EUKARYOTIC CELL, Oct. 2010, p. 1549–1556 Vol. 9, No. 10
1535-9778/10/$12.00 doi:10.1128/EC.00154-10

Transcription Factors in Light and Circadian Clock Signaling Networks Revealed by Genomewide Mapping of Direct Targets for Neurospora White Collar Complex▽†

Kristina M. Smith,[1]‡ Gencer Sancar,[4]‡ Rigzin Dekhang,[5] Christopher M. Sullivan,[2,3] Shaojie Li,[5]§ Andrew G. Tag,[5,6] Cigdem Sancar,[4] Erin L. Bredeweg,[2] Henry D. Priest,[2,3] Ryan F. McCormick,[5] Terry L. Thomas,[5,6] James C. Carrington,[2,3] Jason E. Stajich,[7]¶ Deborah Bell-Pedersen,[5] Michael Brunner,[4] and Michael Freitag[1,2]*

Department of Biochemistry and Biophysics,[1] Center for Genome Research and Biocomputing (CGRB),[2] and Department of Botany and Plant Pathology,[3] Oregon State University, Corvallis, Oregon 97331-7305; Biochemistry Center (BZH), Universität Heidelberg, DE-69120 Heidelberg, Germany[4]; Center for Biological Clocks Research and Program for the Biology of Filamentous Fungi[5] and Laboratory for Functional Genomics,[6] Department of Biology, Texas A&M University, College Station, Texas 77843-3258; and Department of Plant and Microbial Biology, University of California, Berkeley, California 94720-3102[7]

Received 15 June 2010/Accepted 21 July 2010

Light signaling pathways and circadian clocks are inextricably linked and have profound effects on behavior in most organisms. Here, we used chromatin immunoprecipitation (ChIP) sequencing to uncover direct targets of the *Neurospora crassa* circadian regulator White Collar Complex (WCC). The WCC is a blue-light receptor and the key transcription factor of the circadian oscillator. It controls a transcriptional network that regulates ~20% of all genes, generating daily rhythms and responses to light. We found that in response to light, WCC binds to hundreds of genomic regions, including the promoters of previously identified clock- and light-regulated genes. We show that WCC directly controls the expression of 24 transcription factor genes, including the clock-controlled *adv-1* gene, which controls a circadian output pathway required for daily rhythms in development. Our findings provide links between the key circadian activator and effectors in downstream regulatory pathways.

Light perception leads to changes in gene expression that ultimately alter physiology. Even after extensive research, little is known about mechanisms that directly link photoreceptor activation to signaling pathways eliciting light responses. Typically, these regulatory modules are tightly coupled to endogenous circadian clocks (5). In *Neurospora crassa*, a eukaryotic model for light responses and the circadian clock, the blue-light receptor and PAS domain GATA-type transcription factor (TF) WHITE COLLAR-1 (WC-1) dimerizes with a second PAS domain GATA TF, WC-2, to form the White Collar Complex (WCC) (1, 31). In the FRQ/WCC oscillator, the WCC is the positive element that directly activates *frequency* (*frq*) transcription in the morning by binding the *frq* promoter at the proximal light response element (pLRE), required for light regulation of *frq*, and the distal LRE (dLRE) or "clock box" (C box), which is required for clock and light regulation (2, 18, 19, 24). This regulation results in rhythmic expression of *frq* mRNA and FRQ protein. FRQ protein dimerizes and binds to the FRQ-RNA helicase (FRH). The FRQ/FRH complex (FFC) functions as the negative element in the circadian negative-feedback loop (8, 23). Once FRQ protein is made, it becomes progressively phosphorylated, and when fully phosphorylated, it is degraded, allowing the cycle to restart the next morning (16, 32, 40).

The WCC is also required for all known blue-light responses, including resetting the circadian clock, carotenoid synthesis, asexual spore development, formation of female sexual structures, and ascospore release (47). In response to light, activated WCC functions as a TF and binds to LREs to regulate the expression of target genes. More than 100 light-responsive genes have been identified in neurospora, primarily through transcript microarray studies. Importantly, not all of these genes contain an obvious LRE (6, 28, 29, 31), which implies that a genetic regulatory cascade orchestrates expression of the light-responsive gene network. Furthermore, most light-regulated genes are also clock regulated. The identification of key molecules that function in light-regulated/circadian output pathways and control rhythms in target gene expression is one major goal of circadian biology. Despite significant efforts, these molecules have been difficult to identify. While the

* Corresponding author. Mailing address: Department of Biochemistry and Biophysics, ALS 2011, Oregon State University, Corvallis, OR 97331-7305. Phone: (541) 737-4845. Fax: (541) 737-0481. E-mail: freitagm@cgrb.oregonstate.edu.

