Supporting Information

Hurley et al. 10.1073/pnas.1418963111

SI Text

SI Materials and Methods

Primer Design and Synthesis of Promoter Cassettes and Strains. Primer3 was customized for the *Neurospora* genome by Peter Andrews, Institute for Quantitative Biomedical Sciences, Geisel School of Medicine, Hanover, NH, to retrieve regions adjacent to each start/stop codon of each gene, with the program searching around 1,500-bp region upstream (5′) of the translational start codon and around 500-bp region downstream (3′) of gene stop codon. For some genes, a gap range of no greater than 100 bp from start/stop codon was added in order for the selection of appropriate sequences. For 5′ fragments bigger than 1,500 bp, the 5f (forward) primer was designed by hand. 5f+5r primers and 3f+3r primers PCR synthesized 5′ and 3′ flanks, respectively, using WT gDNA as template. Flanking primers had additional sequences for 1,000 bp of cyclophilin locus (5f and 3r) and the luciferase (5r and 3f) cassette (1). All *Neurospora* strains were tested on a luminometer (Turner TD-20e) for signal higher than 0.5, indicating luminescence.

Time Course Preparation. Tissue was prepared for rhythmic time-series analysis as previously described (2). Total RNA was extracted by adding 1 mL of TRIZOL reagent (Invitrogen) to the frozen tissue and then pulverizing it using a homogenizer (Tissuelyzer; Qiagen) containing stainless beads. The tissue was separated from the supernatant by rapid centrifugation (16,000 × *g* for 15 min). The RNA was further purified and DNase treated using the RNeasy Mini Kit from Qiagen. The RNA was quantified using the Quant-It RiboGreen Kit from Invitrogen according to the manufacturer's protocol.

Library Construction for RNA-Seq. Plate-based sample prep was performed on the PerkinElmer Sciclone NGS robotic liquid handling system, where purified mRNA was converted into cDNA library templates of known strand origin for sequencing on the Illumina Sequencer platform. The first step was the purification of mRNA in the total RNA sample using Invitrogen's DynaBeads mRNA purification kit, which uses magnetic beads with bound poly-T oligos to select and purify poly-A–containing mRNA molecules. Following purification, the mRNA was chemically fragmented with 10× fragmentation solution (Ambion) at 70 °C for 3 min to generate fragments ranging in size between 250 and 300 bp. The fragmented RNA was then purified using AMpure SPRI beads (Agencourt) using a ratio of 160:100 beads volume to RNA. The RNA fragments were then synthesized into first-strand cDNA using SuperScript II Reverse Transcriptase (Invitrogen) and random hexamer primers (MBI Fermentas) with Actinomycin D in the master-mix to further ensure strand specificity by preventing spurious DNA dependent synthesis during the first-strand cDNA synthesis. First-strand synthesis thermocycler conditions were incubation at 42 °C for 50 min and then inactivation at 70 °C for 10 min. This was followed by another purification using AMpure SPRI beads at a ratio of 140/100 beads volume to cDNA. Next, second-strand cDNA was synthesized using DNA Polymerase I (Invitrogen), RNase H (Invitrogen), and a nucleotide mix containing dUTP (Roche) that incorporates dUTP in place of dTTP, which ensures strand specificity, in the second strand to generate ds cDNA. Second-strand synthesis thermocycler conditions were incubation at 16 °C for 60 min. The ds cDNA fragments were purified using AMpure SPRI beads at a ratio of 75/100 beads volume to cDNA and a second purification using a ratio of 140/100. Using a DNA sample prep kit designed for Illumina sequencing, the purified ds cDNA fragments were then end-repaired to be blunt-ended, phosphorylated for the addition of a single A base tail, and finally ligation of the Illumina adapters, which includes unique barcode sequences that allow for the multiplexing of the sample libraries in sequencing. To ensure the strand specificity of the cDNA library templates, AmpErase uracil *N*-glycosylase (UNG) (Applied Biosystems) was added to the ds cDNA library fragments and they were incubated at 37 °C for 15 min to cleave and degrade the strand containing dUTP. The ss cDNA was then enriched using 10 cycles of PCR with Illumina TruSeq primers and purified using AMpure SPRI beads at a ratio of 90/100 beads volume to ds cDNA to create the final cDNA library ready for sequencing. The libraries were created in plate format (rather than tube) using a robot. Once libraries were created then they were quantified, pooled, and run on the Hi-seq instrument. The amount of starting RNA material used for each library was 5 μg. Sequencing libraries were diluted to the specifications required by the Illumina/Solexa sequencing machine in use. Manuals and protocols that describe these steps are available from Illumina/Solexa. The *N. crassa* RNA-Seq data are displayed on JGI MycoCosm portal [genome.jgi.doe.gov/Ncrasss (3)] and have been deposited to NCBI SRA under accession number SRP046458.

