*, GalNAc-tagged *HULC* mimics, are in the process of patent application (MDA19-013).

[10]Department of Biochemistry and Biology, University of Houston, Houston, TX, 77030, USA

[11]Current address: UPMC Genome Center, Pittsburgh, PA, 15232, USA

[12]Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030, USA

[13]Center for Epigenetics and Disease Prevention, Institute of Biosciences and Technology, Texas A&M University, Houston, TX, 77030, USA

[14]Graduate Institute of Biomedical Sciences, Research Center for Cancer Biology, and Center for Molecular Medicine, China Medical University, Taichung, 404, Taiwan

[15]Department of Biotechnology, Asia University, Taichung 413, Taiwan

[16]Department of Translational Molecular Pathology, Division of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030, USA

[17]Department of Molecular Medicine and Reference Center for Inborn Errors of Metabolism, University Hospital of Nancy, F-54000, France

[18]INSERM, U1256, NGERE - Nutrition, Genetics, and Environmental Risk Exposure, University of Lorraine, Nancy F-54000, France

[19]University Children's Hospital, University Medical Center Hamburg Eppendorf, Hamburg, 20246, Germany

[20]Division of Metabolism, University Children's Hospital Zurich, CH-8032 Zurich, Switzerland

[21]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, 77030, USA

[22]Necker Hospital, APHP, Reference Center for Inborn Error of Metabolism and Filière G2M, Pediatrics Department, University of Paris, Paris, 75007, France

[23]Inserm UMR_S1163, Institut Imagine, Paris, 75015, France

[24]Pediatric department Reference Center for Inborn Errors of Metabolism Children University Hospital Nancy, Nancy, F-54000, France

[25]The Graduate School of Biomedical Sciences, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030, USA

[26]Center for RNA Interference and Non-Coding RNAs, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030, USA

Abstract

The functional role of long noncoding RNAs (lncRNAs) in inherited metabolic disorders, including phenylketonuria (PKU), is unknown. We demonstrated that the mouse lncRNA *Pair* and human *HULC* associate with phenylalanine hydroxylase (PAH). *Pair*-knockout mice exhibited excessive blood phenylalanine, musty odor, hypopigmentation, growth retardation, and progressive neurological symptoms including seizures, which faithfully models human PKU. *HULC* depletion led to reduced PAH enzymatic activities in human induced pluripotent stem cell (hiPSC)-differentiated hepatocytes. Mechanistically, *HULC* modulated the enzymatic activities

of PAH by facilitating PAH-substrate and PAH-cofactor interactions. To develop a therapeutic strategy for restoring liver lncRNAs, we designed GalNAc-tagged lncRNA mimics that exhibit liver enrichment. Treatment with GalNAc-*HULC* mimics reduced excessive phenylalanine in *Pair*$^{-/-}$ and *Pah*$^{R408W/R408W}$ mice and improved the phenylalanine tolerance of these mice.

One Sentence Summary:

A noncoding RNA chaperone of phenylalanine hydroxylase alleviated phenylketonuria in a mouse model.

The predominant view of human genetic diseases revolves around the identification of mutations/ dysfunctions of coding genes. However, about 98% of human genomic mutations occur within non-coding regions, and an understanding of the functional roles of noncoding RNAs in human genetic diseases is lacking. Phenylketonuria (PKU, OMIM 261600) and its milder variant hyperphenylalaninemia (HPA) are genetic disorders caused by a deficiency in the hydrolysis of L-phenylalanine (Phe) to L-Tyrosine (Tyr) (1). Over 1,000 PAH variants (PAHvdb database; http://www.biopku.org/home/pah.asp) have been identified and associated with PAH deficiency. Most of the variants are missense, usually resulting in protein misfolding and/or impairment of catalytic functions (2). Roughly 1 in every 10,000 infants is affected by this disease (3). Based on blood Phe concentration during diagnosis or screening in the neonatal period, PKU can be categorized as PKU (Phe >900 μmol/L), mild PKU (Phe <900 but > 360 μmol/L), or mild HPA (blood Phe is higher than the normal limit, but <360 μmol/L) (4, 5). Excessive Phe concentrations in patients with untreated PKU cause brain damage and associated mental retardation (6). Although PKU is largely considered a monogenic disorder, clinical evidence has indicated that PAH variants are incompletely correlated with the metabolic phenotypes of patients with PKU (7-9). In some patients diagnosed with PKU, no mutations could be identified in the *PAH* gene, suggesting that unknown factors may contribute to PKU (10). Consistent with this notion, the recent discovery of biallelic mutations in the *DNAJC12* gene in patients exhibiting high blood Phe concentrations without mutations in *PAH* or genes involved in BH_4 (tetrahydrobiopterin, a PAH cofactor) metabolism (11) suggested the possibility that non-*PAH* genes may affect PAH function and subsequently increase blood Phe concentrations.

Newborns diagnosed with PKU are typically treated with a Phe-restricted diet and/or BH_4 supplementation. Patients with untreated PKU exhibit intellectual disability, behavioral issues, seizures, and psychiatric disorders (12). For patients with certain genotypes of mild PKU, supplementation with BH_4 has been suggested as an enzyme enhancement therapy (13). However, patients with a severe form of PKU (Phe concentrations above 900 μmol/L and usually above 1200 μmol/L) require additional treatment considerations. These patients also frequently respond poorly to BH_4 treatment (14). Given the potential immunogenic nature of Phenylalanine ammonia lyase (PAL) enzyme supplementation or substitution therapy (15, 16) and the uncertainty of gene therapy (17), additional therapeutic strategies are urgently needed.

RESULTS

Depletion of lncRNAs drives PKU

Long intergenic noncoding RNAs (lincRNAs) and long non-coding RNAs (lncRNAs) are transcripts with low coding potential. Aiming to investigate the biological importance of lncRNAs, we determined the lncRNA profile of mouse E18.5 embryos and the livers of 2-month-old adult mice (Fig. 1A and fig. S1A to C), finding that 2210408F21Rik (NR_040259) (renamed *Pair:* PAH-activating lincRNA) is one of the most upregulated lncRNAs in the liver of adults compared to the E18.5 embryos and exhibits low coding potential (CNIT score (−0.3544) (18) (fig. S1D to H, and Table S1). To deplete this mouse lncRNA, we introduced a two-nucleotide mutation using CRISPR/Cas9 to interrupt the splicing site after the first exon (Fig. 1B). Northern blotting indicated that *Pair* exhibits two major isoforms with molecular weights 730 bp and 1.5 kb, which were both depleted upon the introduction of the splicing-site mutation (Fig. 1C). Compared to wild-type and heterozygous littermates, $Pair^{-/-}$ mice exhibited similar expression of neighboring genes with no detectable alterations in major organ development (fig. S2A to D).

Both male and female $Pair^{-/-}$ mice exhibited hypopigmentation (Fig. 1D), growth retardation and elevated serum Phe concentrations (Fig. 1E-G) reminiscent of human PKU. $Pair^{-/-}$ livers showed no detectable changes in the expression of BH_4 biogenesis genes or PAH protein abundance (fig. S2E-G). $Pair^{-/-}$ livers exhibited enzymatic deficiency in converting Phe to Tyr, and $Pair^{+/-}$ livers showed impaired PAH enzymatic activity (fig. S2H-I). $Pair^{+/-}$ mice exhibited blood Phe concentrations within the normal range; however, these animals showed elevated blood Phe concentrations upon Phe challenge compared to $Pair^{+/+}$ mice (fig. S2J). These results suggested that $Pair^{+/-}$ livers exhibit partial PAH deficiency. Sanger sequencing of the *Pah* gene suggested that Pah^{enu2} mice harbor a T788C mutation, while $Pair^{-/-}$ mice harbor a wild-type *Pah* gene (Table S2 and fig. S2K and L).

The median life-span of $Pair^{-/-}$ mice is 15.2 months (Fig. 1H), and more than 70% of the mice exhibited seizures starting at the ages of 8-10 months (Fig. 1I-J and Video S1-3). Following cardiopulmonary resuscitation (CPR), *Pair* KO mice experiencing a seizure can be rescued (Video S4). $Pair^{-/-}$ brains were smaller and exhibited reduced tyrosine hydroxylase (TH)-positive neurons compared to wild-type and heterozygous littermates (Fig. 1K-L). $Pair^{-/-}$ livers and serum exhibited diminished tyrosine concentrations. Tyrosine is catalytically produced by PAH in both serum and liver tissue (fig. S3A-E), confirming the presence of PAH deficiency. The role of *Pair* in seizures, in addition to regulating PAH enzymatic activities, could not be ruled out. Our data suggested that $Pair^{-/-}$ mice model human PKU.

