Science Translational Medicine

A database of tissue-specific rhythmically expressed human genes has potential applications in circadian medicine

Marc D. Ruben, Gang Wu, David F. Smith, Robert E. Schmidt, Lauren J. Francey, Yin Yeng Lee, Ron C. Anafi and John B. Hogenesch

Sci Transl Med **10**, eaat8806.
DOI: 10.1126/scitranslmed.aat8806

Body timing

Although the existence of circadian clock–dependent modulation of gene expression in humans has been known for more than a decade, the relevance of the circadian clock in drug response and therapeutic outcome has been only recently appreciated. Now, Ruben *et al*. used an algorithm called cyclic ordering by periodic structure (CYCLOPS) to create a database of cycling genes in 13 human tissues. The authors show that several rhythmically expressed genes code for known drug targets or for proteins involved in drug transport and metabolism. The data represent a useful resource for circadian medicine and strengthen the notion that circadian rhythms should be considered when determining therapeutic interventions.

Science Translational Medicine
AAAS

www.sciencetranslationalmedicine.org/cgi/content/full/10/458/eaat8806/DC1

Supplementary Materials for

A database of tissue-specific rhythmically expressed human genes has potential applications in circadian medicine

Marc D. Ruben, Gang Wu, David F. Smith, Robert E. Schmidt, Lauren J. Francey, Yin Yeng Lee, Ron C. Anafi, John B. Hogenesch*

*Corresponding author. Email: john.hogenesch@cchmc.org

Published 12 September 2018, *Sci. Transl. Med.* **10**, eaat8806 (2018)
DOI: 10.1126/scitranslmed.aat8806

The PDF file includes:

Fig. S1. Pipeline for sample selection and CYCLOPS ordering for 13 tissues.
Fig. S2. Numbers of cycling genes by different statistical thresholds and noncoding gene type.
Fig. S3. Phase set enrichment in human and mouse counterpart tissues.
Fig. S4. Gene expression traces for top cardio-cycling drug targets.
Legends for data files S1 to S3

Other Supplementary Material for this manuscript includes the following:

(available at www.sciencetranslationalmedicine.org/cgi/content/full/10/458/eaat8806/DC1)

Data file S1 (Microsoft Excel format). Cosinor regression output for all genes and all tissues.
Data file S2 (Microsoft Excel format). Set of 54 genes that cycle in 8 or more of the 13 tissues.
Data file S3 (Microsoft Excel format). Set of 136 cardio-cycler drug targets.

