SUPPLEMENTARY MATERIAL

NUMBER OF EARTH REVOLUTIONS SINCE THE ORIGIN OF LIFE

To interpret the pervasiveness and plasticity of circadian oscillations, we need a rough estimate of the number of times the Earth has rotated on its axis since the origin of life. A quick estimate provides a figure of over *one trillion times* (3.5 x 10^9 x 365 = 1.3 x 10^{12}). The exact number is even larger and close to two trillions because the Earth is constantly losing angular velocity and rotational energy through a process called tidal acceleration. This process leads to a slow lengthening of the day–for instance, 620 million years ago a day had only about 21.9$\pm$0.4 hours, providing further evidence for the plasticity of the *period* of circadian rhythms during evolution.

CIRCADIAN ANALYSES

The circadian analyses were conducted using JTK_CYCLE (Hughes *et al.*, 2010), a program implemented in R that can be used to determine cycling events in gene expression and other time series. A gene was considered circadian, if at least one of its transcripts was found to be circadian by JTK_CYCLE. In brief, this algorithm characterizes samples as rhythmic or nonrhythmic using a nonparametric method based on a combination of the JonckheereTerpstra test for monotonic ordering and Kendall's τ test for association of measured quantities. JTK_CYCLE has been reported to be faster and more accurate than other methods, such as COSOPT (Straume, 2004). JTK_CYCLE handles multiple hypothesis testing in two ways. For each time series, JTK_CYCLE produces both p-values and q values. In addition, as stated in Hughes *et al.* (2010) (page 373) "Each minimal p-value is Bonferronia adjusted for multiple testing and consequently, the adjusted minimal p-values reported by JTK-CYCLE are uniformly conservative, i.e., they are never lower than the empirical p-values (Suppl. Fig. S1). ". For circadian analyses, JTK_CYCLE is used with parameters set to a period of 24h with typical default values of $q = 0.05$ and $p = 0.05$. The q value can be adjusted depending on the number of measurements available. Here, for robustness testing, we also varied the p-value cutoff by repeating the entire transcriptome and metabolome analyses at $p = 0.05$, $p = 0.01$, and $p = 0.005$. In combination, these analyses show that the main result is robust: a significant fraction of the transcriptome and the metabolome is capable of circadian oscillations under at least one set of conditions.

ANALYSES OF TRANSCRIPTOMES AND METABOLOMES

Transcriptomic and metabolomic datasets analyzed in this paper are listed in Supplementary Table S.1. Pairwise comparisons of the circadian transcriptomes across experiments carried on *liver only* (10 experiments) are given in Figure S.1. Pairwise comparison of the circadian transcriptomes, determined with a more stringent circadian p-value of 0.01, across all 18 experiments are given in Figure S.2. Pairwise comparisons of the circadian metabolomes, determined with a more stringent circadian p-value of 0.01, across all 10 experiments are given in Figure S.3. Pairwise comparison of the circadian transcriptomes, determined with an even more stringent circadian p-value of 0.005, across all 18 experiments are given in Figure S.4. Pairwise comparisons of the circadian metabolomes, determined with an even more stringent circadian p-value of 0.005, across all 10 experiments are given in Figure S.5.

Table S.1. List of transcriptomic and metabolomic datasets analyzed.