Pair and *HULC* associate with PAH

To understand the molecular mechanism of *Pair*, we performed *Pair* pulldown using biotinylated sense or anti-sense *Pair* in mouse livers (Fig. 2A and Table S3). Sense, but not antisense *Pair,* associated with mouse PAH (Fig. 2A). Beads-only and polyA were used as negative controls, while the association of ELAV1 with AR_3' UTR served as a positive control (19) (Fig. 2A). We then performed CLIP (cross-linking immuno-precipitation) assay

using mouse *Pair*$^{+/+}$ and *Pair*$^{-/-}$ livers or human liver tissues from two healthy donors (fig. S4A and B, Table S4). The presence of PAH-RNA complexes was diminished upon *Pair* knockout (fig. S4A). PAH-lncRNA(s) complexes were detected in the two human liver donors (fig. S4B). The PAH-RNA complexes (fig. S4A-B, blue boxes 1-3) were subjected to reverse transcription and Sanger sequencing (Table S5). As expected, mouse PAH associated with mouse *Pair* (nt. 460-496) in *Pair*$^{+/+}$ livers (Fig. 2B, bottom row). Human PAH associated with one human lncRNA gene, *HULC* (nt. 183-216) (Fig. 2B, top row). Although *HULC* has been suggested to be upregulated in liver cancer (20), northern blotting indicated that *HULC*, similarly to *Pair*, is specifically expressed in normal liver tissues (fig. S4C to E), suggesting the biological relevance of *HULC* in liver homeostasis and function.

RNA immunoprecipitation (RIP) assay confirmed that *Pair* and *HULC* associate with PAH protein in mouse and human livers, respectively (fig. S5A to C). We demonstrated that the regulatory (aa. 1-142) and catalytic (aa. 142-411) domains of PAH are required for PAH-*Pair* interactions (fig. S5D-G). Next, using primary cultured mouse or human hepatocytes, we demonstrated that PAH protein faithfully co-localizes with *Pair* or *HULC* in the cytosol, but not with another cytosolic lncRNA, *Tug1* (Fig. 2C and fig. S5H to I).

We substituted each nucleotide between *Pair* (nt. 460-496) and *HULC* (nt. 183-216) using the most common transition and transversion types (21) and expressed wild-type or mutant *Pair* or *HULC* in *Pair*$^{-/-}$ or *HULC*-deficient hepatocytes (fig. S6A and B). RIP assay suggested that *Pair* (nt. 470-488) and *HULC* (nt. 183-200) are required for PAH-*Pair* and PAH-*HULC* interactions (fig. S6A and B). Furthermore, *Pair* 479A>G and *HULC* 191A>G abolished PAH-*Pair* and PAH-*HULC* interactions, respectively (fig. S6A and B).

SHAPE assays (22, 23) indicated that *HULC* U190 and A191 (capillary electrophoresis size #290 and 289) exhibited chemical labeling, suggesting the presence of a loop structure flanked by low chemical probing (Fig. 2D and fig. S6C). Furthermore, *HULC* harbors 3 additional stem-loop structures, as revealed by SHAPE assays (fig. S7A-D). These 3 stem-loop structures exhibited distinct structures compared to the stem-loop at nt. 184-216 (fig. S7E-H). The chemical labeling of *HULC*^{A191} was verified by in vivo SHAPE (fig. S8A-B). We further looked at the secondary structure of *Pair*, finding that *Pair* nt. 467-484 also exhibits a stem-loop structure, with A479 showing robust chemical probing (fig. S9A-C). *Pair* nt. 467-484 exhibited a similar 3-dimensional structure to *HULC* nt. 184-216 but not the above 3 stem-loops of *HULC* (fig. S9D).

Wild-type *Pair* and wild-type *HULC*, but not the *Pair* 479A>G or *HULC* 191A>G mutants (referred to as *Pair* mut or *HULC* mut, respectively) associated with recombinant PAH, as revealed by RNA electrophoretic mobility shift assay (EMSA) assay (Fig. 2E, F). Expression of MS2-tagged wild-type *Pair/HULC*, but not the mutants, rescued the association with PAH protein in *Pair*$^{-/-}$ hepatocytes (Fig. 2G and fig. S10A). We then performed a rescue CLIP assay by expressing exogenous wild-type *Pair/HULC* or mutants in mouse *Pair*$^{-/-}$ hepatocytes (Fig. 2H and fig. S10B). Exogenous *Pair* or *HULC* WT, but not mutants, associated with PAH similarly to endogenous *Pair* or *HULC* (Fig. 2H). Similar amounts of PAH protein were immuno-precipitated using anti-PAH antibody (Fig.

2H, bottom). These findings suggested that the human lncRNA *HULC* and mouse *Pair* both associate with PAH.

HULC/Pair modulate the enzymatic activity of PAH

To demonstrate the underlying molecular mechanisms of *HULC/Pair* in the enzymatic activity of PAH, we first determined that there are roughly 800 *HULC* RNA, 700 *Pair* RNA, and 4,000 PAH protein molecules per human or mouse hepatocyte (fig. S10C-E). In $Pair^{-/-}$ or *HULC*-deficient hepatocytes, exogenous expression of *Pair/HULC* in a dose-dependent manner led to the reduction of cellular Phe concentrations (fig. S10F and G). The concentrations of amino acids other than Phe and Tyr were minimally affected upon *Pair* depletion (fig. S10H).

The enzymatic activity of PAH requires the presence of a cofactor, BH_4 (24). Mutations of PAH affecting PAH-Phe or PAH-BH_4 interactions impair the catalytic activity of PAH (25). Previous research indicated that the Phe molecule associates with the regulatory domain of PAH as an allosteric activator (26-28). Using the crystal structure information of PAH (PDB 6HYC), protein structural bioinformatics analysis suggested that *HULC* nt. 184-216 associates with the regulatory domain of PAH and allosteric Phe (Fig. 3A). *HULC* A^{191} forms hydrogen bonds with both Thr63 and His64 (Fig. 3A). The mutation of A^{191} to G^{191} causes a change at position 6 from an amino group (6-NH_2) to a carbonyl group (6-CO). This change leads to the loss of the hydrogen bond between A^{191} and His64, because carbonyl groups are hydrogen bond acceptors while amino groups are often hydrogen bond donors. The amino group (2-NH_2) of G^{191} may also cause steric hindrance with Thr63. Phe forms hydrogen bonds with PAH via Asn61 and Leu62, and it also forms stacking interactions with A^{191} (Fig. 3B). These interactions stabilize the whole structural complex in a conformation that makes the active site fully accessible to Phe as a substrate and BH_4 as a cofactor. Our findings suggested that *HULC* serves as an important factor that stabilizes the interaction between allosteric Phe and PAH, as previous hypothesized (26). Aside from A^{191}, *HULC* interacts with PAH via several other residues to achieve the binding specificity of the PAH-*HULC* interaction: A^{195} forms a hydrogen bond with Tyr166; A^{214} forms a hydrogen bond with Arg157; and G^{202} also interacts with Tyr154 via a hydrogen bond (Fig. 3C). *Pair* adapts a binding mode to PAH that is similar to *HULC*: both *Pair* and *HULC* have two nucleotides that stick out to form a T-shape and exhibit stacking interactions with His64 of PAH, and these two nucleotides (A and U) stabilize each other through stacking interactions (Fig. 3D). The only difference between *Pair* and *HULC* with regard to this behavior is that the order of the sequences of these two nucleotides in *Pair* is UA; in *HULC*, the order is AU (Fig. 3D).

We reasoned that the *HULC*-PAH interaction may facilitate the binding of PAH to Phe or BH_4. To address this hypothesis, we synthesized biotinylated-Phe and -BH_4 (referred to as Bio-Phe and Bio-BH_4) (fig. S11A). Compared to $Pair^{+/+}$ livers, $Pair^{-/-}$ livers exhibited a similar abundance of PAH protein; however, the PAH-Phe and PAH-BH_4 interactions were impaired following *Pair* depletion (Fig. 3E). We then determined that digoxin (DIG)-tagged *HULC* showed detectable interactions with bacterially-expressed PAH proteins but undetectable associations with Bio-Phe or Bio-BH_4 (fig. S11B). Wild-type PAH was

included as a positive control (fig. S11B). PAH G46S, F55L, and P281L have been suggested to impair interactions between Bio-Phe and/or Bio-BH_4 (29, 30), which was confirmed (fig. S11B). LncRNA *LINK-A* and the interaction between *LINK-A* and PIP_3 (31) were included as negative controls (fig. S11B).

The N-terminal regulatory domain of PAH can undergo a conformational change to switch between “open” and “closed” states (30). We hypothesized that the binding of *HULC* may stabilize the PAH-Phe-*HULC* complex and stabilize the PAH protein in the “open” state (fig. S11C). We applied limited proteolysis (LiP) followed by liquid chromatography–mass spectrometry (LC–MS) analysis (LiP-LC–MS) (32) to address this hypothesis (fig. S11C). Open loop regions (gray) and Lip-resistant regions (dark blue) were determined (Fig. 3F and Table S6). Region aa 57-66 showed resistance to LiP in the presence of *HULC* but not the *HULC* mut, suggesting that this region associates with *HULC* (Fig. 3F, top panel-magenta). Notably, aa 57-66 of PAH exhibited recovery in the presence of Phe (Fig 3F, bottom panel-magenta), which was consistent with the structural modeling, demonstrating that Thr63 and His64 associate with *HULC* and Phe as a complex. Furthermore, LiP-MS suggested a few regions that were dynamically regulated upon *HULC* binding (Fig.3F, green).

Hence, our data suggested that *HULC* associates with PAH in vitro and facilitates the potential conformational change of PAH. Consistent with this notion, expression of wild-type *Pair* or *HULC* rescued PAH-Phe and PAH-BH_4 interactions and reversed cellular Phe accumulation, while the *Pair* and *HULC* mutants failed to do so (Fig. 3G-H and fig. S11D-E).