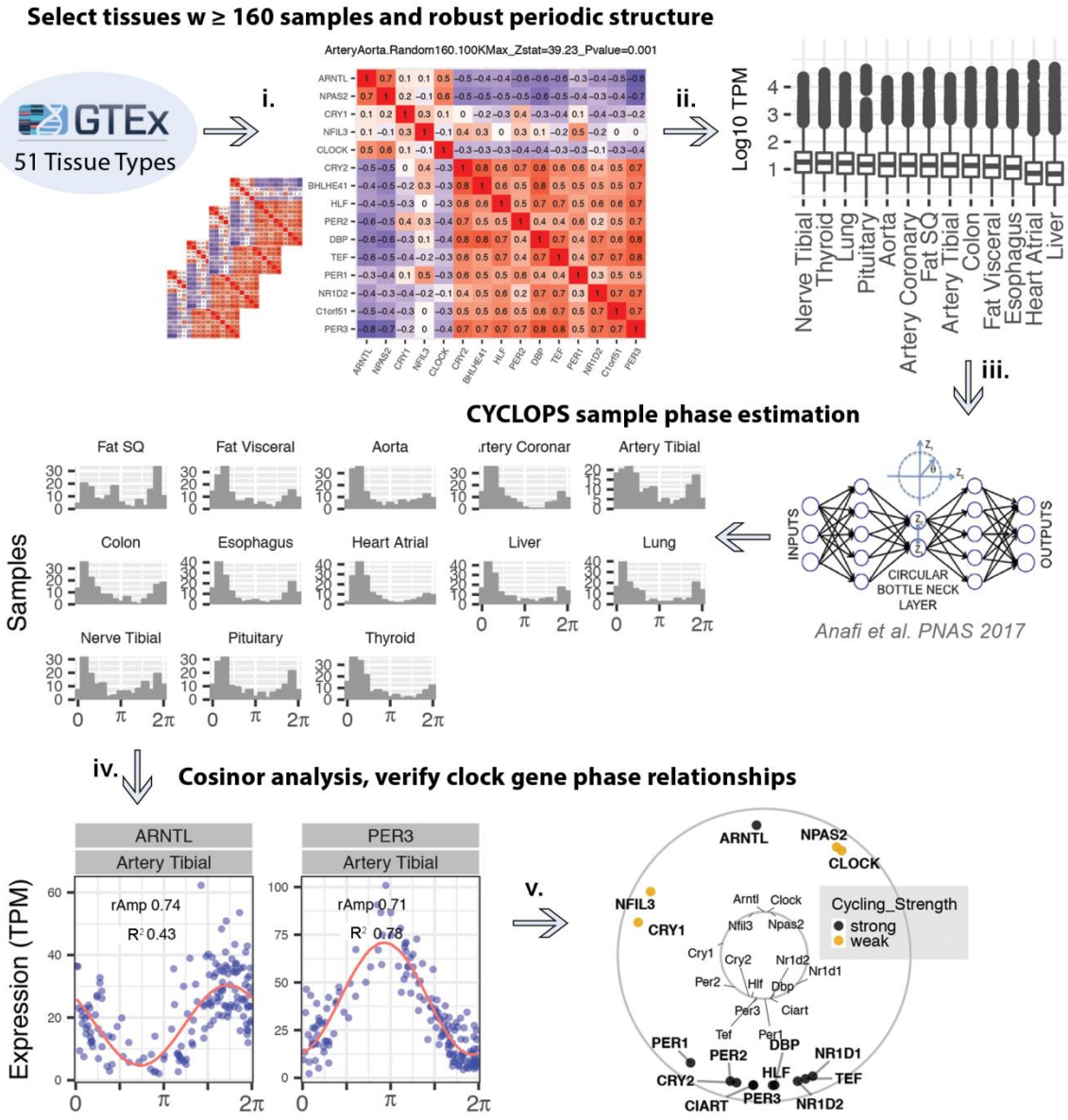

Fig. S1. Pipeline for sample selection and CYCLOPS ordering for 13 tissues. From the GTEx collection, we selected tissues with at least 160 samples *and* evidence of intact molecular clocks. When the mammalian circadian clock is progressing normally, core clock genes show characteristic temporal sequences in expression *(3, 17, 51)*. **i.** To verify presence of hallmark core clock gene phase relationships in each tissue, we calculated pairwise correlation matrices (similar to the approach in *(17)*). This filtering step left us with 13 tissue types. **ii.** To further

enrich for samples with robust intact clocks at time of death, we selected a subset of 160 samples for each tissue that yielded the strongest clock gene correlation matrix following 100K random samplings. This favored sample sets that were on average younger, died suddenly, and had shorter times between death and tissue stabilization. Despite this, tissue sets retained extensive diversity (**Table 1**). **iii.** 160 samples from each tissue were CYCLOPS-ordered following parameters in Anafi et al., with deviations described in Methods. **iv.** To test for gene-level rhythms, cosinor regression was run on Transcripts per Million (TPM) values as a function of estimated sample phases for each of the top 15K expressed genes in each tissue. **v.** CYCLOPS-estimated phase relationships were analyzed for known clock genes to validate the quality of sample ordering.

