The extent of translational control of gene expression in mammalian tissues remains largely unknown. Here we perform genome-wide RNA sequencing and ribosome profiling in heart and liver tissues to investigate strain-specific translational regulation in the spontaneously hypertensive rat (SHR/Ola). For the most part, transcriptional variation is equally apparent at the translational level and there is limited evidence of translational buffering. Remarkably, we observe hundreds of strain-specific differences in translation, almost doubling the number of differentially expressed genes. The integration of genetic, transcriptional and translational data sets reveals distinct signatures in 3'UTR variation, RNA-binding protein motifs and miRNA expression associated with translational regulation of gene expression. We show that a large number of genes associated with heart and liver traits in human genome-wide association studies are primarily translationally regulated. Capturing interindividual differences in the translated genome will lead to new insights into the genes and regulatory pathways underlying disease phenotypes.
heritable RNA expression phenotypes have been studied extensively\(^1\),\(^2\), but the contribution of translational regulation to natural phenotypic variation is largely unknown. Several genetic disorders are believed to be caused by mutations that affect protein translation\(^3\), but so far genome-wide translation has not been studied in complex disease models. It remains a matter of debate how well transcript abundances explain protein levels\(^4\),\(^5\), to what extent RNA expression differences are mirrored by the proteome\(^6\)–\(^10\) and how many genes are under translational control in mammalian tissues.

To obtain quantitative measurements of translation, we adapted ribosome profiling procedures\(^1\),\(^1\)\(^1\)\(^2\) to reliably isolate and sequence ribosome-protected RNA fragments (RPFs) from whole-tissue samples (Ribo-seq). We then investigated translational regulation in the heart and liver of two inbred rat strains, the BN-Lx reference strain \((n = 5)\) and the SHR/Ola \((n = 5)\), a widely studied model of cardiovascular and metabolic disease traits\(^5\)–\(^13\)–\(^17\).

Ribosome profiling of disease tissue reveals widespread interindividual regulation of translation that shapes the molecular landscape of the cardiometabolic phenotype of the SHR/Ola rat. Ribo-seq data are a better proxy for protein levels compared with RNA-seq data and expose almost twice as many differences in gene expression between strains. The integration of transcriptional and post-transcriptional control of gene expression identifies distinct molecular signatures and reveals the contribution of genetic variation, microRNAs (miRNAs) and RNA-binding proteins (RBPs) to translational regulation. Many genes associated with cardiac and hepatic traits in humans are translationally, and not transcriptionally, regulated in disease tissue.

**Results**

RNA transcription and translation in rat heart and liver. The Ribo-seq libraries allowed us to monitor translating ribosomes at sub-codon resolution. We observed codon periodicity (Fig. 1a) and other attributes (Supplementary Fig. 1) that are characteristic for \(in\) \(vivo\) snapshots of active translation. To distinguish translational from transcriptional regulation of gene expression, we integrated two distinct RNA-seq experiments: one polyA-selected data set (RNA-seq1) and one ribosomal RNA (rRNA)-depleted data set (RNA-seq2), which was generated in parallel to ribosome

---

Figure 1 | RNA-seq and Ribo-seq of cardiac and hepatic tissue. (a) Start positions of RNA fragments protected by ribosomes (total of 10 cardiac Ribo-seq libraries, each plotted individually). Ribosomes are located at the start codon and then move sequentially along the triplet code to translate the genetic message (for liver see Supplementary Fig. 1d). Once the ribosomal site occupied by aminoacyl-tRNA (A site) reaches the stop codon, ribosomes are released from the transcripts. For each library, the fraction of reads covering each frame is shown. The majority of ribosomes are located on the codons of the open-reading frame of protein-coding genes. (b,c) Differential expression between BN-Lx and SHR/Ola rat strains (five biological replicates each) in (b) heart and (c) liver. Fold changes of RPF and RNA abundance correlate on a genome-wide scale (Pearson’s, \(P < 1e - 4\)). While the majority of genes showed no significant inter-strain differences (grey dots, \(n_{\text{heart}} = 9,856, n_{\text{liver}} = 9,586\), hundreds of genes were differentially expressed (FDR \(\leq 0.01\)) at the level of transcription only (blue dots, RNA\(_{\text{trans}}\)), translation only (red dots, RIBO\(_{\text{trans}}\)) or both (black dots, RNA + RIBO). (d) Pie charts and tables indicate for each tissue the fractions and absolute numbers of RIBO\(_{\text{trans}}\), RNA\(_{\text{trans}}\) and RNA + RIBO genes of differentially expressed genes, respectively.
profiling from the identical tissue lysate. To adjust for variation in sequencing depth across technologies, mappability and fragment size, we matched the read length and the number of uniquely aligning reads located within exons for all data sets (Supplementary Fig. 2). To exclude strain-specific mapping biases, we used the SHR/Ola and BN-Lx genome sequence information\(^{18-20}\) and mapped the reads of each strain to its respective genome. We defined significant strain-specific differences in RPF and RNA expression levels (false discovery rate (FDR) \(\leq 0.01\)) using DESeq2 (ref. 21) (Supplementary Fig. 3).