Tissue/Condition	Strain	Transcriptome	Metabolome	Reference
Aorta	unknown	Yes	No	Rudic et al, 2004
Gut Normal-Chow	C57BL/6J	No	Yes	Tognini, Murakami et al, 2014
Gut Ketogenic-Diet	C57BL/6J	No	Yes	Tognini, Murakami et al, 2014
Liver	C57BL/6J	Yes	No	Hughes et al, 2009
Liver	C57BL/6J	Yes	No	Panda et al, 2002
Liver *Clock* WT	C57BL/6J	Yes	No	Miller et al, 2007
Liver *Clock* Mutant	C57BL/6J Clock homozygous mutant	Yes	No	Miller et al, 2007
Liver Normal-Chow	C57BL/6J	Yes	Yes	Eckel-Mahan et al, 2013
Liver High-Fat Diet	C57BL/6J	Yes	Yes	Eckel-Mahan et al, 2013
Liver *Sirt1* WT	Mostly C57/B6 with some Black Swiss	Yes	Yes	Masri et al, 2014
Liver *Sirt1* KO	Mostly C57/B6 with some Black Swiss - *Sirt1* knockout	Yes	Yes	Masri et al, 2014
Liver *Sirt6* WT	Mixed C57/B6 and Black Swiss	Yes	Yes	Masri et al, 2014
Liver *Sirt6* KO	Mixed C57/B6 and Black Swiss - *Sirt6* knockout	Yes	Yes	Masri et al, 2014
Muscle *Bmal1* WT	Cre-negative littermates from cross between C57BL/6 with floxed *Bmal1* and C57BL/6 mouse carrying a Cre recombinase transgene.	Yes	Yes	Dyar et al, 2013
Muscle *Bmal1* KO	Cross between C57BL/6 with floxed *Bmal1* and C57BL/6 mouse carrying a Cre recombinase transgene.	Yes	Yes	Dyar et al, 2013
NIH3T3	C57BL/6J	Yes	No	Hughes et al, 2009
Serum Normal-Chow	C57BL/6J	No	Yes	Abbondante et al, 2014
Serum High-Fat Diet	C57BL/6J	No	Yes	Abbondante et al, 2014
SCN	C57BL/6J	Yes	No	Panda et al, 2002
Skeletal Muscle	C57BL/6J	Yes	No	Andrews et al, 2010