Data Analysis. Reads from each RNA-Seq sample were aligned to the reference transcriptome using BWA (parameters used: seed length = 25; maximum hits = 1). The reads were first trimmed to 36 bp, to ensure that reads with a low quality 3′ end will still align. Finally, using the alignment, the number of reads that aligned to each gene was counted, for each sample. Only uniquely mapping reads were counted. The counts were then normalized using reads per kilobase of exon model per million mapped reads (RPKM) normalization. This normalizes both by the number of mapped reads for each library, and by the length for each transcript in the reference file. In other words:

$$\mathrm{RPKM} = \mathrm{raw_count} * (1,000/\mathrm{gene_length}) * (1,000,000/\mathrm{num_mapped_reads}).$$

The data were subjected to quantile normalization to force all expression profiles to have the exactly same distribution and were then log transformed to eliminate outlying points. To generate a list of rhythmically expressed genes, after quantile normalization and log transformation, time points that did not pass quality control were replaced as follows: Time Course 1—DD30 was replaced by DD30 from Time Course 2; Time Course 2—DD28 was replaced by DD28 from Time Course 1; Time Course 3—DD44, DD46, and DD48 were replaced by DD44, DD46, and DD48 from Time Course 2. The data were then smoothed by the following formula: replace expression at t by averaging $t-1, t, t+1$; the exceptions were DD2 (averaging DD2, DD4, and DD44); DD48 (averaging DD46, DD48, and DD6) to correct for the CT time. After the data were processed, each individual time course was subjected to JTK cycle analysis (4). Any gene with a false-discovery rate of $Q < 0.05$ was selected as a rhythmic gene and Venn diagram analysis was used to determine which genes were rhythmic in at least two of the three datasets. To generate heat maps, the three datasets were averaged, and then this single time course was smoothed by the above method and subjected to JTK cycle analysis (4).

DREME analysis (5) was used to identify motifs that were significantly enriched in the promoters of the rhythmic genes compared with the rest of the genome. Briefly, analysis of 1,000 bp of the

promoters of each of the 872 genes identified 36 potential motifs. To test for significant enrichment of these motifs in our rhythmic gene set, we first used a Find Individual Motif Occurrences (FIMO) algorithm (6) to find genes in the entire genome that contained these motifs in their promoters. Typically an individual motif was found in several hundred promoters. Then, at each CT, the number of genes from the rhythmic dataset peaking at that time ±2 h, and that contained a particular motif, was mapped on circular scale. A motif was considered enriched when the number of genes from among the 872, both peaking within a 4-h window and containing the motif, was significantly greater than the number that would be expected from a random distribution of the motif-containing genes in the entire genome. Enrichment of rhythmic genes whose promoters contain a specific motif peaking at an explicit time compared with all of the genes whose promoters contain that motif was examined and a position weight motifs (PWM) probability value was assigned. For every motif, the enrichment was examined at all 24 time points. Then PWMs with minimum P values of <0.001 (out of 24 time points) were selected and a plot was drawn to show the significance of enrichment at each time point.