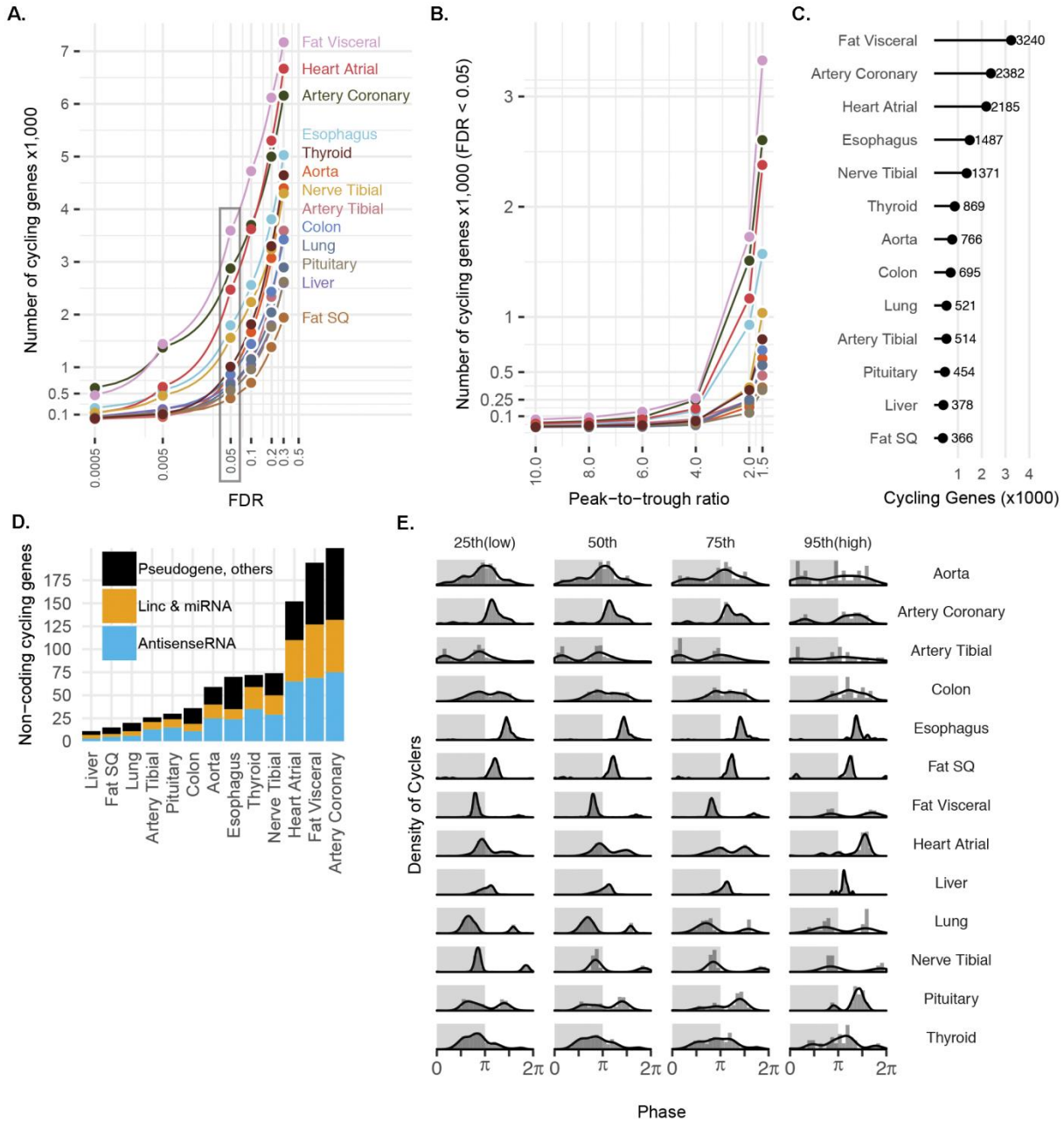

Fig. S2. Numbers of cycling genes by different statistical thresholds and noncoding gene type. (**A**) Cycling genes detected over a series of FDR thresholds, with no boundaries on rAmp or R^2. (**B**) Cycling genes detected over a series of peak-to-trough thresholds, requiring FDR < 0.05. **(C)** Cycler counts at FDR < 0.05, rAmp ≥ 0.1 and R^2 ≥ 0.1. **(D)** Distribution of noncoding cycling genes by class (FDR < 0.05, rAmp ≥ 0.1 and R^2 ≥ 0.1). **(E)** Distribution of peak phases