To determine the functional role of the *HULC*-PAH interaction in vivo, *HULC*-deficient hiPSC were further differentiated into hepatocytes expressing wild-type *HULC* or the A191G mutant (fig. S12A). Expression of wild-type *HULC* or the A191G mutant showed no detectable effects on the abundance of PAH protein (fig. S12B). Upon depletion of *HULC*, the conversion of ^{14}C-Phe to ^{14}C-Tyr was reduced with concurrent elevation of cellular Phe concentrations (fig. S12C and D). Expression of WT *HULC*, but not the A191G mutant, restored the enzymatic activity of PAH and reversed the cellular accumulation of Phe (fig. S12C and D). Similarly, wild-type PAH, but not the TH63-64PN mutant, expressed in *PAH*-deficient hiPSC-hepatocytes rescued the enzymatic deficiency of PAH (fig. S12E-H). Taken together, our findings suggested that *Pair*-PAH and *HULC*-PAH interactions facilitate the PAH-driven catalysis of Phe to Tyr.

HULC mimics restore PAH enzymatic activity

We reasoned that supplementing lncRNAs might improve the catalytic activities of the PAH mutants, leading to reduced serum Phe concentrations and improved symptoms in patients with PKU. We designed Scramble (Scr) and *HULC* mimics representing wild-type *HULC* nt. 181-201 and *HULC* A191G mutated sequences for the following studies. We first selected 17 PAH mutants that were identified from patients with PKU and are known to affect the enzymatic activity of PAH (33, 34), and we collected the bacterially-expressed wild-type and mutant PAH (Fig. 4A and fig. S13A). EMSA assay indicated that 13 of the 17 PAH mutants, as well as wild-type PAH, associated with the *HULC* mimics; on the contrary, PAH TH63-64PN, R157N, N207S, and S349L failed to associate with the *HULC*

mimics (fig. S13B). The denatured wild-type PAH proteins (WT denat.) were included as a negative control (fig. S13B). We further quantified the interactions between PAH proteins (WT/mutants) and *HULC* mimics (WT), finding that PAH protein interacted with the *HULC* mimics with a K_d value of 131.6 nM (fig. S13C). PAH TH63-64PN, R157N, N207S, and S349L, but not the other PAH mutants, exhibited decreased binding affinities (fig. S13C).

We measured the binding affinities between PAH (WT/mutants) and Bio-Phe or Bio-BH_4 in the presence of *HULC* mimics (fig. S14A-O and S15A-O) and summarized the change in binding affinities compared with PAH WT in Figs. 4A-B. LncRNA mimics representing *LINK-A* 1100-1117 (31) were included as a negative control (Figs. 4A-B). For wild-type PAH protein, the presence of *HULC* mimics enhanced PAH-Phe (Fig. 4A) and PAH-BH_4 interactions (Fig. 4B), while *HULC* mut mimics failed to do so (Fig. 4A-B and S14A-O and S15A-O). The PAH mutants we tested all exhibited impaired binding affinities toward Phe and/or BH_4 compared to wild-type PAH protein (Fig. 4A-B and S14A-O and S15A-O). The presence of *HULC* mimics, but not *HULC* mut, enhanced the affinity of Phe and/or BH_4 to the PAH mutants. A PAH catalytic pocket deletion mutant (Δ245-379) was included as a negative control (fig. S14O and S15O).

We further demonstrated that wild-type PAH effectively catalyzed Phe to Tyr, which was abolished in all 17 PAH mutants we tested (Fig. 4C, left section, bar #2). In the presence of *HULC* mimics, but not *LINK-A* mimics or *HULC* mut, wild-type PAH showed enhanced enzymatic activity in converting Phe to Tyr (Fig. 4C, middle section, bar #2). Furthermore, 11 of the 17 PAH mutants also showed improved enzymatic activities in converting Phe to Tyr (Fig. 4C, middle section, as indicated in red). The enzymatic activities of the rest of the PAH mutants were not significantly affected by *HULC* mimics (Fig. 4C, middle section, light blue and brown).

We then determined the k_{cat} of bacterially-expressed human wild-type and mutant PAH in the presence of Scramble (Scr), wild-type *HULC*, or mutant *HULC* mimics (Fig. 4D and fig. S16A-S). Recombinant PAH proteins with deleted catalytic domains (Δcat) were included as a negative control (fig. S16A). In the presence of the wild-type *HULC* mimics, the k_{cat} of wild-type PAH was increased (Fig. 4D and fig. S16B). All PAH mutants we tested exhibited impaired k_{cat} under identical conditions (Fig. 4D and fig. S16C-S). The presence of wild-type *HULC* mimics, but not mutant *HULC* mimics, enhanced the k_{cat} of PAH F39L, A47V, F55L, I65S, P275L, P281L, I283N, F299C, A300S, I318T, and R408W, but not the enzymatic activities of PAH TH63-64PN, R157N, N207S, and S349L (Fig. 4D and fig. S16C-S). Therefore, our data suggested that the presence of *HULC* facilitates the enzymatic activity of wild-type PAH and a cohort of mutants observed in patients with PKU.

HULC/Pair enhances the enzymatic activity of the PAH R408W mutant

19.2-73% of patients with PKU harbor a R408W mutant (35, 36), and patients with this mutation respond poorly to currently available treatment options. We hypothesized that overexpression of *HULC/Pair* might improve the enzymatic activity of PAH R408W. To address this, we first generated a $Pah^{R408W/R408W}$ mouse strain using CRISPR/Cas9 (fig. S17A-B). Similar to $Pair^{-/-}$ mice, $Pah^{R408W/R408W}$ mice exhibited hyperpigmentation, growth retardation, seizures, and elevated blood Phe concentrations, with no detectable

alterations in the protein stability of PAH or the expression of *Pair* (fig. S17C-H). These observations provided genetic evidence confirming that *Pair/HULC* and PAH act in a linear pathway. In primary cultured mouse hepatocytes isolated from $Pair^{-/-}$ or $Pah^{R408W/R408W}$ mice, we determined cellular Phe, Tyr, and tryptophan (Trp) concentrations, finding that expression of wild-type *Pair/HULC*, but not the mutant (fig. S17I), restored PAH-Phe and PAH-BH_4 interactions and cellular Phe and Tyr concentrations, with no effect on Trp status (fig. S17I-L).

We collected skin fibroblasts from healthy donors who harbored wild-type *PAH* genes and from a patient with PKU (ID: NA02406) who harbored F299C and R408W mutations. We reprogrammed these fibroblasts into hiPSCs. These hiPSCs were further induced into hepatocytes (fig. S18A-B). Two hiPSC clones were used, referred to as PKU #7 and PKU #17 (fig. S18A-C). These hiPSC-derived hepatocytes exhibited similar expression of *HULC, PAH,* and hepatic markers (fig. S18B-C). The hiPSC-derived hepatocytes derived from both PKU #7 and PKU #17 clones exhibited increased Phe concentrations compared to the control (healthy donor), confirming PAH enzymatic activity deficiency (fig. S18D). Expression of full-length *HULC* in these hiPSC-derived hepatocytes reduced Phe concentrations in PKU #7 and PKU #17 hepatocytes (fig. S18D-E). This suggested that a supply of *HULC* might enhance the enzymatic activity of PAH in F299C and R408W mutants.

GalNAc-*HULC* mimics improve phenylalanine metabolism in mice

To design lncRNA mimics that could provide therapeutic value in vivo, we synthesized scramble (Scr), wild-type, and mutant *HULC* mimics using 2'-Fluoro (2'-F) modified RNA monomers to provide nuclease resistance in vivo. The potential secondary structure of wild-type or mutant *HULC* mimics remained identical (fig. S19A). Both full-length *HULC* and the mimics exhibited similar binding affinities for PAH (fig. S19B). Furthermore, the *HULC* A191G mutant and *HULC* mut mimics both abolished these interactions (fig. S19B).

To facilitate the liver-enrichment of *HULC* mimics, we applied three types of *HULC* mimics: *HULC* mimics alone, peptides representing ApoE1 (Apo)-tagged *HULC* mimics, and N-Acetylgalactosamine (GalNAc)-tagged *HULC* mimics via intravenous (i.v.) or subcutaneous (SubQ) injection. GalNac-conjugated oligonucleotides have recently been suggested to assist with liver-targeted siRNA delivery (37). The 3'-triantennary GalNAc was conjugated to the *HULC* mimic.

We considered that a three-day treatment trial (Fig. 5A) would serve as the most convenient method to determine the efficacy of these mimics. Since PAH R408W-harboring patients respond to current BH_4 supplementation treatments poorly, we treated female and male $Pah^{R408W/R408W}$ mice with the indicated mimics (fig. S19C to D). GalNac-*HULC* administered via i.v. injection exhibited the highest efficacy in reducing excessive Phe concentrations in both female and male $Pah^{R408W/R408W}$ mice (fig. S19C to D).

We then determined the concentrations of biotin-labeled *HULC* or GalNAc-*HULC* mimics in the major organs, finding that a substantial portion of biotin-labeled GalNAc-tagged *HULC* mimics was detected in the liver between 3 and 72 hours after dosing, while the lungs

and spleen exhibited no detectable accumulation of GalNAc-*HULC* (fig. S19E to G). Biotin was undetectable in the liver 12 hours post-injection (fig. S19H).