‡ K.M.S. and G.S. contributed equally to the experimental data.

§ Present address: Institute of Microbiology, Chinese Academy of Sciences, No. 8 Bei Er Tiao, Zhong Guan Cun, Beijing 100080, People's Republic of China.

¶ Present address: Department of Plant Pathology and Microbiology, University of California, Riverside, CA 92521.

† Supplemental material for this article may be found at http://ec.asm.org.

▽ Published ahead of print on 30 July 2010.

predicted functions of the neurospora light-inducible genes suggest which processes are regulated by light, microarray studies have failed to directly link the action of the WCC to these genes. Identifying the genes directly regulated by the WCC is an essential step toward generating hierarchical or network models to describe the molecular and physiological responses to light and the circadian clock.

To identify the direct targets of the light-activated WCC, we used chromatin immunoprecipitation (ChIP), followed by high-throughput sequencing of bound DNA on an Illumina genome analyzer (ChIP sequencing [ChIP-seq]) (25, 34). We found that the WCC binds to hundreds of sites, predominantly upstream of genes. WCC targets fell into various functional categories, but genes encoding TFs were overrepresented. We tested all TF genes and numerous other genes with WCC binding sites in their promoters for light induction and found that, as expected, most, but not all, responded to light. We show that one of the genes encoding a TF, *adv-1*, is a clock-controlled gene that is necessary for circadian rhythms in development. Our data suggest a "flat" hierarchical network in which ~20% of all annotated neurospora TFs are regulated during the early light response by the WCC, the key TF factor of the circadian clock.

MATERIALS AND METHODS

ChIP. Neurospora cultures (FGSC2489) were grown at 25°C in minimal medium (1× Vogel's salts, 2% glucose) in the light for 2 days, transferred to the dark for 12 h, light induced (160 μE) for 8 min, and cross-linked in constant light with 1% paraformaldehyde for 15 min. We isolated nuclei (33) and performed ChIP as described previously (26) on 6 mg of nuclear fraction with ~2 μg of WC-2 antibody (polyclonal antibody raised to a WC-2 protein fragment expressed in *Escherichia coli*) (44). Histone ChIPs were performed on germinated conidia with dimethylated H3 lysine 4 (H3K4me2) (Upstate; 07-030) antibodies as described previously (45). All ChIP experiments were validated by ChIP-PCR before (not shown) and after (see Fig. S1 in the supplemental material) ChIP-seq library construction to verify previously described results.

ChIP-seq library construction and high-throughput sequencing. DNA was end repaired and ligated to adapters (38). The 200- to 500-bp fragments were gel purified and amplified with 20 cycles of PCR using Phusion polymerase (Finnzymes Oy). the PCR products were gel purified and sequenced on an Illumina 1G sequencer in the Oregon State University Center for Genome Research and Biocomputing (OSU CGRB) core laboratories.

Region-specific ChIP PCR. Duplex PCR with [α^{32}P]dCTP was performed to determine enrichment in the ChIP samples relative to input DNA with region-specific oligonucleotide primers (see Table S1 in the supplemental material) (38, 45). An *hH4-1* segment was used as a control. Phosphorimager screens were exposed to dried gels and analyzed with a GE Storm 820 imager.

Data analysis and visualization. Illumina reads were mapped to the neurospora assembly 7 (http://www.broadinstitute.org/annotation/genome/neurospora/MultiHome.html) genome with SOAP (30), with the following parameters: -s 10 -c 41 -r 1 -g 1 -w 10,000 (seed size, 10; per-read incremental trimming of 3′ basepairs to remove low-quality or nonaligning bases, choosing randomly for equal best locations for a read; only a single random location placement of reads; a maximum gap size of 1; and a maximum of 10,000 mapped locations).

A second analysis used CASHX (17) to map perfect 36-nucleotide (nt) matches to the genome, calculate reads per window across the whole genome, and assign significance values to areas defined as peaks. Only reads matching the genome 1 or 2 times were included in the analysis. We calculated the average read number per 500-bp sliding window, the standard deviation, and a *z* score for each window. ChIP-seq of the relatively compact neurospora genome (~42 Mb) typically results in some background coverage (e.g., ChIP-seq data for H3K4me2 [see Fig. S2 in the supplemental material]), which yields fewer false-negative calls than in larger, more complex genomes. Thus, complete absence of reads with respect to a given region of the reference genome strongly suggests that the region has been deleted from the resequenced strain (J. E. Stajich, K. M. Smith, and M. Freitag, unpublished data).