for cyclers (FDR < 0.05) at four different peak-to-trough ratio (ptr) cutoffs. The cutoffs include cyclers above the: 25th, 50th, 75th, or 95th percentiles for ptr, within each tissue.

Human
Mouse
Heart
Aorta
Lung
Fat
Phase
2π
π
0
CELL_DEATH
CELLULAR_CHEMICAL_HOMEOSTASIS
CELLULAR_HOMEOSTASIS
CELLULAR_RESPONSE_TO_HORMONE_STIMULUS
CHEMICAL_HOMEOSTASIS
DEPHOSPHORYLATION
ENDOCYTOSIS
IMMUNE_SYSTEM_DEVELOPMENT
ION_TRANSPORT
NEGATIVE_REGULATION_OF_INTRACELLULAR_SIGNAL_TRANSDUCTION
NEGATIVE_REGULATION_OF_MULTICELLULAR_ORGANISMAL_PROCESS
NEGATIVE_REGULATION_OF_PHOSPHATE_METABOLIC_PROCESS
NEGATIVE_REGULATION_OF_PHOSPHORUS_METABOLIC_PROCESS
NEGATIVE_REGULATION_OF_PHOSPHORYLATION
ORGANOPHOSPHATE_METABOLIC_PROCESS
PEPTIDYL_AMINO_ACID_MODIFICATION
POSITIVE_REGULATION_OF_TRANSCRIPTION_FROM_RNA_POLYMERASE_II_PROMOTER
PROTEIN_PHOSPHORYLATION
REGULATION_OF_ION_TRANSPORT
RESPONSE_TO_LIPID
CARBOHYDRATE_METABOLIC_PROCESS
CARDIOVASCULAR_SYSTEM_DEVELOPMENT
CATABOLIC_PROCESS
CELLULAR_CATABOLIC_PROCESS
CELLULAR_RESPONSE_TO_ENDOGENOUS_STIMULUS
CELLULAR_RESPONSE_TO_ORGANIC_SUBSTANCE
CIRCULATORY_SYSTEM_DEVELOPMENT
ESTABLISHMENT_OF_PROTEIN_LOCALIZATION
HEART_DEVELOPMENT
ION_TRANSPORT
MITOCHONDRION_ORGANIZATION
MRNA_METABOLIC_PROCESS
NITROGEN_COMPOUND_TRANSPORT
ORGANONITROGEN_COMPOUND_BIOSYNTHETIC_PROCESS
ORGANONITROGEN_COMPOUND_METABOLIC_PROCESS
PHOSPHATE_CONTAINING_COMPOUND_METABOLIC_PROCESS
PROTEIN_LOCALIZATION
REGULATION_OF_ANATOMICAL_STRUCTURE_MORPHOGENESIS
RESPONSE_TO_GROWTH_FACTOR
TRANSMEMBRANE_TRANSPORT
BIOLOGICAL_ADHESION
CELL_MATRIX_ADHESION
CELL_MOTILITY
CELL_SUBSTRATE_ADHESION
ENZYME_LINKED_RECEPTOR_PROTEIN_SIGNALING_PATHWAY
HOMEOSTATIC_PROCESS
LOCALIZATION_OF_CELL
LOCOMOTION
MOVEMENT_OF_CELL_OR_SUBCELLULAR_COMPONENT
PHOSPHORYLATION
POSITIVE_REGULATION_OF_PROTEIN_MODIFICATION_PROCESS
PROTEIN_PHOSPHORYLATION
REGULATION_OF_CELL_ADHESION
REGULATION_OF_CELLULAR_COMPONENT_MOVEMENT
REGULATION_OF_HYDROLASE_ACTIVITY
REGULATION_OF_IMMUNE_SYSTEM_PROCESS
REGULATION_OF_MAPK_CASCADE
RESPONSE_TO_CYTOKINE
RESPONSE_TO_LIPID
RNA_PROCESSING
CELL_MOTILITY
CELLULAR_COMPONENT_DISASSEMBLY
CELLULAR_RESPONSE_TO_HORMONE_STIMULUS
EXTRACELLULAR_STRUCTURE_ORGANIZATION
INTRACELLULAR_SIGNAL_TRANSDUCTION
LOCALIZATION_OF_CELL
NEGATIVE_REGULATION_OF_CELL_COMMUNICATION
POSITIVE_REGULATION_OF_CATALYTIC_ACTIVITY
POSITIVE_REGULATION_OF_MOLECULAR_FUNCTION
POSITIVE_REGULATION_OF_RESPONSE_TO_STIMULUS
POSITIVE_REGULATION_OF_TRANSCRIPTION_FROM_RNA_POLYMERASE_II_PROMOTER
REGULATION_OF_CELL_DIFFERENTIATION
REGULATION_OF_CELLULAR_RESPONSE_TO_STRESS
REGULATION_OF_DEFENSE_RESPONSE
REGULATION_OF_HYDROLASE_ACTIVITY
REGULATION_OF_INTRACELLULAR_SIGNAL_TRANSDUCTION
REGULATION_OF_MULTICELLULAR_ORGANISMAL_DEVELOPMENT
REGULATION_OF_PROTEIN_SERINE_THREONINE_KINASE_ACTIVITY
REGULATION_OF_RESPONSE_TO_STRESS
TISSUE_DEVELOPMENT