As a proof of concept, we applied GalNAc-tagged Scr or *HULC* mimics using short-term (3-day) and medium-term (12-day) treatment regimens (Fig. 5A). To rule out the indirect effects of the GalNAc tag, GalNAc-*HULC* mut mimics were included (Fig. 5B). Both male and female *Pair*$^{-/-}$ mice exhibited reduced serum Phe concentrations 24 hours after injection of GalNAc-*HULC* mimics, but not the Scr or *HULC* mut mimics (Fig. 5B). During the medium-term treatment, administration of GalNAc-*HULC* mimics reduced blood Phe concentrations in *Pair*$^{-/-}$ mice throughout the treatment term (Fig. 5C). In the *Pah*$^{R408W/R408W}$ mice, administration of GalNAc-*HULC* mimics similarly reduced blood Phe concentrations during the short-term and medium-term treatments (Fig. 5D to E). The *Pah*$^{R408W/R408W}$ mice subjected to GalNAc-*HULC* mimics showed increased blood Tyr concentrations compared to animals given scramble mimics (Fig. 5F). To evaluate Phe clearance capacity, *Pah*$^{R408W/R408W}$ mice were subjected to pre-treatment with Scr or *HULC* mimics followed by a Phe challenge (Fig. 5G). The area under the curve (AUC) showed a reduction in Phe concentrations following GalNAc-*HULC* mimic treatment compared to GalNAc-Scr (Fig. 5H).

To determine whether the administration of GalNAc-*HULC* facilitates higher tolerance to dietary Phe intake in PKU animals, *Pah*$^{R408W/R408W}$ were fed a Phe-free diet for three days followed by water containing increasing doses of Phe. Administration of GalNAc-*HULC* mimics allowed *Pah*$^{R408W/R408W}$ animals to maintain relatively low blood Phe concentrations (<600 μM) upon Phe challenge up to 3.0 mg/ml (Fig. 5I), suggesting that GalNAc-*HULC* mimics may also be able to improve tolerance to dietary Phe.

Sapropterin, a synthetic formulation of BH_4, has been used as enzymatic enhancer for patients with HPA (38). Patients with PKU harboring R408W are resistant to sapropterin treatment (14). We tested the efficacy of sapropterin in *Pair*$^{-/-}$ and *Pah*$^{R408W/R408W}$ mice, finding that supplementation of sapropterin showed no effect on the blood Phe of *Pair*$^{-/-}$ and *Pah*$^{R408W/R408W}$ animals (fig. S20A, B). A combinational treatment of GalNAc-*HULC* and sapropterin showed a cooperative effect in reducing the blood Phe concentrations of *Pah*$^{R408W/R408W}$ animals (Fig. 5J). These findings suggested that a supply of GalNAc-*HULC* mimics may enhance the association between PAH and BH_4, thereby improving the therapeutic effect of sapropterin. The *HULC* mimics showed no detectable effects on the body weight, liver function, and kidney function of *Pair*$^{-/-}$ and *Pah*$^{R408W/R408W}$ mice (fig. S20C-G). Taken together, our data suggested that application of *HULC* mimics could enhance the enzymatic activities of certain mutated PAH proteins, offering a potential intervention for patients with PKU.

DISCUSSION

PKU has long been considered a single-gene disease with an autosomal recessive inheritance pattern, in which mutations at the PAH locus lead to impaired enzymatic function and contribute to a hyperphenylalaninemia metabolic phenotype, subsequently resulting in cognitive phenotypes such as mental retardation. Recent genotype-phenotype correlation

analyses of PAH mutations suggested a substantial discrepancy between genotypes and their predicted metabolic or cognitive phenotypes (8, 9, 39, 40). Variations in genomic, epigenomic, transcriptomic, proteomic, and metabolic systems could all contribute to the PAH deficiency of patients with PKU (41). It has been shown that patients, even siblings sharing identical mutant PAH genotypes, can have greatly differing cognitive and metabolic phenotypes (10). This highly suggests that a PAH genotype cannot consistently and reliably predict the monogenic phenotype (10). Consistent with this notion, the recent discovery of biallelic mutations in the *DNAJC12* gene in patients with high Phe concentrations suggested the possibility that non-PAH genes contribute to this disease (42). Hence, factors modulating PAH expression, PAH protein stability, BH_4 biogenesis, and PAH enzymatic activity may play important roles in maintaining PAH enzymatic proficiency. More than 95% of genetic mutations occur at the non-coding regions of the human genome, yet the functional importance of lncRNAs in human genetic diseases remains elusive. We demonstrated that depletion of the lncRNA *Pair* leads to phenotypes that model human PKU, including hypopigmentation, growth retardation, seizures, and neuronal loss. Whether the presence of the *HULC* genomic variants contributes to the severity of PKU requires further investigation, given the heterogeneity of *PAH* genotypes and the lack of knowledge of *HULC* genomic variant statuses in current data collections.

The nature of genetic diseases has resulted in the development of compensatory lifestyle considerations that have marked, potentially negative impacts on the overall quality of life of patients. BH_4 and its synthetic analog, sapropterin, have been used as enzyme cofactors for a subset of patients with PKU, particularly those with mild PKU (43-45). However, patients with PKU and Phe >1200 μmol/L are less likely to respond to BH_4 supplementation (46). PKU could potentially be genetically corrected via adenoviral or related vectors (47); nevertheless, their specificity, stability, and the adenoviral genes responsible for immune response require further research studies (48). Enzyme replacement or substitution therapy, in which PAH or PAL fusion proteins (Palynziq) are intravenously injected, leads to decreased plasma Phe (15, 16, 49, 50) but also potentially concerning immune response-related symptoms experienced by a percentage of patients with PKU (51, 52). Hence, development of additional therapeutic strategies is needed, such as an approach using liver-enriched lncRNA mimics that specifically enhance PAH-substrate and PAH-cofactor binding affinities and enhance enzymatic activity in vitro and in vivo.

The *HULC* mimics we developed target lncRNAs as a therapeutic strategy against a human inherited metabolic disorder. Considering lncRNA mimics as a potential therapeutic option has the following advantages: 1) Flexibility in synthesizing lncRNA mimics according to the target sequence and tagging with organ-targeting peptides for tissue-specific distribution. The target sequence could be designed to improve the enzymatic activity of key enzyme(s) in other metabolic disorders. 2) Profound stability in vivo due to the modified covalent bonds between nucleotides, which are resistant to DNase and RNase. 3) Liver-enriched tissue distribution assisted by the GalNAc-tag. 4) Low potential for organ toxicity, as supported by our observation that administration of lncRNA mimics resulted in no detectable effects on liver or kidney function. It is likely that combining *HULC* mimics with current dietary restrictions, sapropterin supplements, and enzyme replacement or

substitution therapies would further improve patient outcomes, particularly for patients with severe PKU.

Low conservation between mouse and human lncRNAs has greatly hindered the discovery of disease-driving lncRNAs and their roles in human diseases; however, it is possible that RNA "domains" might be conserved between human and mouse lncRNAs. *HULC* and *Pair* associate with PAH protein at the N-terminal regulatory domain. The presence of *HULC* could rescue PAH enzymatic activity in *Pair*-deficient cells and vice versa, shedding light on the therapeutic potential of the vast, currently-unconsidered, and poorly conserved lncRNAs.

PAH is self-regulated by the N-terminal regulatory domain. The crystal structure of PAH suggests that the C-terminal helix is required for the tetramerization of PAH. *HULC/Pair* is unlikely to affect the oligomerization of PAH, as supported by the finding that *HULC* mimics showed no notable effect on the enzymatic activities of PAH Y414C and Y417H. The N-terminal regulatory domain can undergo a conformational change to switch between "open" and "closed" states (30). We reason that the binding of *HULC* to the N-terminal regulatory domain of PAH may stabilize the PAH-Phe-*HULC* complex, leading the PAH protein to operate in the "open" state and allowing Phe and BH_4 to access the enzyme. Our finding is consistent with the previous hypothesis that an "unknown" allosteric factor stabilizes the interaction between PAH and Phe, serving as a therapeutic chaperone (26). Taken together, our findings suggested the functional importance of lncRNAs in phenylalanine metabolism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This project was partially supported by the Human Stem Cell Core (HSCC) at Baylor College of Medicine. Human iPSCs was generated at the HSCC by Dr. Jean J. Kim.

Funding:

The Proteomics Facility was supported in part by NIH grant number 1S10OD012304-01 (D.H.). This project was partially supported by University of Houston Seq-N-Edit Core with funding from UH Division of Research; UH College of NSM and Department of Biology & Biochemistry; NRUF MINOR CORE 17 Grant to P. Gunaratne; UH Small Core Equipment Program Grant to P. Gunaratne. We thank the core facilities at BCM: Metabolomics Core, (NIH P30CA125123), CPRIT Proteomics and Metabolomics Core Facility (N.P.) (RP170005), and Dan L. Duncan Cancer Center. This project was also supported by the grant of Cancer Prevention & Research Institute of Texas (CPRIT) RP150085 and RP190570 (L.H.). Z.T. and S. Z. were partially supported by CPRIT RP170333, NIH R01CA225955, and NSF CHE-1411859. This work was supported by C.R.L.'s NIH (1R01CA218025-01, 1R01CA231011-01), CPRIT (RP180259), and Department of Defense (DoD) (BC180196), as well as L.Q.Y.'s NIH (1R01CA218036-01), CPRIT (RP200423), DoD (BC181384), the American Association for Cancer Research Grant (20-60-51 Yang), Andrew Sabin Family Foundation Fellows award and MD Anderson Cancer Center Institutional Research Grant.