We compared differences in messenger RNA (mRNA) levels to variation in RPF abundances between strains (Fig. 1b–d; Supplementary Fig. 4). Three classes of genes represent different modes of regulation (Supplementary Data 1 and 2): (i) Forwarded inter-strain differences in RNA expression that were carried onward to translation (RNA + RIBO; \(n_{\text{liver}} = 292; n_{\text{heart}} = 441\)); (ii) buffered RNA expression differences that were not detected on the level of translation but were present in both RNA-seq experiments (RNA only; \(n_{\text{liver}} = 52; n_{\text{heart}} = 191\)); (iii) reinforced strain-specific differences that were not detected in either RNA-seq experiments but were apparent at the translational level and thus exclusive to ribosome profiling data (RIBO only; \(n_{\text{liver}} = 354; n_{\text{heart}} = 498\)). Whereas in yeast a variable extent of translational buffering was observed\(^{22-24}\), the majority of RNA expression differences were forwarded to the level of translation in mammalian heart (70%) and liver (85%). In addition, we detected hundreds of strain-specific differences in mRNA translation that were not observed at the RNA level almost doubling the number of differentially expressed genes found between strains compared with RNA-seq data alone.

We then quantified the amount of translational regulation by calculating the slopes (standardized major-axis estimation)\(^{25}\) between RNA-seq and Ribo-seq fold changes between strains for genes with differential expression (Fig. 2a). This enabled us to determine the global contribution of translational regulation to strain-specific differences in gene usage in addition to transcriptional regulation. For RNA + RIBO events, transcriptional and translational differences co-occurred and slopes approximated 1 (slope\(_{\text{heart}} = 1.09\); slope\(_{\text{liver}} = 0.97\)). Overall, the majority of RNA expression differences in tissues are associated with equal variation in RPF abundance of the same magnitude and direction.

The slopes for the RIBO only genes were significantly greater than RNA + RIBO slopes (likelihood ratio test; slope\(_{\text{heart}} = 1.71\); slope\(_{\text{liver}} = 1.93\); \(p < 2.2e-16\) for both tissues, indicating reinforced strain differences that arise through post-transcriptional control.

**Translational regulation shapes the proteome.** To assess whether genome-wide strain-specific differences in RPF abundance were indeed reflected on the proteome, we integrated two mass-spectrometry based quantitative proteome data sets from liver of SHR/Ola and BN-Lx rat strains\(^{26}\).

We globally assessed whether our data supported the expected flow of genetic information from transcription through

---

**Figure 2 | Characteristics of genes under translational regulation.** (a) Black slopes approximated 1 (\(x = y\)) and indicated equal fold changes of expression differences between SHR/Ola and BN-Lx across both levels of regulation for RNA + RIBO genes (standardized major-axis estimation; SMA). Significant (\(p < 2.2e-16\)) divergence of blue and red slopes demonstrated translational regulation for RNA only and RIBO only genes, respectively. (b) Strain-specific differences of translationally regulated RIBO only (\(n = 179\)) genes were confirmed in the liver proteome (Wilcoxon–Mann–Whitney; \(p < 1e - 4\)). Variation in translation captured by Ribo-seq resulted in significantly different protein abundances across strains (whiskers indicate 5th and 95th percentiles). RNA only genes have a lower (c) translational efficiency and a significantly increased (d) 3’UTR length compared with RIBO only and RNA + RIBO genes (mean with 95% confidence interval).
translation to protein levels. Each step of biological information processing introduces a new set of specific regulatory changes (for example, post-transcriptional or -translational) that affect the abundance of a gene product. However, consequences of events that occur in later stages are not expected to be evident in preceding stages. In particular, translation levels reflect both transcriptional and translational regulation events, whereas transcript levels only reflect transcriptional—and not translational—regulation events. Therefore, when comparing RNA levels and protein levels conditioning on the translation levels, we would expect no additional information about the protein levels to be present in the RNA and thus they should be conditionally independent. On the other hand, when comparing translation levels and protein levels conditioning on the RNA levels, we would still expect dependences between translation levels and protein levels due to translational regulation events. We first tested the conditional independence between RNA and protein levels, given the Ribo-seq data using partial correlations\(^2\). In both strains the partial correlation was not significantly different from zero (Supplementary Table 1). Second, for each strain we found highly significant partial correlations between Ribo-seq data and protein levels when we conditioned on RNA transcript levels (Supplementary Table 1). This corroborated previous observations\(^1\) that ribosome footprint abundances correlated better with genome-wide protein levels than RNA-seq data, and indicated that Ribo-seq provides a better proxy for protein levels than RNA-seq in our data (Supplementary Fig. 5). Transcript levels provided no additional information to explain protein expression. Strain-specific regulation of translation (RIBO\(_{\text{only}}\)), not detected by RNA expression profiling, was found to account for significant differences in protein abundances across strains (Fig. 2b; Wilcoxon–Mann–Whitney; \(P < 1e-4\)). Ribo-seq data effectively revealed additional variation in gene usage that explains protein levels (Supplementary Fig. 6).