Real-Time CCD Recording. Liquid nitrogen or electronically cooled cameras from Roper Scientific (VersArray 1300 B/LN and PIXIS 1020) were used to follow luminescence. Camera runs were done at 25 °C with cultures grown in 96-well plates using black opaque tubes. Liquid suspensions of conidia were plated onto race tube (RT) medium (0.1% glucose, 0.17% arginine), maltose medium (0.1% maltose, 0.17% arginine), and/or QA medium (0.03% glucose, 0.05% arginine, 0.01 M QA, pH 5.75) containing 25 μM luciferin (Dataset S3) and covered with a Breathe-Easy strip (USA Scientific). Plates were then subjected to 48 h of 12:12 dark:light cycle conditions at 25 °C before transferring the plates to constant-dark conditions in the camera setup. Signals were accumulated for either 5 or 10 min every hour, depending on the intensity of the luciferase signal. A customized Excel Macro to process the images was developed (7). Raw data arising from each time series were then processed using custom-written software to detrend the data and then to normalize amplitudes.

Validation of Gene Expression by Real-Time PCR. A total of 1,500 ng from each sample of the RNA extracted as described above was used to prepare cDNA using the SuperScript III First-Strand synthesis kit (Invitrogen) followed by real-time PCR using the Fast SYBR green master mix kit in an ABI 7500 real-time cycler (Applied Biosystems). Primer sequences for the various genes tested were generated using an in-house primer designing software. The primer combinations used are listed in Dataset S12, and the final concentration of primers in the reaction mix was 0.5 μM. The following cycling parameters were used: step 1: 95 °C for 5 min; and step 2: 95 °C for 10 s, 60 °C for 30 s for 40 cycles. Ct values were calculated using software provided by the instrument manufacturer. *rac-1* gene was used as a control noncycling gene for the determination of relative mRNA levels for all genes tested. The relative mRNA levels for each time point were calculated using at least two out of three biological replicates in the case where one of the replicates differed from either of the other two by more than threefold.

ChIP Assay to Identify WC-2 Binding Sites. Conidia from the wild-type strain (OR74A) was germinated and grown in 200 mL of Bird medium at 25 °C in the light, before appropriately timed transfer to dark to get samples ranging from DD4 to DD32—representing a 32-h time course with a 4-h resolution. The *Neurospora* mycelium was cross-linked in the light/dark with 1% paraformaldehyde for 15 min and quenched with 0.125 M glycine for an additional 5 min. ChIP was performed as described previously (8, 9) on 500 mg of tissue with WC-2 antibody with some modifications (10). To improve the sonication, the tissue was probe-sonicated for 8 s at 30% amplitude for three times, and then sonicated with a Biorupter for four times with 30-s on/off cycles for 5 min at high power to get an average chromatin size of 100–500 bp. The sheared chromatin was obtained by centrifuging at $11{,}000 \times g$ for 5 min and incubated with protein A dynabeads and WC-2 antibody overnight at 4 °C with constant rotation. The beads were washed once with wash buffer 1 (20 mM Tris·HCl, pH 8.0, 2 mM EDTA, 1% Triton X-100, 0.1% SDS, 150 mM NaCl), once with wash buffer 2 (20 mM Tris·HCl, pH 8.0, 2 mM EDTA, 0.1% Triton-X-100, 0.1% SDS, 500 mM NaCl), once with wash buffer 3 (10 mM Tris·HCl, pH 8.0, 1 mM EDTA, 1% Nonidet P-40, 1% sodium deoxycholate, 0.25 M LiCl), once with wash buffer 4 (10 mM Tris·HCl, pH 8.0, 1 mM EDTA), and eluted with 0.1 M sodium bicarbonate and 1% SDS, and the cross-link was reversed by incubation at 65 °C for 8 h. The DNA was recovered by Qiagen PCR purification kit. To verify the quality of the ChIP time course, RT-PCR was performed with the purified DNA using the following primers: *frq*-PLRE-F, 5′-atacttgtaggcccgctccccatc-3′; *frq*-PLRE-R, 5′-gctggtcatctcctcagcattttgtc-3′; *frq*-CBOX-F, 5′-caattttgcagcgtcatcggtct-3′; and *frq*-CBOX-R, 5′-tcaagtcaagctcgtacccacatc-3′ using the cycling conditions described before relative enrichment of pLRE and C-box of the *frq* promoter region in the samples was determined as a percentage of input.