Data availability:

The raw lncRNA array data for this manuscript are available at GEO under the accession number GSE120207. All other data are available in the main text or the supplementary materials. Due to the COVID-19 pandemic, our genetically modified mouse colonies died

out, and we are in the process of reestablishing our homozygous mouse populations. The skin fibroblasts and hiPSCs were obtained from the Coriell Institute with MTA and are subjected to restrictions on redistribution and sharing. Request of these cells shall be directly obtained from Coriell Institute. All other research materials are available from the corresponding author upon reasonable request.

References and notes:

1. Folling I, The discovery of phenylketonuria. Acta Paediatr Suppl 407, 4–10 (1994). [PubMed: 7766954]
2. Blau N, Genetics of Phenylketonuria: Then and Now. Hum Mutat 37, 508–515 (2016). [PubMed: 26919687]
3. Routine neonatal screening for phenylketonuria in the United Kingdom 1964–78. Medical Research Council Steering Committee for the MRC/DHSS Phenylketonuria Register. Br Med J (Clin Res Ed) 282, 1680–1684 (1981).
4. van Spronsen FJ et al. , Key European guidelines for the diagnosis and management of patients with phenylketonuria. Lancet Diabetes Endocrinol 5, 743–756 (2017). [PubMed: 28082082]
5. Camp KM et al. , Phenylketonuria Scientific Review Conference: state of the science and future research needs. Mol Genet Metab 112, 87–122 (2014). [PubMed: 24667081]
6. Martynyuk AE, van Spronsen FJ, Van der Zee EA, Animal models of brain dysfunction in phenylketonuria. Mol Genet Metab 99 Suppl 1, S100–105 (2010). [PubMed: 20123463]
7. Pey AL, Desviat LR, Gamez A, Ugarte M, Perez B, Phenylketonuria: genotype-phenotype correlations based on expression analysis of structural and functional mutations in PAH. Hum Mutat 21, 370–378 (2003). [PubMed: 12655546]
8. Kayaalp E et al. , Human phenylalanine hydroxylase mutations and hyperphenylalaninemia phenotypes: a metanalysis of genotype-phenotype correlations. Am J Hum Genet 61, 1309–1317 (1997). [PubMed: 9399896]
9. Murad H, Dabboul A, Moassas F, Alasmar D, Al-Achkar W, Mutation spectrum of phenylketonuria in Syrian population: genotype-phenotype correlation. Gene 528, 241–247 (2013). [PubMed: 23856132]
10. Scriver CR, Waters PJ, Monogenic traits are not simple: lessons from phenylketonuria. Trends in genetics : TIG 15, 267–272 (1999). [PubMed: 10390625]
11. Anikster Y et al. , Biallelic Mutations in DNAJC12 Cause Hyperphenylalaninemia, Dystonia, and Intellectual Disability. Am J Hum Genet 100, 257–266 (2017). [PubMed: 28132689]
12. Penrose L, Quastel JH, Metabolic studies in phenylketonuria. The Biochemical journal 31, 266–274 (1937). [PubMed: 16746333]
13. Kure S et al. , Tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. J Pediatr 135, 375–378 (1999). [PubMed: 10484807]
14. Karacic I et al. , Genotype-predicted tetrahydrobiopterin (BH4)-responsiveness and molecular genetics in Croatian patients with phenylalanine hydroxylase (PAH) deficiency. Mol Genet Metab 97, 165–171 (2009). [PubMed: 19394257]
15. Eavri R, Lorberboum-Galski H, A novel approach for enzyme replacement therapy. The use of phenylalanine hydroxylase-based fusion proteins for the treatment of phenylketonuria. The Journal of biological chemistry 282, 23402–23409 (2007). [PubMed: 17565982]
16. Hoskins JA et al. , Enzymatic control of phenylalanine intake in phenylketonuria. Lancet 1, 392–394 (1980). [PubMed: 6101846]
17. Rebuffat A, Harding CO, Ding Z, Thony B, Comparison of adeno-associated virus pseudotype 1, 2, and 8 vectors administered by intramuscular injection in the treatment of murine phenylketonuria. Hum Gene Ther 21, 463–477 (2010). [PubMed: 19916803]
18. Guo JC et al. , CNIT: a fast and accurate web tool for identifying protein-coding and long non-coding transcripts based on intrinsic sequence composition. Nucleic acids research 47, W516–W522 (2019). [PubMed: 31147700]

19. Wang J et al. , Multiple functions of the RNA-binding protein HuR in cancer progression, treatment responses and prognosis. Int J Mol Sci 14, 10015–10041 (2013). [PubMed: 23665903]
20. Panzitt K et al. , Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. Gastroenterology 132, 330–342 (2007). [PubMed: 17241883]
21. Collins DW, Jukes TH, Rates of transition and transversion in coding sequences since the human-rodent divergence. Genomics 20, 386–396 (1994). [PubMed: 8034311]
22. Watters KE, Yu AM, Strobel EJ, Settle AH, Lucks JB, Characterizing RNA structures in vitro and in vivo with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). Methods 103, 34–48 (2016). [PubMed: 27064082]
23. Kubota M, Tran C, Spitale RC, Progress and challenges for chemical probing of RNA structure inside living cells. Nat Chem Biol 11, 933–941 (2015). [PubMed: 26575240]
24. Thony B, Ding Z, Martinez A, Tetrahydrobiopterin protects phenylalanine hydroxylase activity in vivo: implications for tetrahydrobiopterin-responsive hyperphenylalaninemia. FEBS letters 577, 507–511 (2004). [PubMed: 15556637]
25. Erlandsen H, Stevens RC, A structural hypothesis for BH4 responsiveness in patients with mild forms of hyperphenylalaninaemia and phenylketonuria. J Inherit Metab Dis 24, 213–230 (2001). [PubMed: 11405341]
26. Jaffe EK, New protein structures provide an updated understanding of phenylketonuria. Mol Genet Metab 121, 289–296 (2017). [PubMed: 28645531]
27. Zhang S, Fitzpatrick PF, Identification of the Allosteric Site for Phenylalanine in Rat Phenylalanine Hydroxylase. The Journal of biological chemistry 291, 7418–7425 (2016). [PubMed: 26823465]
28. Frederick CB, Dooley KL, Kodell RL, Sheldon WG, Kadlubar FF, The effect of lifetime sodium saccharin dosing on mice initiated with the carcinogen 2-acetylaminofluorene. Fundam Appl Toxicol 12, 346–357 (1989). [PubMed: 2714533]
29. Gjetting T, Petersen M, Guldberg P, Guttler F, Missense mutations in the N-terminal domain of human phenylalanine hydroxylase interfere with binding of regulatory phenylalanine. Am J Hum Genet 68, 1353–1360 (2001). [PubMed: 11326337]
30. Carluccio C, Fraternali F, Salvatore F, Fornili A, Zagari A, Structural features of the regulatory ACT domain of phenylalanine hydroxylase. PloS one 8, e79482 (2013). [PubMed: 24244510]
31. Lin A et al. , The LINK-A lncRNA interacts with PtdIns(3,4,5)P3 to hyperactivate AKT and confer resistance to AKT inhibitors. Nature cell biology 19, 238–251 (2017). [PubMed: 28218907]
32. Feng Y et al. , Global analysis of protein structural changes in complex proteomes. Nature biotechnology 32, 1036–1044 (2014).
33. Waters PJ, How PAH gene mutations cause hyper-phenylalaninemia and why mechanism matters: insights from in vitro expression. Hum Mutat 21, 357–369 (2003). [PubMed: 12655545]
34. Gamez A, Perez B, Ugarte M, Desviat LR, Expression analysis of phenylketonuria mutations. Effect on folding and stability of the phenylalanine hydroxylase protein. The Journal of biological chemistry 275, 29737–29742 (2000). [PubMed: 10875932]
35. Zschocke J, Graham CA, Carson DJ, Nevin NC, Phenylketonuria mutation analysis in Northern Ireland: a rapid stepwise approach. Am J Hum Genet 57, 1311–1317 (1995). [PubMed: 8533759]
36. O'Neill CA et al. , Molecular analysis of PKU in Ireland. Acta Paediatr Suppl 407, 43–44 (1994). [PubMed: 7766956]
37. Huang Y, Preclinical and Clinical Advances of GalNAc-Decorated Nucleic Acid Therapeutics. Mol Ther Nucleic Acids 6, 116–132 (2017). [PubMed: 28325278]
38. Sanford M, Keating GM, Sapropterin: a review of its use in the treatment of primary hyperphenylalaninaemia. Drugs 69, 461–476 (2009). [PubMed: 19323589]
39. Santos LL et al. , Variations in genotype-phenotype correlations in phenylketonuria patients. Genet Mol Res 9, 1–8 (2010). [PubMed: 20082265]
40. Bercovich D et al. , Genotype-phenotype correlations analysis of mutations in the phenylalanine hydroxylase (PAH) gene. J Hum Genet 53, 407–418 (2008). [PubMed: 18299955]
41. Garbade SF et al. , Allelic phenotype values: a model for genotype-based phenotype prediction in phenylketonuria. Genet Med 21, 580–590 (2019). [PubMed: 29997390]