For 25% of differentially transcribed genes across both tissues, we did not detect strain-specific differences in RPF abundances. This can in part be explained by translational buffering of RNA expression differences for these RNA\(_{\text{only}}\) genes (likelihood ratio test: slope\(_{\text{heart}}\) = 0.66; slope\(_{\text{liver}}\) = 0.8; \(P < 2.2e-16\)), reducing the inter-strain variation in gene expression. In addition, RNA\(_{\text{only}}\) genes were translated at a low efficiency (Fig. 2c), resulting in fewer ribosome footprints per transcript. Translational efficiency (TE) is defined as the number of RPFs compared with RNA-seq reads covering the coding sequence (CDS) of genes (see Methods). Low TE effectively reduced the power to detect differences using Ribo-seq compared with RNA-seq. Thus, these genes cannot be as effectively quantified during and after translation, which may result in an overestimation of buffering events. RNA\(_{\text{only}}\) genes possess significantly longer 3’untranslated regions (UTRs) (Fig. 2d; Wilcoxon–Mann–Whitney test; Table 1).

**Figure 3 | Genetic variation and strain-specific differences in translation.** (a,b) Linkage analysis using RNA-seq data of the HxB/BxH recombinant inbred panel\(^3\) derived from the BN-Lx and SHR/Ola strains revealed RNA expression differences under genetic control (eQTLs) for the (a) heart and the (b) liver. Genes with differential RNA expression (forwarded: black bar; buffered: blue bar) between the parental strains were enriched for eQTLs. As expected, genes regulated on the translational level only (red bar) were not significantly enriched for eQTLs (for details see Supplementary Table 2). The majority of RNA expression differences were forwarded to the translational level and detected in both RNA-seq and Ribo-seq experiments. This percentage increased when RNA expression differences were under genetic control. (c) SNP density in the regulatory regions of RNA\(_{\text{only}}\) (blue bar) and RIBO\(_{\text{only}}\) (red bar) genes compared with genes without RNA or RPF strain-specific differences (grey bar). SNP density was significantly higher in the 3’UTR of translationally regulated genes (Wilcoxon–Mann–Whitney). Error bars indicate 95% confidence interval. (d) Known motifs of RNA-binding proteins such as Sf3b4 were more often altered by sequence variants in the 3’UTR compared with genes that did not undergo translational regulation.
Distinct regulation across levels of gene expression. It is a matter of debate how well inherited RNA expression traits (eQTLs) are mirrored by the proteome and how much post-transcriptional processes contribute. Previous studies suggested that \(~32–35\%\) of eQTLs lead to inherited differences in protein abundance\(^{10,30}\) or more recently, green fluorescent protein-tagged\(^{17,18}\) or micro-western array-based approaches\(^3\) focused on subsets of proteins and show that more than 50\% of eQTLs also have a corresponding protein QTL. Here, we assessed the proportion of eQTLs that are forwarded to differential translation at the genome-wide level in mammalian tissues. We previously performed RNA expression linkage analysis of the HxB/BxH recombinant inbred panel derived from the parental BN-Lx and SHR/Ola strains\(^{31}\) (see Methods). Of 489 differentially transcribed genes in the heart, 143 have an eQTL (\(\chi^2\)-test; \(P_{\text{heart}}\leq 2.2\times 10^{-16}\); in liver: 106 out of 238; \(P_{\text{liver}}\leq 2.2\times 10^{-16}\); see also Supplementary Table 2). More than 80\% of genes with an eQTL are differentially translated (Fig. 3a) between the two parental strains, demonstrating that genetically induced RNA expression changes are largely carried forward to differences in RPF levels and are not frequently buffered during translation. We also detected hundreds of RIBO\(_{\text{only}}\) events that do not overlap with eQTLs and yet affect protein abundance in the parental strains (Fig. 3a,b). These post-transcriptional gene expression differences cannot be captured by traditional eQTL approaches but may likely contribute to phenotypic variation.

Accurate quantification of translation in the context of natural phenotypic variation created the opportunity to begin to identify regulators of post-transcriptional processes. We observed an increase in the single-nucleotide polymorphism (SNP) density in the 3’UTR of genes under translational control (Fig. 3c), suggesting the presence of genetic variation in cis-regulatory elements. We tested whether known motifs of RBPs are significantly enriched for genetic variation, but the analysis also suggested a role for splicing factors such as SF3B4 (Wilcoxon–Mann–Whitney test; \(P_{\text{corrected}} = 0.026\)) were significantly enriched for genetic variation, but the analysis also suggested a role for splicing factors such as SF3B4 (Wilcoxon–Mann–Whitney test; \(P_{\text{corrected}} = 0.033\)) as well as other factors in translation by binding to cis-regulatory elements (Fig. 3d; see also Supplementary Tables 3 and 4).