Fig. S3. Phase set enrichment in human and mouse counterpart tissues. PSEA was run on rhythmically-expressed genes in human (FDR < 0.05, rAmp ≥ 0.1, $R^2 \geq 0.1$) and mouse-matched (JTK Q-value < 0.05) tissues. Peak phases of expression are plotted for the top 20 pathways that were identified as phase-enriched (Kuiper's test, p-value ≤ 0.05) in both species. Phase of *ARNTL* (human) or *Arntl* (mouse) was set to 0 for comparison. Human:mouse tissue comparisons were heart-atrium:whole heart; aorta:aorta; lung:lung; visceral fat:white fat.

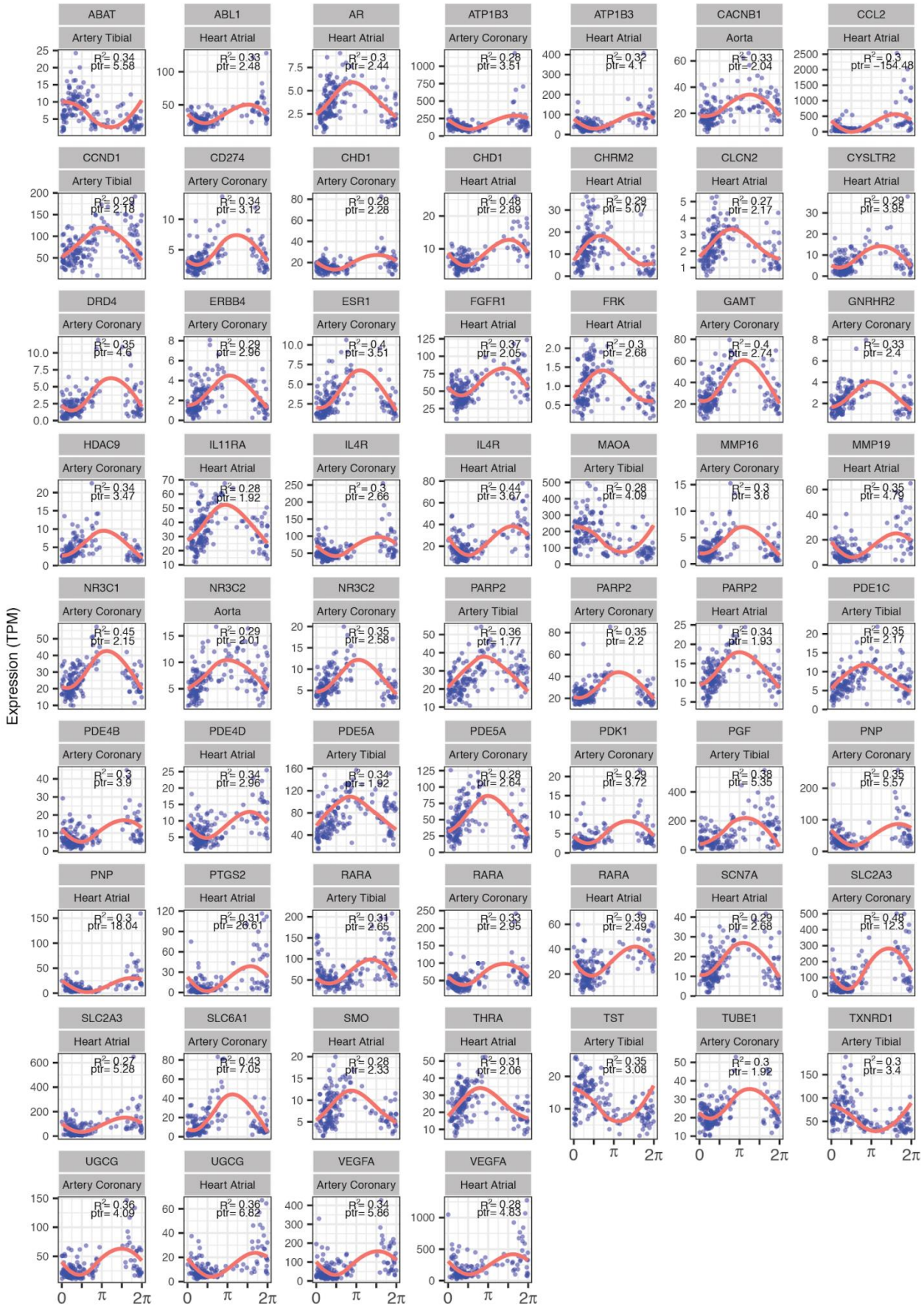
Expression (TPM)
ABAT
Artery Tibial
R2= 0.34
ptr= 5.58
ABL1
Heart Atrial
R2= 0.33
ptr= 2.48
AR
Heart Atrial
R2= 0.3
ptr= 2.44
ATP1B3
Artery Coronary
R2= 0.28
ptr= 3.51
ATP1B3
Heart Atrial
R2= 0.32
ptr= 4.1
CACNB1
Aorta
R2= 0.33
ptr= 2.04
CCL2
Heart Atrial
R2= 0.3
ptr= −154.48
CCND1
Artery Tibial
R2= 0.29
ptr= 2.19
CD274
Artery Coronary
R2= 0.34
ptr= 3.12
CHD1
Artery Coronary
R2= 0.28
ptr= 2.28
CHD1
Heart Atrial
R2= 0.48
ptr= 2.89
CHRM2
Heart Atrial
R2= 0.29
ptr= 5.07
CLCN2
Heart Atrial
R2= 0.27
ptr= 2.17
CYSLTR2
Heart Atrial
R2= 0.29
ptr= 3.95
DRD4
Artery Coronary
R2= 0.35
ptr= 4.6
ERBB4