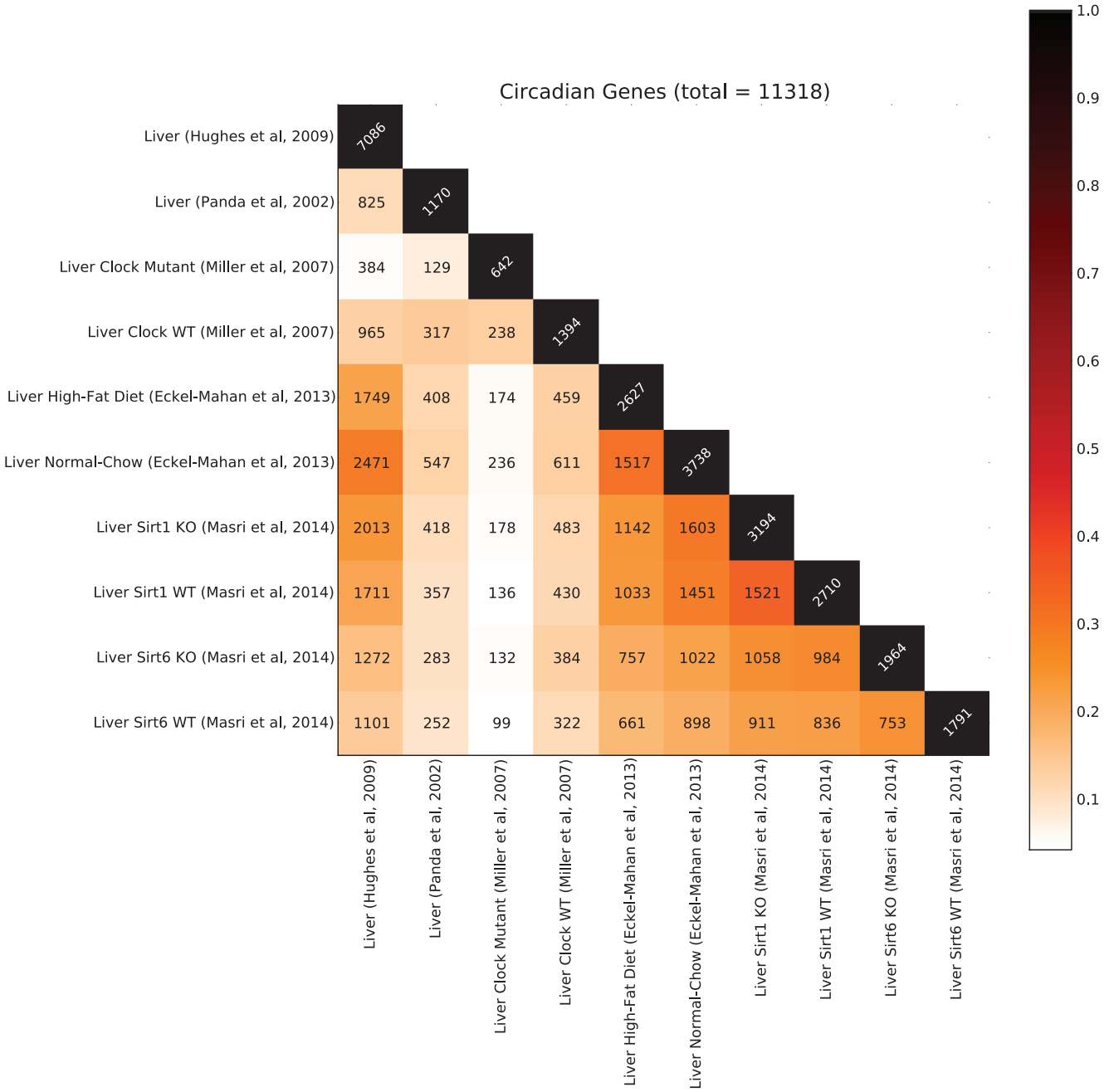

Fig. S.1. Pairwise comparison matrix across all experiments performed on liver tissue only. The numbers correspond to the number of protein coding genes with at least one circadian transcript ($p \leq 0.05$) that are common to both perturbations (i.e. $|A \cap B|$). The color intensity corresponds to the Tanimoto-Jaccard index ($|A \cap B|/|A \cup B|$). In total, there are 11318 ($\sim$56%) protein coding genes that can produce circadian transcripts in at least one condition.