Figure 4 | Genes associated with human cardiac and hepatic complex traits are under transcriptional and/or translational control. The SHR/Ola rat strain is an animal model for traits investigated in human GWAS studies. Genes associated with ‘heart rate’: Myh6, Gja1 and Fads1 (ref. 36), ‘metabolite levels’: Acadl, Tat and Abca8b (ref. 36, 37), ‘high-density lipoprotein cholesterol’: Abca8b\(^{63–65}\) are under transcriptional and/or translational control. Representative examples of two genes for each group (RNA\(_{\text{only}}\), RNA + RIBO and RIBO\(_{\text{only}}\)) are shown. For each gene, read counts for five biological replicates per strain across three sequencing experiments are plotted: ribosome profiling (Ribo-seq), polyA-selected RNA sequencing (RNA-seq1) and rRNA-depleted total RNA sequencing (RNA-seq2). Error bars indicate s.d. \(P\) values were corrected for multiple testing and indicate significant differences in expression (DESeq2 (ref. 45); FDR \(\leq 0.01\)).
both tissues from SHR/Ola (n = 4) and BN-Lx rat strains (n = 4). Binding sites for differentially transcribed miRNAs (FDR <0.05) were enriched in the 3′UTRs of differentially translated (RIBOseq) genes when compared with differentially transcribed (RNA + RIBO) genes (χ²-test; P = 0.008; Supplementary Table 5). This may suggest a widespread role of miRNAs in strain-specific translational regulation in mammalian tissues.

Discussion

We aimed to estimate the relative contribution of translational and transcriptional regulation to the expression differences in our rat model for cardiac and metabolic disease. First, we performed gene set enrichment analyses to identify disease-specific pathways that are differentially regulated between strains (Supplementary Data 3–5). Of 37 and 31 enriched KEGG (kyoto encyclopedia of genes and genomes) pathways in the heart and liver, respectively, the majority, 70.3% in the heart and 77.4% in the liver, were only detected with Ribo-seq data. Thus, the investigation of Ribo-seq data revealed a large proportion of potentially disease-specific genes and genomes (network) pathways in the heart and liver, respectively.

Data 3–5). Of 37 and 31 enriched KEGG (kyoto encyclopedia of genes and genomes) pathways in the heart and liver, respectively, the majority, 70.3% in the heart and 77.4% in the liver, were only detected with Ribo-seq data. Thus, the investigation of Ribo-seq data revealed a large proportion of potentially disease-specific genes and genomes (network) pathways in the heart and liver, respectively.

Parallel generation of ribosome profiling and RNA libraries. For each animal ~100 mg of frozen tissue were pulverized manually under liquid nitrogen and lysed in 1 ml lysis buffer (1 × ARTrseq mammalian polysome buffer (Expressent), 1% Triton X-100, 0.1% NP-40, 1 mM diethiothreitol, 10 U ml⁻¹ DNase I and nuclease-free H₂O to final volume). To impede post-lysis translation, the lysis buffer was supplemented with cycloheximide (Sigma), previously dissolved in EtOH, at a final concentration of 0.1 mg ml⁻¹. The tissue was homogenized further by repeatedly passing the lysate through a 21-gauge syringe needle. For complete lysis, the samples were kept on ice for 10 min and subsequently centrifuged at 20,000g to precipitate cell debris.

To accurately dissect translation and transcription, we prepared both Ribo-seq and RNA-seq libraries for each biological replicate from the identical lysate. Ribosomal footprints were generated by treating part of the lysate with proprietary ARTrseq ribosome profiling and RNA sequencing library preparation kit (Epicentre). Sequencing adapters were ligated before the samples were retrotranscribed and again PAGE purified. Circularized complementary DNA (cDNA) templates were amplified with 12 PCR cycles using Phusion polymerase (NEB). Following an additional native PAGE purification (8% TGE gel) step, libraries were quantified using the Qubit fluorometer, while the quality and average fragment size were estimated at the Bioanalyzer (High Sensitivity assay, Agilent). If not described otherwise, steps of the library generation were performed according to the mammalian ARTrseq kit.

Barcodes were used to perform multiplex sequencing and create sequencing pools containing at least eight different samples and always an equal amount of both RNA and RFP libraries. We then sequenced sample pools on several lanes on the HiSeq 2000 platform using 50 bp sequencing chemistry to reduce barcode-, lane-, or slide-related bias in the sequencing data.

Sequencing data processing and quality control. Raw sequencing data were demultiplexed with the CASAVA 1.8 pipeline and the 3′-end adapter was removed using the FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), retaining reads of 20 nt or longer post-clipping length: fastx_clipper -a 5′-AGATCGGAAGAGCACACGTCT-3′ -l 20 -n -v -Q33 followed by fastx_trimmer -Q33 -f 1. We then removed redundant reads and created a custom index for all samples and performed a custom alignment using bowtie2 (see Supplementary Table 6). We then identified differences in read abundance. To avoid differences between RNA-seq and Ribo-seq that arise through differing attributes of the sequencing data such as read length or sequencing depth, we matched the data before detecting strain-specific differences in read abundance. To avoid differences between RNA-seq and Ribo-seq that arise through mapping artefacts, we then trimmed all Ribo-seq reads to 29 nt length before alignment: fastx_trimmer -l 29 -Q33. This is the most common length of Ribo-seq reads (Supplementary Fig. 1b).

To avoid mapping artefacts due to split-read alignment of the short reads, we defined splice junctions of ribosome profiling data across strains and categorize differential expression events specific to either both or only one data set type. To avoid differences that arise through differing attributes of the sequencing data such as read length or sequencing depth, we matched the data before detecting strain-specific differences in read abundance. To avoid differences between RNA-seq and Ribo-seq that arise through mapping artefacts, we then trimmed all Ribo-seq reads to 29 nt length before alignment: fastx_trimmer -l 29 -Q33. This is the most common length of Ribo-seq reads (Supplementary Fig. 1b).