1. Gooch VD, et al. (2008) Fully codon-optimized luciferase uncovers novel temperature characteristics of the *Neurospora* clock. *Eukaryot Cell* 7(1):28–37.
2. Loros JJ, Denome SA, Dunlap JC (1989) Molecular cloning of genes under control of the circadian clock in *Neurospora*. *Science* 243(4889):385–388.
3. Grigoriev IV, et al. (2014) MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res* 42(Database issue):D699–D704.
4. Hughes ME, Hogenesch JB, Kornacker K (2010) JTK_CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms* 25(5):372–380.
5. Bailey TL (2011) DREME: Motif discovery in transcription factor ChIP-seq data. *Bioinformatics* 27(12):1653–1659.
6. Grant CE, Bailey TL, Noble WS (2011) FIMO: Scanning for occurrences of a given motif. *Bioinformatics* 27(7):1017–1018.
7. Larrondo LF, Loros JJ, Dunlap JC (2012) High-resolution spatiotemporal analysis of gene expression in real time: In vivo analysis of circadian rhythms in *Neurospora crassa* using a FREQUENCY-luciferase translational reporter. *Fungal Genet Biol* 49(9):681–683.
8. Johnson L, Cao X, Jacobsen S (2002) Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. *Curr Biol* 12(16):1360–1367.
9. Belden WJ, Lewis ZA, Selker EU, Loros JJ, Dunlap JC (2011) CHD1 remodels chromatin and influences transient DNA methylation at the clock gene frequency. *PLoS Genet* 7(7):e1002166.
10. Denault DL, Loros JJ, Dunlap JC (2001) WC-2 mediates WC-1-FRQ interaction within the PAS protein-linked circadian feedback loop of *Neurospora*. *EMBO J* 20(1-2):109–117.

Fig. S1. *Neurospora* time course displays circadian rhythms. (*A*) A diagrammatic explanation of the harvest protocol used to obtain the three time courses. Briefly, conidia were inoculated into Bird medium, grown 24 h, and then discs were cut from the resultant tissue, transferred to individual flasks, and kept in light and dark for times appropriate to yield a 48-h time course. The age of all cultures was the same ±4 h. (*B*) Representative RT-PCR of *frq* mRNA normalized to *rac-1* from one of the three circadian time courses, described with both CT and DD time points.

Fig. S2. Setup of strains for luciferase imaging. (*A*) Schematic of the method used to construct the luciferase reporters used to track the promoter activity oscillations. (*B*) Description of the data acquisition of each strain using a 96-well format.

PNAS PNAS PNAS PNAS

Fig. S3. Luciferase traces of *ccg*s and core clock genes. (*A*) Luciferase traces for three technical replicates for each of three biological replicates of the four core clock genes without detrending and rescaling. (*B*) Luciferase traces acquired for all of the published *ccg*s after detrending and rescaling.

Fig. S4. FunCat analysis demonstrated time-of-day category enrichment. Breakdown of FunCat terms from genes designated rhythmic in this dataset with a heat map of the specific genes. Categories highlighted red and marked with an asterisk are considered significant with a value of $P < 0.05$, blue categories have a value of $P > 0.05$, and green represents genes that are unclassified. *A–F* are the genes from each significant FunCat category: (*A*) Metabolism, (*B*) Protein Fate, (*C*) Protein Synthesis, (*D*) Cell Rescue, (*E*) Protein Synthesis, and (*F*) Cell Cycle/DNA Processing, respectively. FunCat category number noted next to each slice.