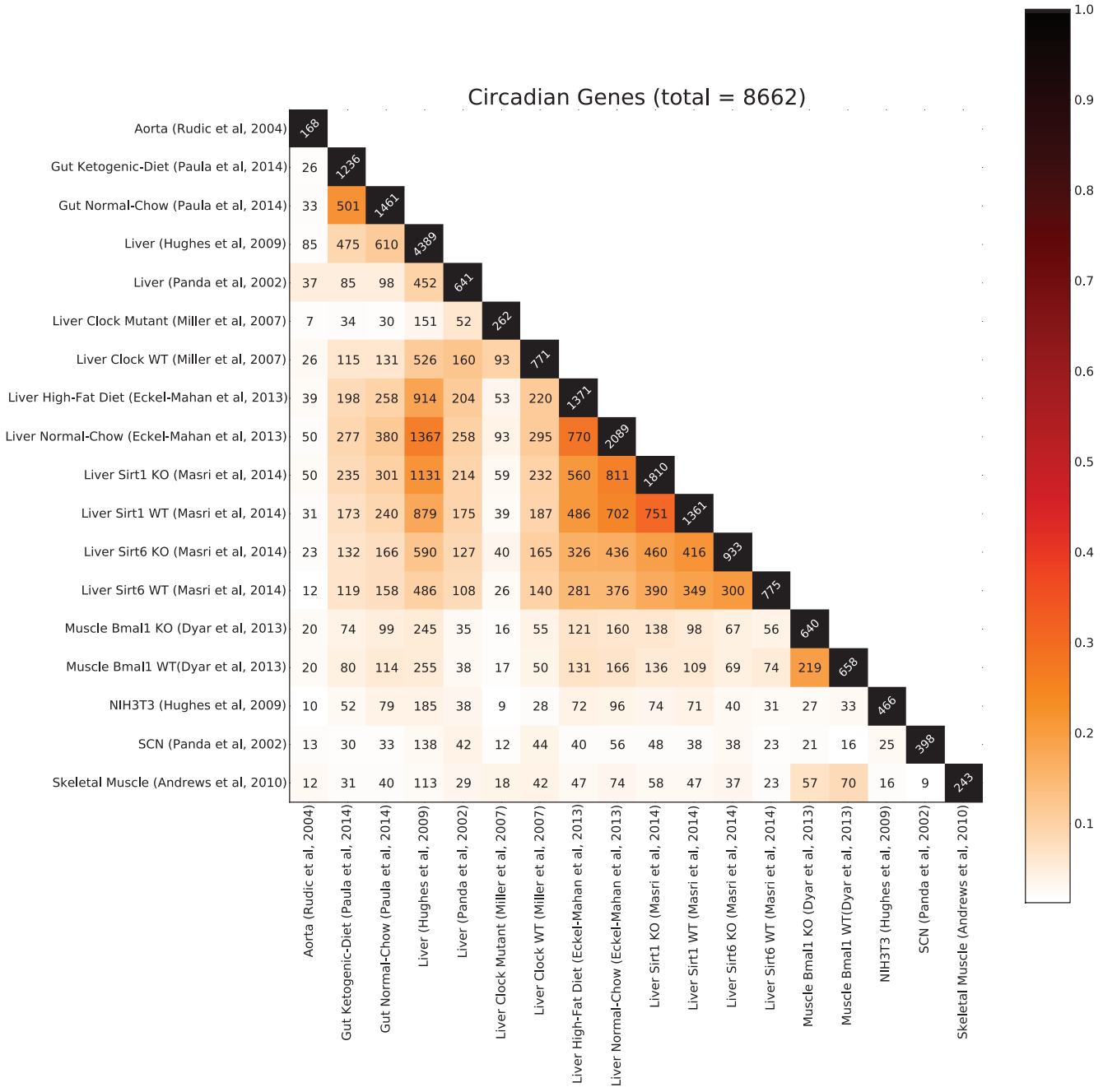

Fig. S.2. Pairwise comparison matrix across 18 transcriptomic experiments with more stringent oscillatory cutoff. The numbers correspond to the number of protein coding genes with at least one circadian transcript ($p \leq 0.01$) that are common to both tissues/conditions (i.e. $|A \cap B|$). The color intensity corresponds to the Tanimoto-Jaccard index ($|A \cap B|/|A \cup B|$). In total, there are 8662 ($\sim$43%) protein coding genes with cam produce circadian transcripts in at least one tissue or condition.