To avoid mapping artefacts due to split-read alignment of the short reads, we defined splice junctions of ribosome profiling data across strains and categorize differential expression events specific to either both or only one data set type. To avoid differences that arise through differing attributes of the sequencing data such as read length or sequencing depth, we matched the data before detecting strain-specific differences in read abundance. To avoid differences between RNA-seq and Ribo-seq that arise through mapping artefacts, we then trimmed all Ribo-seq reads to 29 nt length before alignment: fastx_trimmer -l 29 -Q33. This is the most common length of Ribo-seq reads (Supplementary Fig. 1b).

To avoid mapping artefacts due to split-read alignment of the short reads, we defined splice junctions of ribosome profiling data across strains and categorize differential expression events specific to either both or only one data set type. To avoid differences that arise through differing attributes of the sequencing data such as read length or sequencing depth, we matched the data before detecting strain-specific differences in read abundance. To avoid differences between RNA-seq and Ribo-seq that arise through mapping artefacts, we then trimmed all Ribo-seq reads to 29 nt length before alignment: fastx_trimmer -l 29 -Q33. This is the most common length of Ribo-seq reads (Supplementary Fig. 1b).

To avoid mapping artefacts due to split-read alignment of the short reads, we defined splice junctions of ribosome profiling data across strains and categorize differential expression events specific to either both or only one data set type. To avoid differences that arise through differing attributes of the sequencing data such as read length or sequencing depth, we matched the data before detecting strain-specific differences in read abundance. To avoid differences between RNA-seq and Ribo-seq that arise through mapping artefacts, we then trimmed all Ribo-seq reads to 29 nt length before alignment: fastx_trimmer -l 29 -Q33. This is the most common length of Ribo-seq reads (Supplementary Fig. 1b).

To avoid mapping artefacts due to split-read alignment of the short reads, we defined splice junctions of ribosome profiling data across strains and categorize differential expression events specific to either both or only one data set type. To avoid differences that arise through differing attributes of the sequencing data such as read length or sequencing depth, we matched the data before detecting strain-specific differences in read abundance. To avoid differences between RNA-seq and Ribo-seq that arise through mapping artefacts, we then trimmed all Ribo-seq reads to 29 nt length before alignment: fastx_trimmer -l 29 -Q33. This is the most common length of Ribo-seq reads (Supplementary Fig. 1b).
detecting expression differences introduced by variation in sequencing depth across technologies. For each animal, we matched and downsampl ed both RNA-seq and the Ribo-seq libraries to the size of the library with the lowest read counts within each animal (Supplementary Fig. 2).

All sequencing data sets were then matched in depth (effective depth of reads contributing to the analysis) and read length to ensure that exclusive characteristics of Ribo-seq or RNA-seq experiments reflect true biological processes and not technical properties of the different methods.

To assess the quality of the Ribo-seq data sets, we calculated the length-normalized, genome-wide average expression of the UTR and coding regions. For each library, we counted the number of reads mapping to either 5'UTR, Coding Sequence or 3'UTR and normalized it by the total length of the respective feature and the library size. Ribosome footprints of the heart and liver were distributed across gene bodies as expected and mainly covered the coding sequence of genes or to a lesser extent the 5'UTR. The 3'UTR was depleted in reads of the Ribo-seq data sets (Supplementary Fig. 1a). We also determined the position of read starts in the proximity of the start and stop codon in the Ribo-seq data. For all 20 Ribo-seq data sets, we observed clear triplet periodicity, distinctive for actively translating ribosomes, and a strong peak of 29-nt oligomers at 12 bp upstream of the start codon, indicating ribosomes located at the translation start site. To estimate sample-to-sample distances, we clustered all samples within each data set according to the Euclidean distances using the DESeq2 (refs 21,45) package (Supplementary Fig. 2).

Differential expression analysis. To quantify expression of genes, we first assigned uniquely mapping reads to their genomic feature (Ensembl database release 72 (ref. 41)) and treated spliced reads between exons as a single fragment using HTSeq. htseq-count-stranded = no -t type = exon -idattr = gene_id. DESeq2 (refs 21,45) with standard parameters was then used to deduce significant alterations in transcription and translation of cellular transcripts (Supplementary Fig. 3; Bonferroni adj. \( P \leq 0.01 \)).

To determine the relationship of transcriptional and translational regulation, we defined three groups of genes: (i) RNA + Ribo: denotes forwarded significant differences in expression across strains that are detectable in both RNA-seq experiments as well as in the Ribo-seq data (Bonferroni adj. \( P \leq 0.01 \)). Differential RNA expression of these genes is promoted to the translational level and has a direct effect on protein synthesis. (ii) RNA only: are buffered genes that do not exhibit significant inter-strain variation on the translational level but that were detected as differentially transcribed in both RNA-seq experiments (Bonferroni adj. \( P \leq 0.01 \)). Translational regulation counteracts changes on the RNA level and buffers differences in between the strains. (iii) Ribo only: indicates reinforced and translationally induced strain-specific differences, which are exclusive to Ribo-seq and cannot be detected in either of the two RNA-seq experiments (Bonferroni adj. \( P \leq 0.01 \)). Ribosome profiling can reveal differential usage of genes that is either not present or under-represented in RNA expression data sets.