42. Blau N, Martinez A, Hoffmann GF, Thony B, DNAJC12 deficiency: A new strategy in the diagnosis of hyperphenylalaninemias. Mol Genet Metab 123, 1–5 (2018). [PubMed: 29174366]
43. Schaub J et al. , Tetrahydrobiopterin therapy of atypical phenylketonuria due to defective dihydrobiopterin biosynthesis. Arch Dis Child 53, 674–676 (1978). [PubMed: 708106]
44. Blau N, Erlandsen H, The metabolic and molecular bases of tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. Mol Genet Metab 82, 101–111 (2004). [PubMed: 15171997]
45. Trefz FK, Scheible D, Frauendienst-Egger G, Korall H, Blau N, Long-term treatment of patients with mild and classical phenylketonuria by tetrahydrobiopterin. Mol Genet Metab 86 Suppl 1, S75–80 (2005). [PubMed: 16242984]
46. Gramer G, Burgard P, Garbade SF, Lindner M, Effects and clinical significance of tetrahydrobiopterin supplementation in phenylalanine hydroxylase-deficient hyperphenylalaninaemia. J Inherit Metab Dis 30, 556–562 (2007). [PubMed: 17680344]
47. Eisensmith RC, Woo SL, Gene therapy for phenylketonuria. Acta Paediatr Suppl 407, 124–129 (1994). [PubMed: 7766948]
48. Ding Z, Harding CO, Thony B, State-of-the-art 2003 on PKU gene therapy. Mol Genet Metab 81, 3–8 (2004). [PubMed: 14728985]
49. Sarkissian CN et al. , A different approach to treatment of phenylketonuria: phenylalanine degradation with recombinant phenylalanine ammonia lyase. Proceedings of the National Academy of Sciences of the United States of America 96, 2339–2344 (1999). [PubMed: 10051643]
50. Hoskins JA, Gray J, Phenylalanine ammonia lyase in the management of phenylketonuria: the relationship between ingested cinnamate and urinary hippurate in humans. Res Commun Chem Pathol Pharmacol 35, 275–282 (1982). [PubMed: 7071414]
51. Longo N et al. , Long-term safety and efficacy of pegvaliase for the treatment of phenylketonuria in adults: combined phase 2 outcomes through PAL-003 extension study. Orphanet J Rare Dis 13, 108 (2018). [PubMed: 29973227]
52. Mahan KC, Gandhi MA, Anand S, Pegvaliase: a novel treatment option for adults with phenylketonuria. Curr Med Res Opin 35, 647–651 (2019). [PubMed: 30247930]
53. Harding CO et al. , Complete correction of hyperphenylalaninemia following liver-directed, recombinant AAV2/8 vector-mediated gene therapy in murine phenylketonuria. Gene Ther 13, 457–462 (2006). [PubMed: 16319949]
54. Smola MJ et al. , SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. Proceedings of the National Academy of Sciences of the United States of America 113, 10322–10327 (2016). [PubMed: 27578869]
55. Kwok CK, Ding Y, Tang Y, Assmann SM, Bevilacqua PC, Determination of in vivo RNA structure in low-abundance transcripts. Nat Commun 4, 2971 (2013). [PubMed: 24336128]
56. Qi HH et al. , Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development. Nature 466, 503–507 (2010). [PubMed: 20622853]
57. Fusaki N, Ban H, Nishiyama A, Saeki K, Hasegawa M, Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proc Jpn Acad Ser B Phys Biol Sci 85, 348–362 (2009).
58. Sanjana NE, Shalem O, Zhang F, Improved vectors and genome-wide libraries for CRISPR screening. Nature methods 11, 783–784 (2014). [PubMed: 25075903]
59. Lin A et al. , The LINK-A lncRNA activates normoxic HIF1alpha signalling in triple-negative breast cancer. Nature cell biology 18, 213–224 (2016). [PubMed: 26751287]
60. Ule J, Jensen K, Mele A, Darnell RB, CLIP: a method for identifying protein-RNA interaction sites in living cells. Methods 37, 376–386 (2005). [PubMed: 16314267]
61. Yang L et al. , ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. Cell 147, 773–788 (2011). [PubMed: 22078878]
62. Fey A et al. , Establishment of a real-time PCR-based approach for accurate quantification of bacterial RNA targets in water, using Salmonella as a model organism. Appl Environ Microbiol 70, 3618–3623 (2004). [PubMed: 15184165]

63. Xing Z et al. , lncRNA Directs Cooperative Epigenetic Regulation Downstream of Chemokine Signals. Cell 159, 1110–1125 (2014). [PubMed: 25416949]

64. Yang L et al. , lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature 500, 598–602 (2013). [PubMed: 23945587]

65. Shiman R, Xia T, Hill MA, Gray DW, Regulation of rat liver phenylalanine hydroxylase. II. Substrate binding and the role of activation in the control of enzymatic activity. The Journal of biological chemistry 269, 24647–24656 (1994). [PubMed: 7929136]

66. Gersting SW et al. , Activation of phenylalanine hydroxylase induces positive cooperativity toward the natural cofactor. The Journal of biological chemistry 285, 30686–30697 (2010). [PubMed: 20667834]

67. Milstien S, Kaufman S, Studies on the phenylalanine hydroxylase system in liver slices. The Journal of biological chemistry 250, 4777–4781 (1975). [PubMed: 124732]

68. Sali A, Blundell TL, Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234, 779–815 (1993). [PubMed: 8254673]

69. Parisien M, Major F, The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. Nature 452, 51–55 (2008). [PubMed: 18322526]

70. Huang Y, Liu S, Guo D, Li L, Xiao Y, A novel protocol for three-dimensional structure prediction of RNA-protein complexes. Sci Rep 3, 1887 (2013). [PubMed: 23712416]

71. Yan Y, Zhang D, Zhou P, Li B, Huang SY, HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. Nucleic acids research 45, W365–W373 (2017). [PubMed: 28521030]

72. Tuszynska I, Magnus M, Jonak K, Dawson W, Bujnicki JM, NPDock: a web server for protein-nucleic acid docking. Nucleic acids research 43, W425–430 (2015). [PubMed: 25977296]

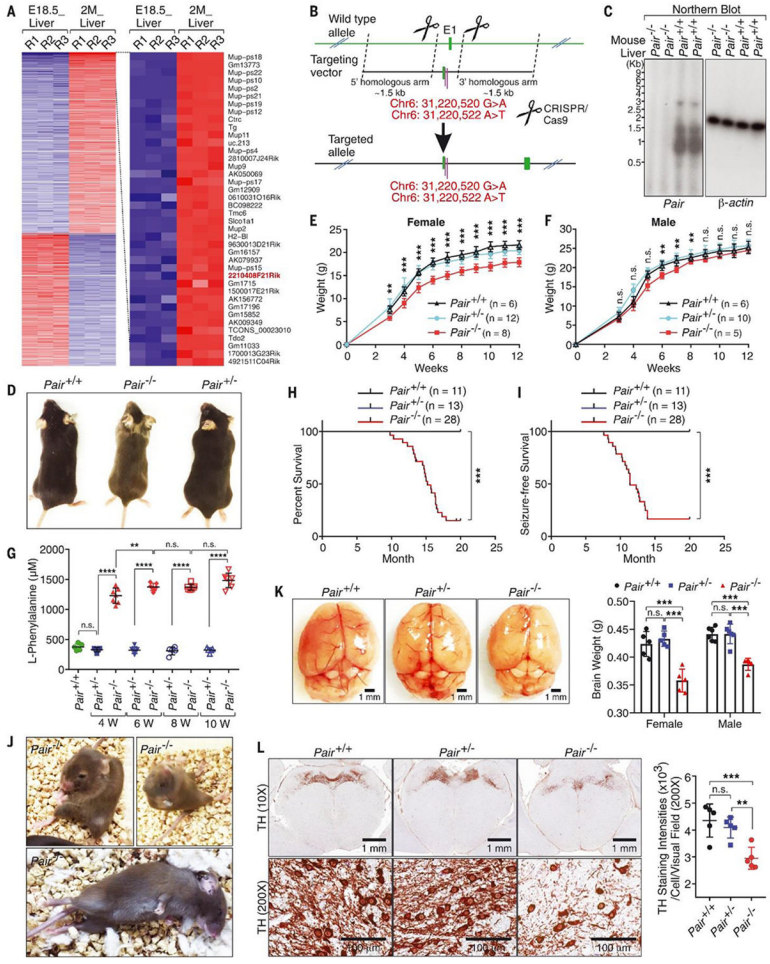

Fig. 1. *Pair*$^{-/-}$ mouse mimics human PKU disease.