Artery Coronary
R2= 0.29
ptr= 2.96
ESR1
Artery Coronary
R2= 0.4
ptr= 3.51
FGFR1
Heart Atrial
R2= 0.37
ptr= 2.05
FRK
Heart Atrial
R2= 0.3
ptr= 2.68
GAMT
Artery Coronary
R2= 0.4
ptr= 2.74
GNRHR2
Artery Coronary
R2= 0.33
ptr= 2.4
HDAC9
Artery Coronary
R2= 0.34
ptr= 3.47
IL11RA
Heart Atrial
R2= 0.28
ptr= 1.92
IL4R
Artery Coronary
R2= 0.3
ptr= 2.66
IL4R
Heart Atrial
R2= 0.44
ptr= 3.67
MAOA
Artery Tibial
R2= 0.28
ptr= 4.09
MMP16
Artery Coronary
R2= 0.3
ptr= 3.6
MMP19
Heart Atrial
R2= 0.35
ptr= 4.79
NR3C1
Artery Coronary
R2= 0.45
ptr= 2.15
NR3C2
Aorta
R2= 0.29
ptr= 2.01
NR3C2
Artery Coronary
R2= 0.35
ptr= 2.58
PARP2
Artery Tibial
R2= 0.36
ptr= 1.77
PARP2
Artery Coronary
R2= 0.35
ptr= 2.2
PARP2
Heart Atrial
R2= 0.34
ptr= 1.93
PDE1C
Artery Tibial
R2= 0.35
ptr= 2.17
PDE4B
Artery Coronary
R2= 0.3
ptr= 3.9
PDE4D
Heart Atrial
R2= 0.34
ptr= 2.96
PDE5A
Artery Tibial
R2= 0.34
ptr= 1.92
PDE5A
Artery Coronary
R2= 0.28
ptr= 2.64
PDK1
Artery Coronary
R2= 0.29
ptr= 3.72
PGF
Artery Tibial
R2= 0.38
ptr= 5.35
PNP
Artery Coronary
R2= 0.35
ptr= 5.57
PNP
Heart Atrial
R2= 0.3
ptr= 18.04
PTGS2
Heart Atrial
R2= 0.31
ptr= 20.61
RARA
Artery Tibial
R2= 0.31
ptr= 2.65
RARA
Artery Coronary
R2= 0.33
ptr= 2.95