To define differences between Ribo-seq and RNA-seq data that arise due to translational regulation, we performed RNA sequencing using two different methodologies: sequencing of RNA-depleted total RNA (RNA-seq2) and polyadenylated transcripts (RNA-seq1).

RNA-seq: technical control for Ribo-seq. The libraries were created in parallel from the identical tissue lysate with a procedure very similar to Ribo-seq. However, ribosomes occupy mostly processed transcripts, whereas total RNA-seq also contains nuclear RNA. Thus we also used RNA-seq1 (PolyA-selected RNA-seq) as a control data set to determine whether processed transcripts are differentially expressed on the RNA level.

For each data set, we also calculated fold changes of normalized read counts per gene for SHR/Olaa over BN-Lx. RNA fold changes are the average of RNA-seq1 and RNA-seq2. To assess whether RNA-seq and Ribo-seq fold changes correlate significantly on a global scale, we computed the Pearson’s correlation for both tissues (\( r = 0.4 \)).

Comparing RNA-seq and Ribo-seq fold changes uncovers the extent of translational regulation for each gene. When differences on the RNA level induce equivalent changes in the ribosome occupancy, no translational regulation is present and strain-specific differences are promoted across regulatory layers. An increase in the magnitude of changes across transcripts indicates translational reinforcement or buffering. To estimate the extent of translational regulation for the RNA only, RNA + Ribo and Ribo only gene sets, we computed the slopes of three linear models:

\[
\log FC (\text{Ribo-seq SHR}/BN) = \beta_{\text{RNA only}} + \beta_{\text{Ribo only}} \times \log FC (\text{RNA-seq SHR}/BN)
\]

\[
\log FC (\text{Ribo-seq SHR}/BN) = \beta_{\text{RNA only}} + \beta_{\text{Ribo only}} \times \log FC (\text{RNA-seq SHR}/BN)
\]

\[
\log FC (\text{Ribo-seq SHR}/BN) = \beta_{\text{RNA only}} + \beta_{\text{Ribo only}} \times \log FC (\text{RNA-seq SHR}/BN)
\]

using the major-axis estimation method. We then tested the null hypothesis of \( \beta_{\text{Ribo only}} = \beta_{\text{RNA only}} = 0 \) using the likelihood ratio statistic implemented in the smatr 3 R package.

Proteome integration. Global protein abundances were estimated in the liver of BN-Lx and SHR/Ola using liquid chromatography-tandem mass spectrometry (LC-MS/MS) as reported by Low et al.\(^{22}\) We calculated the expression of all genes based on RNA-seq (average of RNA-seq 1 and RNA-seq2) as well as on Ribo-seq data sets of both strains in the liver (per segment per kilobase per million reads mapped ≥1) and correlated these values with MS-based proteomic data (all data were log transformed). Correlation between Ribo-seq and protein levels was higher than correlation between RNA-seq and protein levels in both strains (Supplementary Fig. 5). To assess whether our data support the expected model for the flow of genetic information from transcription, through translation, to protein levels, we tested for conditional independence using partial correlations.

We transvers the three range variables ‘RNA’, ‘Ribo’ and ‘protein’. For each strain we tested all pairs of the three variables ‘RNA’, ‘Ribo’ and ‘protein’ for conditional independence, given the third variable. Intuitively, conditional independence between ‘RNA’ and ‘protein’ conditional on ‘Ribo’ means that knowledge about ‘RNA’ is irrelevant for knowledge about ‘protein’ if we know ‘Ribo’. We assume that tuples (‘RNA’, ‘Ribo’, ‘protein’) follow a multivariate Gaussian distribution. Then the conditional distribution of two variables, given the third, is a bivariate Gaussian distribution, with the partial correlation coefficient describing the relation between the two variables. In this model, conditional independence is equivalent to a zero partial correlation coefficient. Partial correlation coefficients can be obtained from the inverse of the covariance matrix or by correlating the residuals obtained by regressing each of the two variables against the third. Here we used the second approach and tested the significance of the partial correlation coefficient using the t-distribution.\(^{97}\)

To test whether variation in ribosome occupancy—in the absence of differential RNA expression—is also reflected on the protein level, we investigated the relative expression of 179 proteins in more detail. These proteins of interest were chosen from different translational regulation (Ribo only) in the liver and were also quantified by MS in both strains. Translationally induced strain-specific differences resulted in significant concordant changes (Wilcoxon–Mann–Whitney; \( P < 1e-4 \)) on the protein level (Fig. 2b).

Translational efficiency. The TE score was calculated for the different gene groups (RNA only, Ribo only, and RNA + Ribo) as the ratio of normalized Ribo-seq reads in the CDS over normalized RNA-seq reads (average of RNA-seq1 and RNA-seq2) in the CDS, thus avoiding the length of UTRs to influence our calculation. Ribo-seq and RNA-seq data cover different parts of genes making it necessary to focus on the CDS, a region covered by both methodologies, when performing quantitative comparisons across both technologies.

UTR length and SNP enrichment. We determined the length and position of UTRs as annotated by the Ensembl release 72 database and tested for differential length and SNP density using the Wilcoxon–Mann–Whitney method.