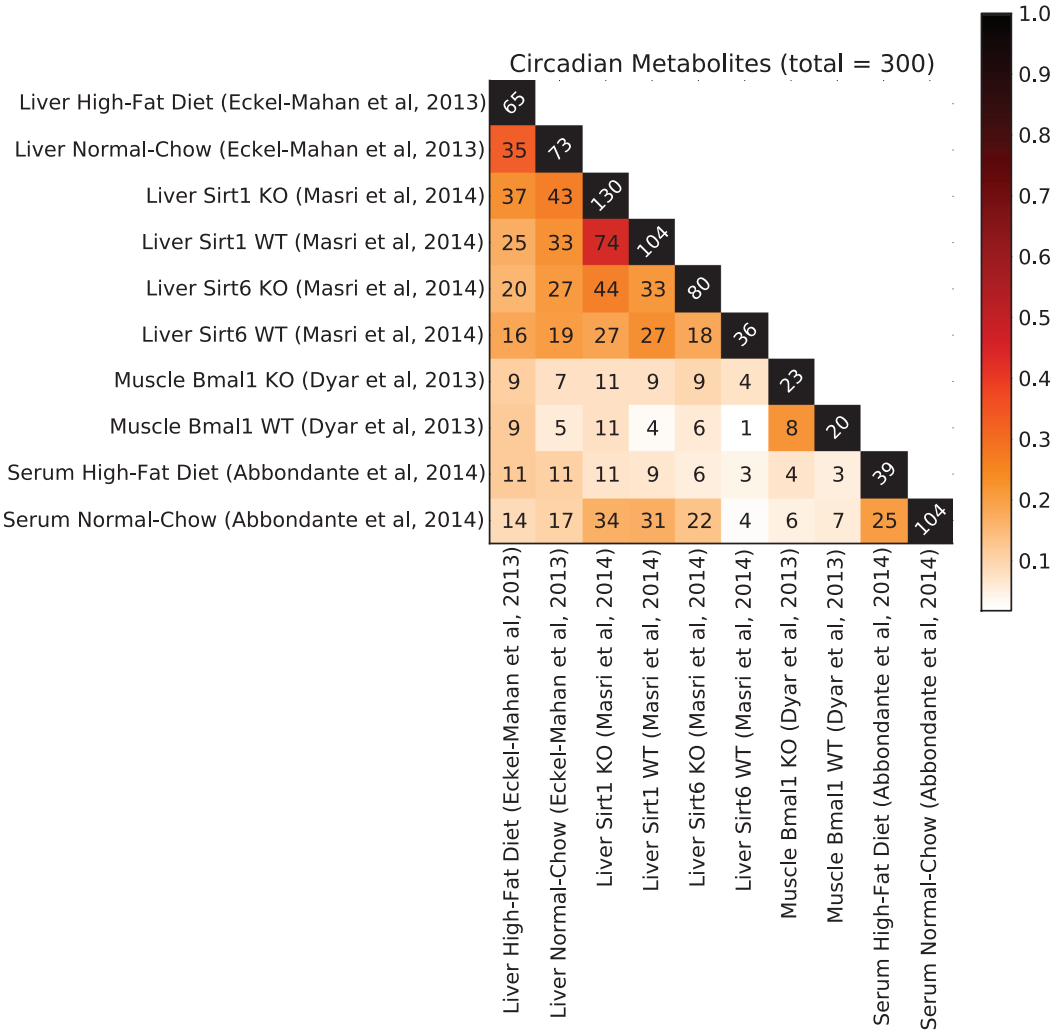

Fig. S.3. Pairwise comparison matrix across 10 metabolomic experiments with more stringent oscillatory cutoff. The numbers correspond to the number of oscillating metabolites ($p \leq 0.01$) that are common to both tissues/conditions (i.e. $|A \cap B|$). The color intensity corresponds to the Tanimoto-Jaccard index ($|A \cap B|/|A \cup B|$). In total, there are 300 ($\sim$54%) measured metabolites that oscillate in at least one tissue or condition in a circadian manner.