RARA
Heart Atrial
R2= 0.39
ptr= 2.49
SCN7A
Heart Atrial
R2= 0.29
ptr= 2.68
SLC2A3
Artery Coronary
R2= 0.48
ptr= 12.3
SLC2A3
Heart Atrial
R2= 0.27
ptr= 5.28
SLC6A1
Artery Coronary
R2= 0.43
ptr= 7.05
SMO
Heart Atrial
R2= 0.28
ptr= 2.33
THRA
Heart Atrial
R2= 0.31
ptr= 2.06
TST
Artery Tibial
R2= 0.35
ptr= 3.08
TUBE1
Artery Coronary
R2= 0.3
ptr= 1.92
TXNRD1
Artery Tibial
R2= 0.3
ptr= 3.4
UGCG
Artery Coronary
R2= 0.36
ptr= 4.09
UGCG
Heart Atrial
R2= 0.36
ptr= 6.82
VEGFA
Artery Coronary
R2= 0.34
ptr= 5.86
VEGFA
Heart Atrial
R2= 0.28
ptr= 4.83
0
π
2π

Fig. S4. Gene expression traces for top cardio-cycling drug targets. Top 60 cardio-cycling drug targets by amplitude. Coefficient of determination (R^2) and peak-to-trough ratio (ptr) are indicated for each gene. A more permissive FDR threshold (≤ 0.1) was applied here to detect potentially noisier, but high amplitude oscillations.

Data file S1. Cosinor regression output for all genes and all tissues. Provided as a separate excel file.

Data file S2. Set of 54 genes that cycle in 8 or more of the 13 tissues. Provided as a separate excel file.

Data file S3. Set of 136 cardio-cycler drug targets. Provided as a separate excel file.