Supplemental Figures S1 to S12

Widespread dysregulation of long non-coding genes associated with fatty acid metabolism, cell division, and immune response gene networks in xenobiotic-exposed rat liver

Kritika Karri and David J. Waxman
Dept. of Biology and Bioinformatics Program, Boston University
Boston, MA 02215 USA

Figure S1. Module and sub−module hierarchy for MEGENA modules. MEGENA uses a divisive clustering approach and discovers co-expression modules in a multi-layer manner. The innermost core, C1_1, contains all 2,637 PCGs and 1,447 xeno-lncs, which are clustered into 13 gene modules in layer 1 (C1_3 to C1_15). Gene modules in layer 1 are further clustered into smaller compact sub−modules in layer 2. This process continues until no further compact child clusters are formed.

Figure S2. Gene expression data for xeno-lncs (**A**) and protein coding genes (PCGs) (**B**) that are consistently induced or repressed by $\geq$20 of the 27 chemicals examined. Data are shown as $\log_2$ fold-change (FC) values along the Y-axis. Bars shown in white, FDR < 0.05.

A. Five lncRNAs (rlnc4657, rlnc3088, rlnc715, rlnc1425, and rlnc2750) showed down regulation in 20 or more chemicals.

B. Sult2a was up regulated by 22 out of 27 chemicals, and Ltc4s gene was down-regulated by 22 out of 27 chemicals. Acot1 was up-regulated by 21 chemicals, but was down-regulated by two of the three AhR agonists, and by aflatoxin−B1, which also has AhR agonist activity (see text). Bars are colored according to the MOA of each chemical.

Figure S3. Gene expression profiles across all 27 chemicals for representative MOA-selective marker genes, including PCGs (**A**) and xeno-lncs (**B**). Gene expression data is in $\log_2$ FC values (y-axis), and bars are colored according to the MOA of each chemical. White bars, FDR < 0.05.

Figure S4. Rat−mouse ortholog responses to xenobiotics that are activators of CAR or PXR.

A, Heatmap of 140 rat xeno-lncs whose mouse orthologs was significantly dysregulated by a CAR or PXR agonists in one of the mouse datasets (see text). Data are displayed by hierarchical clustering using Euclidean distance metric and Ward.d2 minimum variance criterion. Each row represents a lncRNA rat−mouse ortholog pair and each column represents one gene expression dataset.

B, Expression data for select rat-mouse xeno-lnc orthog pairs (top) and of four PCGs co-expressed with the orthologs pair rlnc2209-mlnc3859 (bottom).

Figure S5. Module C10, which is highly enriched in ER marker genes. Xeno-lncs occupying central position as hubs and bottlenecks for module C10 that contained all 50 ER markers.

Figure S6. Complete network of module C9, which is enriched for fatty acid metabolism terms. Fatty acid metabolism genes are represented by nodes shown in blue. Network submodules (C−59, C−60, C−61) are represented by different colors.

Figure S7. Oncogenic sub−network derived from module C13. This module includes 35 PCGs with cancer−associated roles, either as oncogenic drivers or tumor suppressors (TSGs). Five xeno-lncs are connected directly to these genes.

Figure S8. Subnetworks involving hub and bottleneck genes in module C7.
A. rlnc2830, a hub gene from module C7, is positively co-expressed with nine PCGs involved in immune response.
B. Responses of rlnc2830 and its PCG partners across 27 chemicals (in log2 FC).
C. rlnc1130, a hub gene connected to six genes in module C7 and rlnc1023, a hub−bottleneck gene with connections to nine genes. rlnc1023 negatively correlated with Arg1, an immunosuppressive gene (Arlauckas *et al.*, 2018), and with Emp2, a tumor suppressor (Li *et al.*, 2013). These two regulatory xeno-lncs were negatively correlated with Sox4, a heptaocarcinogenic driver (Hur *et al.*, 2010), and showed positive associations with Il6R and Mat1a.
D. Responses of rlnc1130 and rlnc1023 and their PCG partners across 27 chemicals (in log2 FC).

Figure S9. lncRNA−PCG causal network enriched for different biological processes. Each directed edge (arrows) represents a causal effect (absolute causal effect value > 0.5) of a xeno-lnc (diamond shapes) on the expression of a PCG. Ortholog information is represented by different node colors with node description added for functionally well−characterized lncRNAs.

Figure S10. Apoptosis PCG-xeno-lnc co-expression networks.

A. Shown are network based on 40 apoptosis-related PCGs that respond to one or more of the 27 xenobiotics and made direct connections with other apoptosis-related genes or with 96 xeno-lncs based on correlation $\geq$ 0.8.

B. Heat map presenting 40 apoptosis-related PCGs (black text) that are exclusively connected to a set of 96 xeno-lncs (red text). In addition, we identified several known lncRNA orthologs (red arrows).

Figure S11. Liver cirrhosis PCG-xeno-lnc co-expression networks.

A. Shown are network based on subset 174 liver cirrhosis-related PCGs that respond to one or more of the 27 xenobiotics and made direct connections with other cirrhosis-related genes or with 58 xeno-lncs based on correlation $\geq$ 0.8.

B. Heat map presenting 60 cirrhosis-related PCGs (black text) that are exclusively connected to a set of 58 xeno-lncs (red text). In addition, we identified several known lncRNA orthologs (red arrows).

Figure S12. A. Heat map presenting gene response for oncogenic genes from Module C7, which is enriched for immune response genes. The 125 oncogenes (black text) genes displayed here are connected to one or more of 49 xeno-lncs (gold text). We observed a small cluster (marked at the bottom as a dotted rectangle) that was robustly down-regulated across all chemical exposures. Xeno-

lnc rlnc4110 (blue arrow) was induced across all conditions. In addition, we identified several known lncRNA orthologs (red arrows). **B**. Oncogenic gene sub-network, excerpted from module C7. This network presents oncogenic genes and their direct xeno-lnc neighbors. Three of the onco-lnc orthologs shown, lnc-CYTOR, Linc00941, RP11-405F3.4, are connected to critical node genes in the network.

References in Supplementary Figure legends:

Arlauckas, S. P., Garren, S. B., Garris, C. S., Kohler, R. H., Oh, J., Pittet, M. J., and Weissleder, R. (2018). Arg1 expression defines immunosuppressive subsets of tumor-associated macrophages. *Theranostics* **8**(21), 5842-5854.

Hur, W., Rhim, H., Jung, C. K., Kim, J. D., Bae, S. H., Jang, J. W., Yang, J. M., Oh, S. T., Kim, D. G., Wang, H. J.*, et al.* (2010). SOX4 overexpression regulates the p53-mediated apoptosis in hepatocellular carcinoma: clinical implication and functional analysis in vitro. *Carcinogenesis* **31**(7), 1298-307.

Li, C.-F., Chen, L.-T., Lin, C.-Y., Wang, Y.-H., Huang, H.-Y., Hsing, C.-H., Tsai, C.-J., and Shiue, Y.-L. (2013). Loss of epithelial membrane protein-2 expression confers an independent prognosticator in gallbladder carcinoma. *Biomarkers and Genomic Medicine* **5**(1), 31-38.