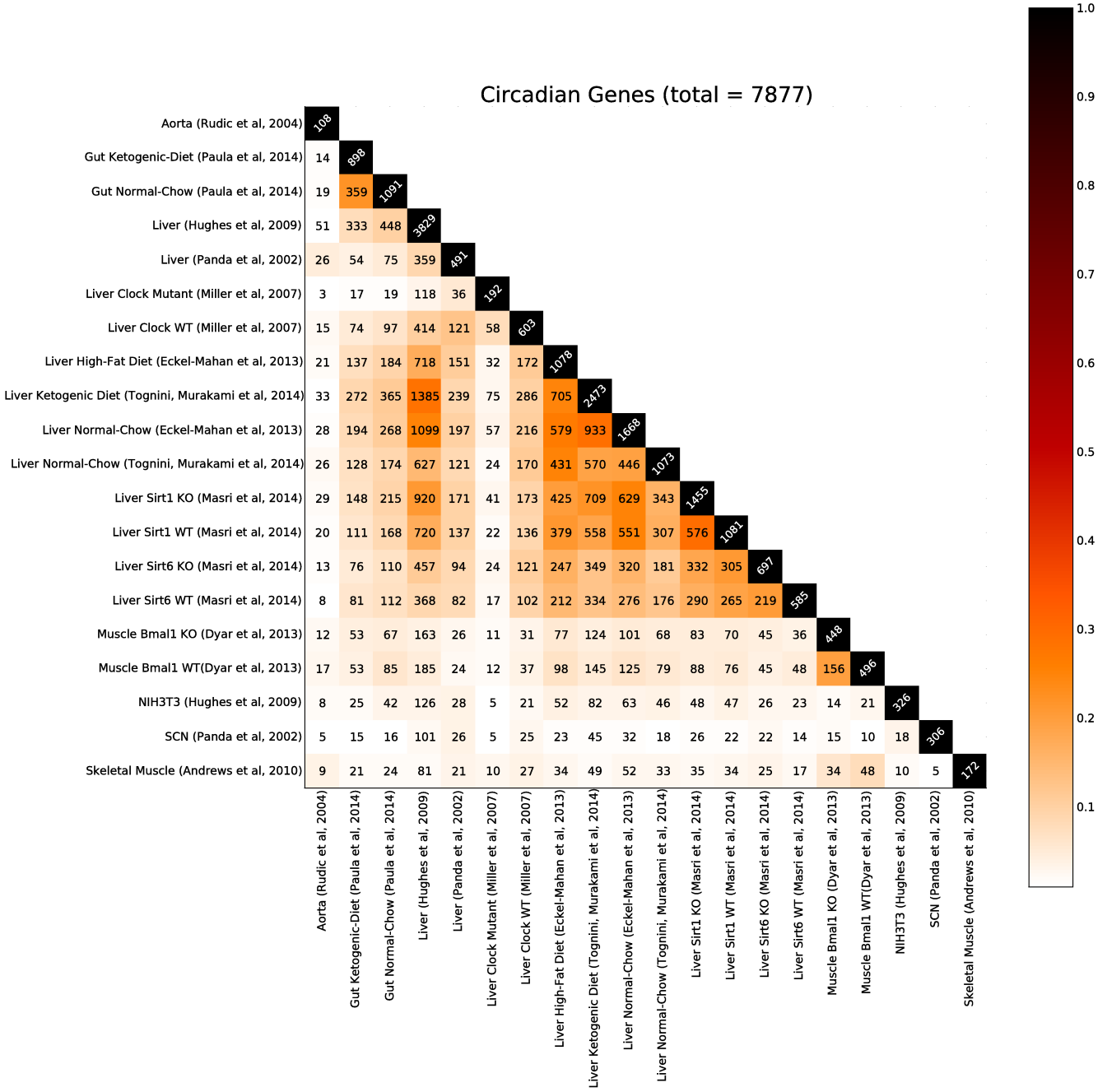

Fig. S.4. Pairwise comparison matrix across 18 transcriptomic experiments with even more stringent oscillatory cutoff. The numbers correspond to the number of protein coding genes with at least one circadian transcript ($p \leq 0.005$) that are common to both tissues/conditions (i.e. $|A \cap B|$). The color intensity corresponds to the Tanimoto-Jaccard index ($|A \cap B|/|A \cup B|$). In total, there are 7877 (∼39%) protein coding genes that can produce circadian transcripts in at least one tissue or condition.