The updated *N. crassa* annotation tracks (Fig. 1 and 2; see Fig. S2 in the supplemental material) include extended transcript predictions of untranslated regions (UTRs) by using 279,323 *Neurospora tetrasperma* and 453,559 *Neurospora discreta* expressed sequence tags (ESTs) generated with 454 technology for the respective genome projects (Joint Genome Institute and J. E. Stajich, D. J. Jacobson, D. O. Natvig, N. L. Glass, and J. W. Taylor, unpublished data), as well as 84,309 *N. crassa* ESTs and full-length cDNAs available from GenBank and the NCBI Trace archive. This updated *N. crassa* annotation was produced with the PASA annotation system (3) with assembly 7 and annotation from GenBank (AABX02000000; 10 September 2007). Visualization of annotations and mapped read densities was done with Gbrowse (42) with wiggle track format.

Sequence analysis to find WCC consensus binding sites. Known WCC-bound promoters of light-induced genes (*frq*, *vvd*, and *al-3*), plus an additional 26 WCC binding sites found in this study at genes confirmed to be light induced, were extracted from assembly 7. One-kilobase regions centered on each peak were searched *de novo* for consensus *cis* elements with Weeder version 1.3.1 (parameters, NC large A M S T15) (37) and SCOPE (4; http://genie.dartmouth.edu/scope/). The consensus motif logos in Fig. 1B were generated using WebLogo version 2 (http://weblogo.berkeley.edu/logo.cgi).

Reverse transcription (RT)-PCR to assay light induction. We used standard laboratory *ras^{bd}* (22) clock strains and Δ*wc2 ras^{bd}* (10) mutant strains for light induction experiments. The *ras^{bd}* mutation enhances the developmental rhythm in cultures (39) but does not affect light responses or molecular rhythms in gene expression. The strains were grown for 24 h in 200 ml of Vogel's minimal medium (14) with 2% glucose and 0.5% arginine in constant dark and subjected to 15 min of full-spectrum light (160 μE). We chose a 15-min time point to better compare our results with published results and to allow time for RNA accumulation. Total RNA was prepared using the peqGOLD TriFAST kit (peqLab, Erlangen, Germany). We generated cDNA by reverse transcription with primers supplied with the QuantiTect Reverse Transcription Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Expression was analyzed by quantitative real-time PCR (22) with the primers listed in Table S2 in the supplemental material. We amplified cDNA for *act*, *frq*, and *frq* antisense RNA with specific primers (see Table S2 in the supplemental material).

Circadian rhythm assays. We crossed Δ*adv-1* to *ras^{bd}* to generate *ras^{bd}* Δ*adv-1* strains. The *ras^{bd}* Δ*adv-1* strain was assayed on standard race tube medium containing 1× Vogel's salts, 0.1% glucose, 0.5% arginine, and 1.5% agar. Race tube assays were performed in controlled environmental chambers in constant darkness at 25°C. For RNA and protein analyses, mycelial mats were grown in shaking liquid culture (100 rpm) in 25 ml of Vogel's minimal medium (1× Vogel's salts, 2% glucose, 0.5% arginine) and processed as described previously (21). Transfer of cultures was performed so that the cultures were approximately the same age (within 8 h) at the time of harvest but the circadian times of the cultures varied (12). The ADV-1:V5-tagged strain was a generous gift from Jay Dunlap's laboratory and the Neurospora Program Project Grant and contained a C-terminal V5 tag. The tagged gene was recombined into the endogenous *adv-1* locus. Antibody for V5 was obtained from Invitrogen (Carlsbad, CA).

RESULTS AND DISCUSSION

WCC ChIP. To identify direct targets of the WCC, we performed ChIP-seq with anti-WC-2 antibody on cultures subjected to an 8-min light pulse. We verified that the ChIP-seq library was enriched for known targets of the WCC (see Fig. S1 in the supplemental material). Of 4,866,015 32-nt-long ELAND-processed (25) sequences, 92% were mapped to assembly 7 of the neurospora genome by SOAP (30). Sequence reads are available through the NCBI sequence read archive (SRA010801.1).

The WCC binding sites were compared to ChIP-seq data for H3K4me2 (K. M. Smith, C. M. Sullivan, K. R. Pomraning, and M. Freitag, unpublished data) because the presence of this epigenetic modification is correlated with transcriptional activity. To establish a cutoff value for statistically significant WCC binding sites, we used the CASHX mapping algorithm (17) and sliding-window read counts to calculate the mean and standard deviation of reads per 500 bp for the entire genome. A count of 83 reads per 500-bp window was significantly above the mean level of background signal ($P < 0.001$ and $z = 3.09$).