Enrichment of KEGG pathways. We compared enrichment analyses based on gene sets differentially regulated according to either technology: RNA-seq (RNA + Ribo and RNA only) or Ribo-seq (RNA + Ribo and Ribo only). We identified a significant over-representation of KEGG pathways in differentially used genes for both Ribo-seq and RNA-seq data using webGestalt.\(^{43}\) At least two genes for each pathway were required. Pathways had to be significantly enriched (\( P_{\text{adjusted}} \leq 0.01 \)) enriched according to a hypergeometric test. All pathways listed by the KEGG database (as of 23 March 2011) were tested for a significant over-representation of differentially expressed genes. Pathways were either detected using both methodologies (black), or exclusively enriched in either Ribo-seq (red) or RNA-seq (blue) data (Supplementary Data 3–5).

miRNA expression. To test for differential miRNA expression, we analysed four biological replicates of each strain. Total RNA was extracted from the left ventricle tissue using the mirVana miRNA Isolation kit (Life Technologies). RNA was quantified and the quality assessed using Agilent’s Bioanalyzer. To enrich for 15–35-bp small RNAs, 10 μg total RNA was purified on 15% denaturing PAA gel (7 M urea). pLoud miRNA libraries were prepared using the Small RNA Seq kit (Life Technologies) following the manufacturer’s guidelines. After adapter ligation, cDNA was purified using the MinElute PCR Purification kit (Qiagen) and size selected on a 10% TBE-urea gel. cDNA fragments were then amplified using barcoded primers for multiplex sequencing. The final cDNA libraries were quantified using the Qubit fluorometer and assessed with the DNA 1,000 kit (Agilent) and sequencing was carried out on a SOLiD 3 system.

Genome mapping and miRNA quantification. Adapter-clipped reads were mapped against mm with Bowtie.\(^{28}\) (Bowtie options: -l 17 -a -m 2 -n 2 --best --strata -C -f) to accurately quantify annotated rat miRNAs. We assessed the expression of 680 annotated mature miRNAs. We performed a quantile-based scaling to normalize read counts as described previously. Normalized read counts per miRNA were then compared between BN-Lx and SHR/Ola using a generalized linear model with the quasi-Poisson family. The FDR was controlled using the Benjamini–Hochberg procedure. We then assessed whether differential
miRNA expression was associated preferentially with transcriptional or post-transcriptional control of protein synthesis by comparing the union of the target sets of all differentially expressed miRNAs with the prediction of genes based on differential translation and transcription. We found that genes with variation in translational regulation (RIBO\textsubscript{mRNA}) were more often predicted targets\(^{36}\) of differentially expressed miRNAs (FDR <0.05) than differentially expressed (RNA \(\leftrightarrow\) RIBO) genes (Supplementary Table 5) using Fisher’s exact test in each tissue separately. Finally, we combined the results from both tissues using the Fisher\texttextsubscript{2} method\(^{37}\).

**eQTL analysis.** To identify loci that determine RNA expression levels and the respective target genes, we previously performed an eQTL study for both tissues based on RNA-seq data of the HxB/Bsh recombinant inbred panel\(^{37}\). Briefly, we sequenced the RNA of one animal for each of the 30 recombinant inbred strains, each on one lane of a HiSeq 2000 Illumina instrument with TruSeq 2 \(\times\) 100 bp paired-end chemistry. Reads were aligned to the reference genome RGC3.34 using TopHat v1.2.0 (ref. 43). Gene expression levels were estimated by counting reads within gene bodies. Gene expression levels were normalized using a quantile-based scaling method\(^{36}\). QTLs were identified by trait–marker regression using a negative binomial regression model38 similar to the model used for differential analysis in DESeq. The genetic map consisted of 1,384 non-redundant SNP genotype profiles. For each pair of trait i and marker j, we computed the likelihood ratio statistic of the full model containing the genotype variable against a null model containing only an intercept term. To determine the significance of likelihood ratio statistic scores while accounting for linkage disequilibrium, we used a permutation strategy. Finally, we adjusted the QTL \(P\) values for multiple testing using the Benjamini–Hochberg method\(^{55}\).

We then compared the classification of genes based on differential translation and transcription to the set of genes that have eQTLs using the \(\chi^2\)-test. As expected, we found a strong enrichment of eQTLs for differentially expressed genes on the RNA level in the founder strains (Supplementary Table 2). Genes that have an eQTL on the RNA level are mostly detected as being also differential using Ribo-seq, which indicates that genetically induced variation in RNA expression is carried forward to transcription (Fig. 3a,b). Genes predicted to be mostly under translational control (RIBO\textsubscript{mRNA}) were not enriched for eQTLs, thus representing a layer of regulation not captured by RNA expression-based linkage studies.

**Motif analysis.** To determine whether the binding sites of expressed RNA-RPs were preferentially altered by SNPs in 3\’UTRs of translationally regulated (RIBO\textsubscript{mRNA}) genes as compared with SNPs in 3\’UTR of genes that were regulated on both levels (RNA \(\leftrightarrow\) RIBO), we first obtained 50 bp of flanking sequences for all SNPs located in all 3\’UTR regions subject to the integrated analysis in each tissue. For each position weight matrix from the Hughes RBP database\(^{59}\), we scored the difference of binding affinities of the reference and alternative alleles using the e1001393 (2011).