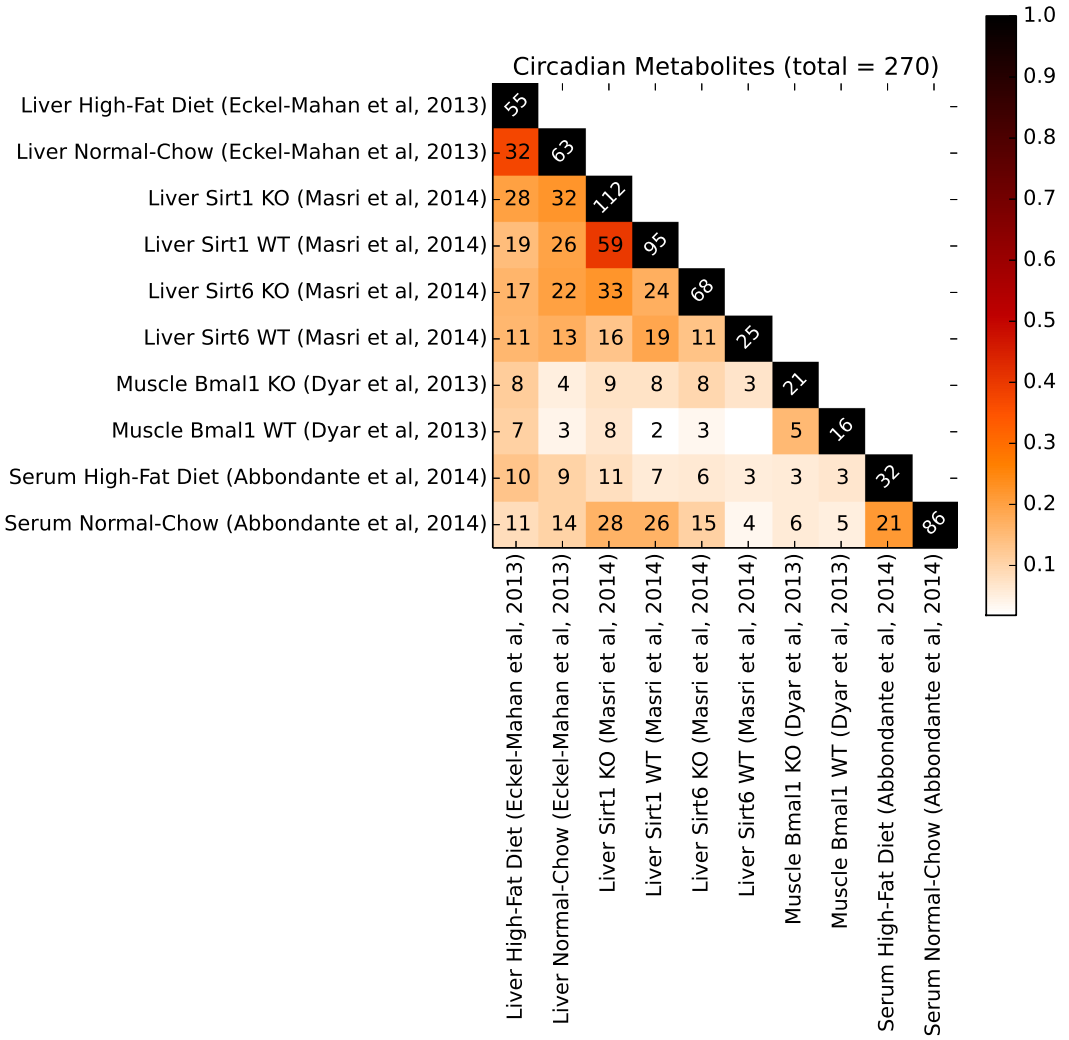

Fig. S.5. Pairwise comparison matrix across 10 metabolomic experiments with more stringent oscillatory cutoff. The numbers correspond to the number of oscillating metabolites ($p \leq 0.005$) that are common to both tissues/conditions (i.e. $|A \cap B|$). The color intensity corresponds to the Tanimoto-Jaccard index ($|A \cap B|/|A \cup B|$). In total, there are 270 ($\sim$47%) measured metabolites that oscillate in at least one tissue or condition in a circadian manner.

NETWORK ANALYSIS

Using CircadiOmics, we constructed a network containing only regulatory (transcriptional) edges and protein-protein interaction edges to get rough estimates of important network properties, in particular the number of loops of various types and the centrality of the clock hub. The network consisted of 21826 genes/proteins, with 120988 edges. There were 114493 regulatory edges and 6495 protein-protein interactions (only physical interactions were considered). The diameter of the network is 8. Regulatory edges are unidirectional, while protein-protein interaction edges are considered bi-directional.

The clock machinery is centrally located and capable of potentially acting on a large fraction of the genome. We estimated the network-distance between *Clock* or *Bmal1* and all other gene/proteins. We computed these on three different networks. First, the network where the direction of the edges is ignored. Second, the network with uni-directional regulatory edges and bi-directional protein-protein interactions. And third, the network containing only the uni-directional regulatory edges. In all cases, it can be seen in Table S.2 that $\sim$10% of genes are one hop away from *Clock*/*Bmal1* and $\sim$60-70% genes are two hops away.

Table S.2. Network distances from *Clock*/*Bmal1*.

Distance	# Gens in undirected network	# Genes in directed network with PPI	# Genes in directed network without PPI
1	2300	2293	2286
2	15249	13744	13339
3	758	1750	1035
4	17	437	50
5	1	2	-

In general, molecular species in isolation do not oscillate. Oscillations require directed loops of interactions (cycles). Cycles were counted by enumerating all paths, with no repeated nodes, which start and end at *Clock*/*Bmal1*. The numbers are given in Table S.3.

Table S.3. Number of cycles containing *Clock* or *Bmal1*.