(**A**) lncRNA profiling of livers from E18.5 embryos and from 2-month-old adult mice (n = 3). 2210408F21Rik is highlighted in red. (**B**) Schematic of using CRISPR-Cas9 to generate a *Pair*$^{-/-}$ mouse model. (**C**) Northern blot detection of the expression of *Pair* in the indicated mouse livers. β-actin was used as a loading control. (**D**) Representative images of 12-month-old *Pair* mice. (**E** and **F**) Comparison of body weights of the indicated female (E) and male (F) mice at the age of 3 to 12 weeks. Data are shown as mean ± SD, one-way ANOVA. (**G**) Blood Phe concentrations in the indicated mice was tested every 2 weeks starting from 4 weeks of age (n = 5 *Pair*$^{+/+}$, n = 5 *Pair*$^{-/-}$, n = 7 *Pair*$^{+/-}$ animals). Data are shown as mean ± SD, one-way ANOVA. (**H**) Cumulative survival curve of cohorts of indicated littermates (log-rank test). (**I**) Cumulative seizure-free survival curve of cohorts of indicated littermates (log-rank test). (**J**) Representative images of *Pair*$^{-/-}$ mice experiencing seizures. (**K**) Left: representative images of brains from indicated mice at the age of 12 months (n = 5). Right: scatter plots represent brain weight quantification for both female and male indicated littermates at the age of 12 months (n = 5). Mean ± SD, one-way ANOVA. (**L**) Coronal sections of adult mouse brains subjected to immunohistochemical staining and quantitative data for TH^{+} neurons in the substantia nigra compact/ventral tegmental area of the indicated mice at the age of 12 months (n = 5). Mean ± SD, one-way ANOVA. Top row,

10× magnification; bottom row, 200× magnification. n.s., not significant at $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$.

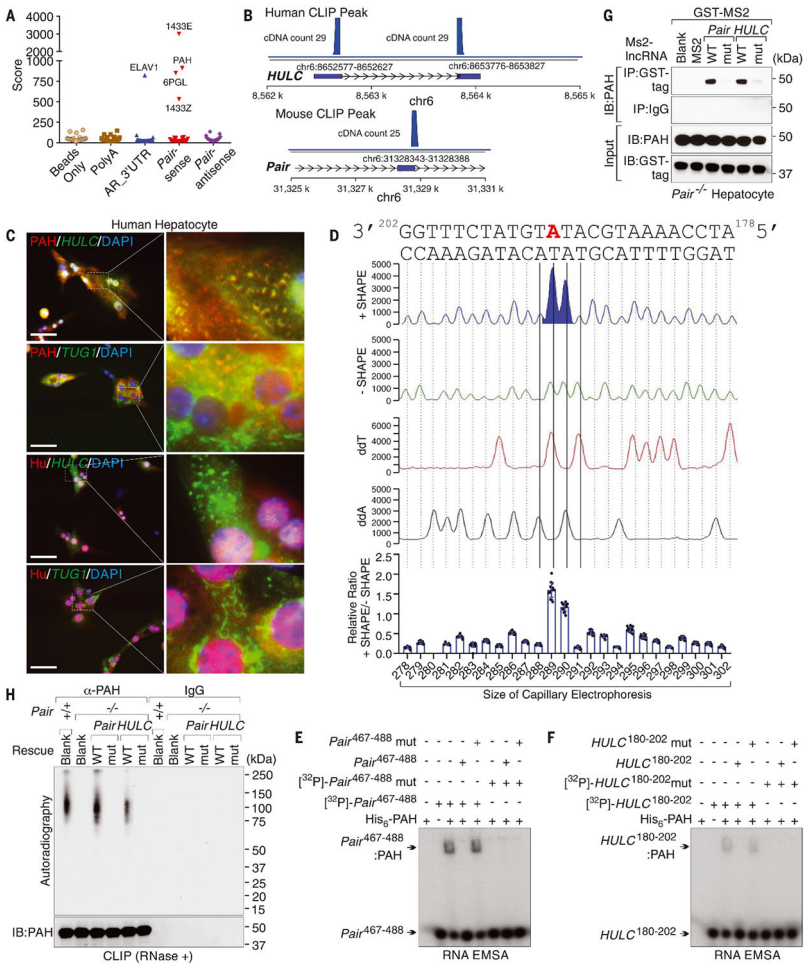

Fig. 2. *Pair* and *HULC* associate with PAH.

(**A**) Protein candidates interacting with *Pair* were revealed by LC-MS. The *x*-axis indicates different experimental groups. (**B**) PAH-binding sites along *HULC* and *Pair* identified by CLIP assay. The figure represents the read coverage along *HULC* and *Pair* transcripts obtained by mapping Sanger sequencing reads to the representative transcript. cDNA counts are shown. (**C**) Representative images of immunolabeling with fluorescent in situ hybridization to simultaneously detect colocalization of the indicated RNA and protein molecules in human primary hepatocytes. *TUG1* and Hu were used as RNA and protein controls, respectively. Scale bars, 50 μm. (**D**) Resolution of SHAPE reactivity by capillary electrophoresis. Four-color electropherograms for the products of one multicolor run by 6-FAM-labeled NMIA modified RNA (+SHAPE), VIC-labeled control RNA (−SHAPE) of *HULC* nucleotides 178 to 202, NED-labeled ddA, and a PET-labeled ddT are shown. Bottom panel: normalized SHAPE reactivity of *HULC* nucleotides 178 to 202. *x*-axis indicates the sizes of capillary electrophoresis. Data are shown as mean ± SD of n = 15 independent experiments. (**E** and **F**) EMSA using recombinant His-tagged PAH and [γ-32P]–labeled *Pair* nucleotides 467 to 488 (E) or *HULC* nucleotides 180 to 202 (F) WT or mutant oligonucleotides. Unlabeled *Pair* or *HULC* WT or mutant RNA oligonucleotides were included as competitors. (**G**) MS2-TRAP assay using *Pair*$^{-/-}$ hepatocytes expressing

indicated plasmids performed by immunoblotting using the indicated antibodies. (**H**) Immunoblotting (bottom) or autoradiography (top) of CLIP assay using the indicated hepatocytes expressing the indicated plasmids.

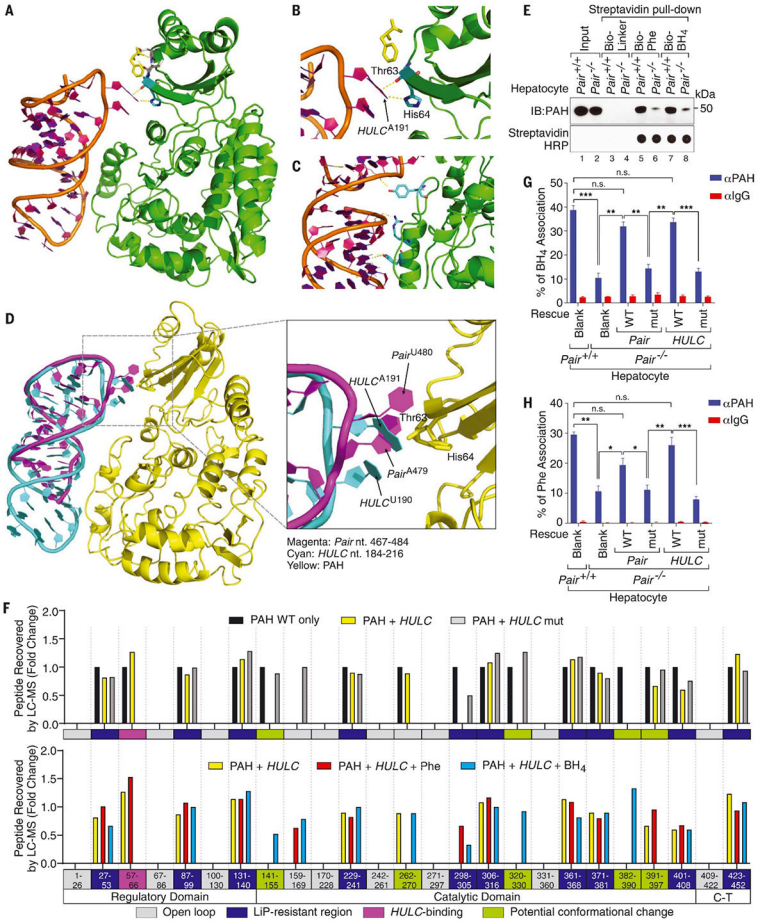

Fig. 3. *HULC* and *Pair* modulate the enzymatic activities of PAH.

(**A**) *HULC* binds to PAH to stabilize the allosteric Phe-induced open conformation of PAH. (**B**) Magnified view of *HULC*'s interaction with the regulatory domain of PAH. *HULC* A^{191} forms hydrogen bonds with His^{64} and polar interaction with Thr^{63}. Phe forms stacking interaction with *HULC* A^{191}. Color representation in both (A) and (B): green indicates PAH; orange and magenta indicate *HULC*; yellow indicates allosteric Phe; and cyan indicates Thr^{63} and His^{64}. (**C**) *HULC* A^{195} forms a hydrogen bond with Tyr^{166}; *HULC* A^{214} forms a hydrogen bond with Arg^{157}; *HULC* G^{202} forms a hydrogen bond with Tyr^{154}. Green indicates PAH; orange and magenta indicate *HULC*; and Cyan indicates Tyr^{154}, Arg^{157}, and Tyr^{166}. (**D**) Superimposition of the three-dimensional structures of *HULC* nucleotides 184 to 216 (cyan) and *Pair* nucleotides 467 to 484 (magenta) after docking to PAH (yellow). *Pair* A^{479}, U^{480} and *HULC* U^{190}, A^{191} are indicated by arrows. (**E**) Streptavidin pull-down using Bio-Phe/Bio-BH_4 followed by immunoblotting detection using anti-PAH antibody in the indicated hepatocytes. (**F**) Fold change of peptides recovered from LiP-MS. Top panel: comparison of PAH WT only, PAH incubated with *HULC*, and PAH incubated with a *HULC* mutant sequence. Bottom panel: comparison of PAH incubated with *HULC*, PAH incubated with *HULC* and Phe, and PAH incubated with *HULC* and BH_4. The *x*-axis indicates the amino acid position of full-length PAH; the *y*-axis indicates the fold change of peptide recovery number. (**G** and **H**) ELISA measurement of the percentage of PAH-associated BH_4 (G) or PAH-associated Phe (H) in the indicated hepatocytes expressing the indicated mimics.