13. Petretto, E. & Califano, A. Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass. Nat. Genet. 34, 546–552 (2006).
14. Gravett, R. A. et al. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428, 493–521 (2004).
15. Atanur, S. S. et al. The genome sequence of the spontaneously hypertensive rat. J. Genet. 156, 158 (2001).
16. Pravenec, M. et al. Transgenic rescue of defective Cd36 ameliorates insulin resistance in spontaneously hypertensive rats. Nat. Genet. 27, 156–158 (2001).
17. Pravenec, M. et al. Identification of renal Cd36 as a determinant of blood pressure and risk for hypertension. Nat. Genet. 40, 952–954 (2008).
18. Pacheco, E. J. et al. Genome sequencing reveals loci under artificial selection that underlie disease phenotypes in the laboratory rat. Cell 154, 691–703 (2013).
19. Anders, S. & Huber, W. Differential expression analysis for sequence count data. Genome Biol. 11, R106 (2010).
20. McManus, C. I., May, G. E., Spellman, P. & Steyton, A. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. Genome Res. 24, 422–430 (2014).
21. Artieres, C. & Fraser, H. Evolution at two levels of gene expression in yeast. Genome Res. 24, 411–421 (2014).
22. Albert, F. W., Muzzey, D., Weissman, J. S. & Kruglyak, L. Genetic influences on translation in yeast. PLoS Genet. 10, e1004692 (2014).
23. Warton, D. I., Duursma, R. A., Falster, D. S. & Taskinen, S. smatr 3 - an R package for estimation and inference about allometric lines. Methods Ecol. Evol. 3, 257–259 (2012).
24. Low, T. et al. Quantitative and qualitative proteome characteristics extracted from in-depth integrated genomics and proteomics analysis. Cell Rep. 5, 1469–1478 (2013).
25. Edwards, D. Introduction to Graphical Modelling (Springer, 2000).
26. Sandberg, R., Neilson, J., Sarma, A., Sharp, P. & Burge, C. Profiling eukaryotic mRNAs with short-read 3\’ untranslated regions and fewer microRNA target sites. Science 320, 1643–1647 (2008).
27. Mayr, C. & Bartel, D. Widespread shortening of 3\’UTR by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 138, 673–684 (2009).
28. Marfat, A. et al. Comparative analysis of proteome and transcriptome variation in mouse. PLoS Genet. 7, e1001393 (2011).
29. Rintisch, C. et al. Natural variation of histone modification and its impact on gene expression in the rat. Genome Res. 24, 942–953 (2014).
32. Richter, J. D. CPEB: a life in translation. Trends Biochem. Sci. 32, 279–285 (2007).
33. Guo, H., Ingolia, N., Weissman, J. & Bartel, D. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466, 835–840 (2010).
34. Bazzini, A., Lee, M. & Giraldez, A. Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. Science 336, 233–237 (2012).
35. Wartenberg, D. et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 42, D1001–D1006 (2014).
36. Hoed, M. den & Illig, T. A genome-wide perspective of genetic variation in human metabolism. Nat. Genet. 42, 137–141 (2010).
37. Flicek, P. et al. A genome-wide assessment of variability in human serum metabolite composition. Hum. Mutat. 34, S15–S24 (2013).
38. Richter, J. D. CPEB: a life in translation. Nature Communications 5, 1088 (2014).
39. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, R25 (2014).
40. Langmead, B. & Trapnell, C. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012).
41. Flicek, P. et al. Ensembl 2013. Nucleic Acids Res. 41, D84–D95 (2013).
42. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964 (1997).
43. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111 (2009).
44. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36 (2013).
45. 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).
46. Warton, D. & Weber, N. Common slope tests for bivariate errors-in-variables models. Biometrika 97, 279–285 (2010).
47. Kettunen, J. et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat. Genet. 44, 269–276 (2012).
48. Graybill, F. A. Theory and Application of the Linear Model (Duxbury, 2000).
49. Johnstone, P. J. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964 (1997).
50. Wang, K., Li, M. & Hakonarson, H. Cufflinks: improved methods for genome-wide transcriptome analysis. Genome Biol. 11, R13 (2010).
51. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25 (2009).
52. Schulte, J. H. et al. Deep sequencing reveals differential expression of microRNAs in favorable versus unfavorable neuroblastoma. Nucleic Acids Res. 38, 5919–5928 (2010).
53. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B. 57, 289–300 (1995).
54. Fisher, R. A. Statistical Methods for Research Workers (Oliver and Boyd, 1925).
55. Venables, W. N. & Ripley, B. D. Modern Applied Statistics With S (Springer, 2002).
56. Ray, D. et al. A compendium of RNA-binding motifs for decoding gene regulation. Nature 499, 172–177 (2013).
57. Liu, J. et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat. Genet. 44, 269–276 (2012).
58. Schulte, J. H. et al. Deep sequencing reveals differential expression of microRNAs in favorable versus unfavorable neuroblastoma. Nucleic Acids Res. 38, 5919–5928 (2010).
59. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B. 57, 289–300 (1995).