Cycle length	# Cycles with *Clock*	# Cycles with *Bmal1*
2	10	14
3	73	90
4	1007	1097
5	15512	15641
6	260973	253615
7	4570219	4324732

FORMAL MODELS OF COUPLED OSCILLATORS

Detailed mathematical models for specific molecular oscillators have been developed (Goldbeter, 1995, 1997). While useful, these are not capable of providing a system view of circadian oscillations, or making system level predictions. Furthermore, even the basic clock is not fully understood in all its biochemical details. For example, many details of the composition and mode of action of the inhibiting complex containing the CRY and PER proteins remain to be elucidated.

Thus one may want to consider more global models of coupled oscillators. Arrays of coupled-oscillators have been studied in physics and other areas (Baldi and Meir, 1990; Strogatz, 2000; Goel and Ermentrout, 2002; Brandt *et al.*, 2006). A fairly general class of models can be written in the form

$$\frac{\partial \theta_i}{\partial t} = \omega_i + K \sum_i^n f_i(\theta_i, \theta_j) \tag{1}$$

where θ_i is the phase of the *i*-th oscillator, ω_i represents its frequency, and f_i is the coupling function. For instance, in the well-studied Kuramoto model the coupling is given by $f(\theta_i, \theta_j) = sin(\theta_j - \theta_i)$. However such a model seem too simple and homogeneous in order to properly model the molecular oscillators described in this paper which are not homogeneous nor regularly organized on some kind of lattice.

Other relatively simple models could use Boolean functions or neural networks with Hill-like functions (Baldi and Atiya, 1989; Scheper *et al.*, 1999; Akman *et al.*, 2012) to model the state or concentration of a molecular species as a function of its interacting neighbors. For instance, the concentration y_i of species i could be a non-linear function of its activation x_i with

$$y_i = f(x_i) = \frac{1}{1 + ce^{-\lambda x_i}} \quad \text{and} \quad \frac{dx_i}{dt} = \frac{-x_i}{\tau_i} + \sum_j w_{ij} y_j \tag{2}$$

here τ_i is the decay time-constant of species *i* and w_{ij} is a matrix of weights capturing the interactions with neighboring species. However it is not clear whether enough data is available to fit the parameters of such models reliably, notwithstanding that many molecular mechanisms are far more subtle than Equation 2.

REFERENCES

Akman, O. E., Watterson, S., Parton, A., Binns, N., Millar, A. J., and Ghazal, P. (2012). Digital clocks: simple boolean models can quantitatively describe circadian systems. *J R Soc Interface*, **9**(74), 2365–82.

Baldi, P. and Atiya, A. (1989). Oscillations and synchronizations in neural networks: An exploration of the labeling hypothesis. *International Journal of Neural Systems*, **01**(02), 103–124.

Baldi, P. and Meir, R. (1990). Computing with arrays of coupled oscillators: an application to preattentive texture discrimination. *Neural Computation*, **2**(4), 458–471.

Brandt, S., Dellen, B., and Wessel, R. (2006). Synchronization from disordered driving forces in arrays of coupled oscillators. *Physical review letters*, **96**(3), 34104.

Goel, P. and Ermentrout, B. (2002). Synchrony, stability, and firing patterns in pulse-coupled oscillators. *Physica D: Nonlinear Phenomena*, **163**(3), 191–216.

Goldbeter, A. (1995). A model for circadian oscillations in the drosophila period protein (per). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **261**(1362), 319–324.

Goldbeter, A. (1997). Biochemical oscillations and cellular rhythms: the molecular bases of periodic and chaotic behaviour.

Hughes, M. E., Hogenesch, J. B., and Kornacker, K. (2010). JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms*, **25**(5), 372–80.

Scheper, T., Klinkenberg, D., Pennartz, C., and van Pelt, J. (1999). A mathematical model for the intracellular circadian rhythm generator. *J Neurosci*, **19**(1), 40–7.

Straume, M. (2004). Dna microarray time series analysis: automated statistical assessment of circadian rhythms in gene expression patterning. *Methods in enzymology*, **383**, 149.

Strogatz, S. (2000). From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, **143**(1), 1–20.