Data are shown as mean ± SEM of $n = 5$ independent experiments, one-way ANOVA. n.s., not significant at $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

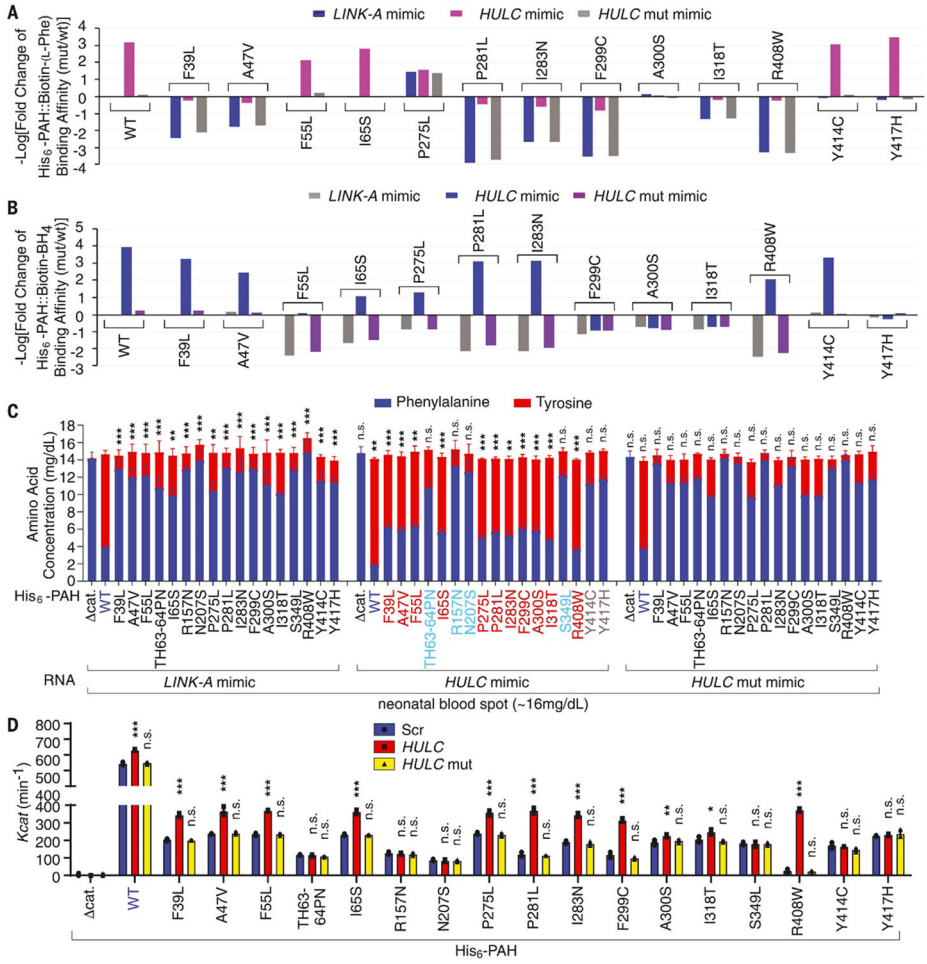

Fig. 4. *HULC* mimics facilitate PAH-Phe and PAH-BH_4 interactions.

(**A** and **B**) Log_2 of relative fold change of His-PAH WT/mutants and Biotin-Phe (A) or His-PAH WT/mutants and biotin-BH_4 (B) binding affinity in the presence of the indicated lncRNA mimics. The fold change was normalized using His-PAH WT in the presence of the *LINK-A* mimic. (**C**) Phe and Tyr concentrations in neonatal blood spots in the presence of His-tagged PAH WT or indicated mutants and the indicated lncRNA mimics. Data are shown as mean ± SD of $n = 3$ independent experiments, Student's t test. The lncRNA mimic representing *LINK-A* nucleotides 1100 to 1117 was included as a negative control. (**D**) Determination of k_{cat} of recombinant PAH WT or indicated mutant proteins in the presence of the indicated mimics. Data are shown as the mean ± SD of $n = 3$ independent experiments, one-way ANOVA. n.s., not significant at $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

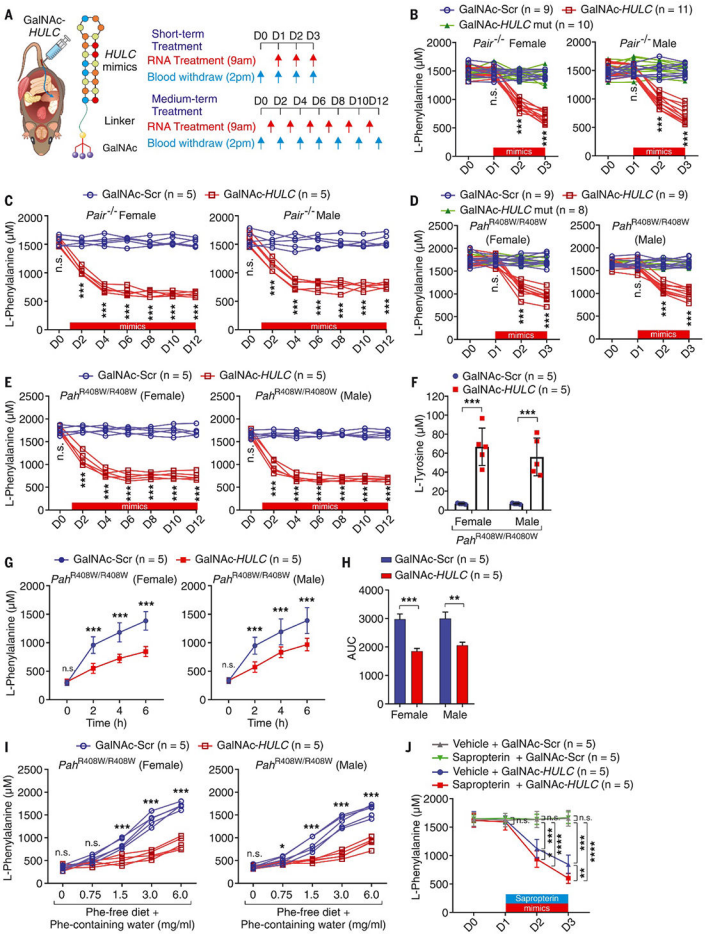

Fig. 5. GalNAc-*HULC* lncRNA mimics alleviate PKU symptoms.

(**A**) Graphic illustration of GalNAc-tagged *HULC* mimics (left) and the treatment schedules (right). (**B**) Blood Phe concentrations were monitored every day for short-term treatment in female and male *Pair*$^{-/-}$ mice with the indicated mimics. Data were analyzed with Student's *t* test. (**C**) Blood Phe concentrations were monitored every other day for medium-term treatment in female and male *Pair*$^{-/-}$ mice treated with the indicated mimics. Data were analyzed with Student's *t* test. (**D**) Blood Phe concentrations were monitored every day for short-term treatment in female or male *Pah*$^{R408W/R408W}$ mice treated with the indicated mimics. Data were analyzed with Student's *t* test. (**E**) Blood Phe concentrations were monitored every other day for medium-term treatment in female or male *Pah*$^{R408W/R408W}$ mice treated with the indicated mimics. Data were analyzed with Student's t test. (**F**) Measurement of blood tyrosine concentrations for medium-term treatment in female and male *Pah*$^{R408W/R408W}$ mice treated with the indicated mimics, Data were analyzed with Student's *t* test. (**G** and **H**) Phe clearance test (G) or area under the curve (AUC) (H) of female or male *Pah*$^{R408W/R408W}$ mice subjected to a Phe-free diet and pretreatment with the indicated mimics. Data are shown as mean ± SD, Student's *t* test. (**I**) Phe tolerance test of female or male *Pah*$^{R408W/R408W}$ mice subjected to a Phe-free diet and pretreatment with indicated mimics for 3 days, followed by water containing 0, 0.75, 1.5, 3.0, or 6.0 mg/ml of Phe (with the dose increasing every 2 days). Data were analyzed with Student's

t test. (**J**) Blood Phe concentrations were monitored every day for short-term treatment in *Pah*$^{R408W/R408W}$ mice subjected to sapropterin alone or in combination with the indicated mimics. Data are shown as mean ± SD, one-way ANOVA. n.s., not significant at $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$.