A. PYRUVATE METABOLISM

B. STEROID BIOSYNTHESIS

C. BIOSYNTHESIS OF UNSATURATED FATTY ACIDS

Fig. S5. Identification of rhythmic genes in the KEGG pathway. (*A–C*) Diagrammatic representation of the KEGG pathways for, pyruvate metabolism, steroid biosynthesis, and biosynthesis of unsaturated fatty acids [modified from the Kyoto Encyclopedia of Genes and Genomes website (1)]. Each gene identified as rhythmic in our analysis and playing a role in the pathway has been highlighted in red in the pathway map and its peak expression time is noted in blue next to it.

1. Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28(1):27–30.

Gene	Rhythmic by RNA-Seq	Rhythmic by RT-PCR	RNA-Seq correlates with RT-PCR	Rhythmic by Luc	RNA-Seq correlates with luciferase
NCU00285	+	+	✓	+	✓
NCU00289	-	-	✓	+	✗
NCU00499	-	-	✓	+	✗
NCU00582	-	-	✓	+	✗
NCU00855	-	-	✓	+	✗
NCU01528 (*ccg*-7)	-	-	✓	+	✗
NCU01810	-	-	✓	+	✗
NCU01898	-	-	✓	+	✗
NCU02084	-	-	✓	-	✓
NCU02193	-	-	✓	+	✗
NCU02238	-	-	✓	+	✗
NCU02265	+	+	✓	+	✓
NCU03362	-	-	✓	+	✗
NCU03753 (*ccg*-1)	+	+	✓	-	✗
NCU04041	-	-	✓	-	✓
NCU04923	+	-	✗	+	✓
NCU06799	+	-	✗	+	✓
NCU06961	+	+	✓	+	✓
NCU07139	-	-	✓	+	✗
NCU07787 (*ccg*-14)	-	-	✓	-	✓
NCU07817	+	+	✓	+	✓
NCU08141	-	-	✓	+	✗
NCU08457 (*ccg*-2)	+	+	✓	-	✗
NCU08769	+	+	✓	+	✓
NCU08791	-	-	✓	+	✗
NCU08907 (*ccg*-13)	+	+	✓	+	✓
NCU09559 (*ccg*-9)	+	+	✓	-	✗
NCU09686 (*ccg*-8)	-	-	✓	-	✓
NCU09976	+	+	✓	+	✓

Rhythmic in both
Rhythmic in RNA-Seq
Rhythmic in Luc only
Rhythmic in neither

Fig. S6. Weak correlation between RNA-Seq and promoter activation. A list of all of the genes investigated by RNA-Seq, luciferase, and RT-PCR. Genes in red are rhythmic by RNA-Seq and luciferase, in green by RNA-Seq only, and in blue by Luc only and white by neither method.

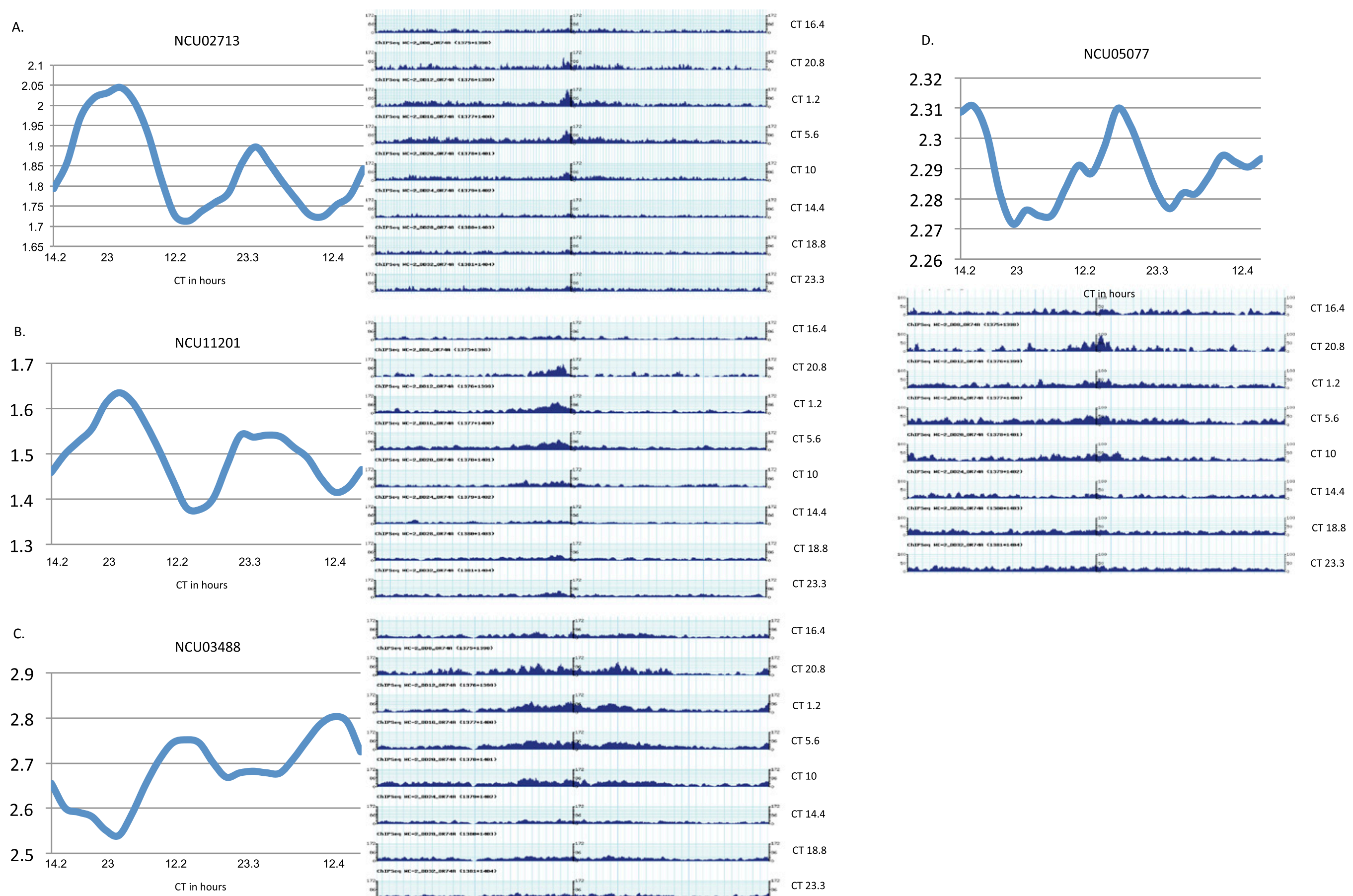

Fig. S7. Comparison of mRNA levels to ChIP-Seq data. (*A–D*) A comparison between the RNA-Seq data and ChIP-Seq data for four genes identified from each dataset. *A* and *B* show peaks in mRNA levels that correspond to the peak in WC-2 binding, whereas *C* and *D* demonstrate that the WC-2 binding peak does not always correlate with the peak of mRNA levels.

Dataset S1. Genes identified as rhythmic by JTK analysis

Dataset S1

Dataset S2. Genes identified as rhythmic by luciferase trace

Dataset S2

Dataset S3. Media used in luciferase assays

Dataset S3

Dataset S4. Genes identified as rhythmic by all methods

Dataset S4

Dataset S5. FunCat term category *P* values

Dataset S5

Dataset S6. Breakdown of each gene and the significance in each FunCat category

Dataset S6

Dataset S7. Genes in significantly enriched FunCat term categories for RNA-Seq and luciferase rhythmic genes

Dataset S7

Dataset S8. Rhythmic genes in each time-of-day–specific circadian category

Dataset S8

Dataset S9. Genes in significantly enriched FunCat term categories for DREME motifs

Dataset S9

Dataset S10. Genes that differ between the RNA-Seq and luciferase data

Dataset S10

Dataset S11. Genes with peaks identified in WC-2 circadian chip

Dataset S11

Dataset S12. Primers used in real-time PCR

Dataset S12