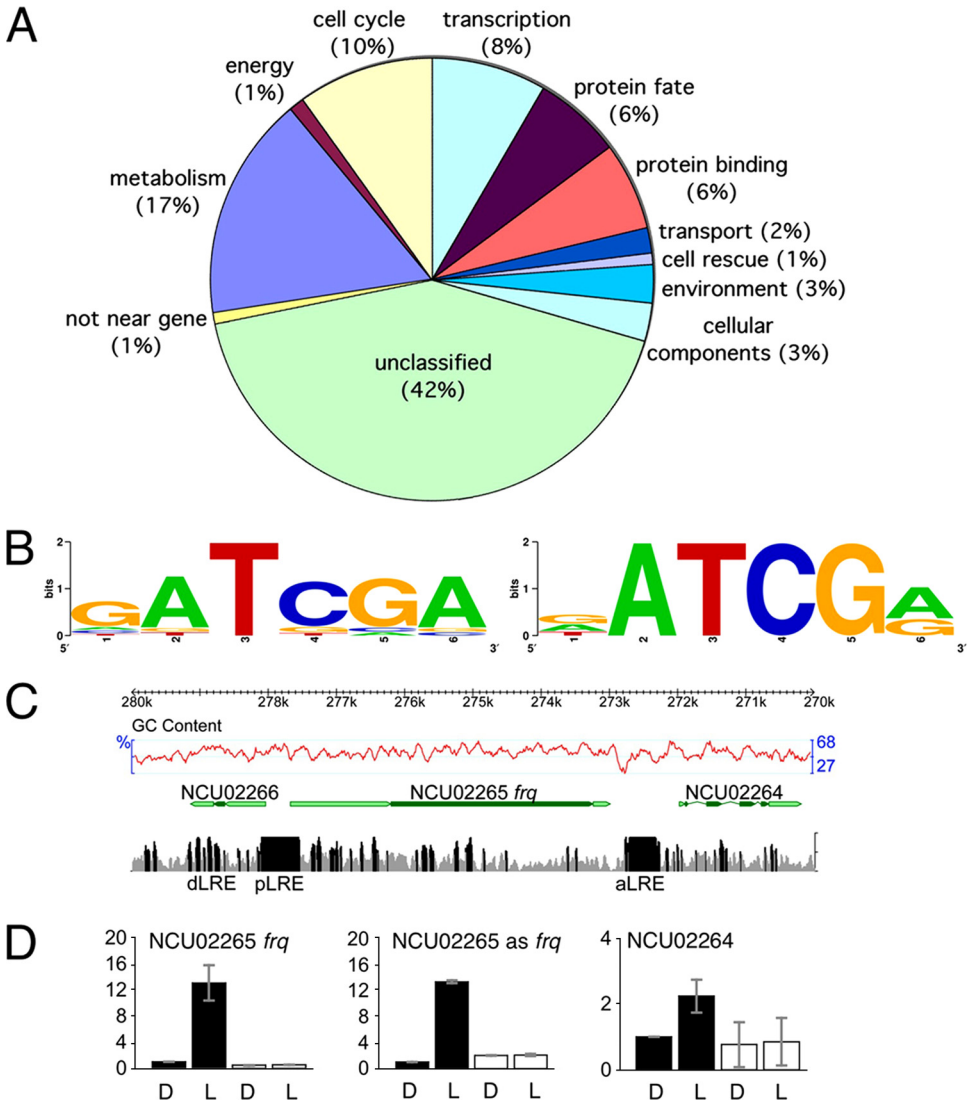

FIG. 1. Genomewide analysis of WCC binding. (A) Genes with WCC binding sites fall into diverse functional categories (indicated by the percentage of genes with binding sites within a class). Only the most significant 109 WCC binding sites (z score > 5; $P < 2.9e{-}7$) were included here, but a comparison between functional categories of the best targets and all targets is shown in Fig. S3 in the supplemental material. (B) Analyses of WCC binding sites found by ChIP-seq with Weeder (37) (left) and SCOPE (4) (right) revealed a consensus binding site for the WCC. The relative height of each nucleotide (shown in 5′-to-3′ direction on the x axis) indicates the degree of sequence conservation, with a maximal score of 2 (as indicated on the y axis). (C) ChIP-seq verified known light response elements (dLRE and pLRE) and revealed a novel binding site (aLRE) for the WCC in the *frq* region. SOAP-mapped sequence reads from the WC-2 ChIP library were plotted as a histogram of sequence coverage per base along a 10-kb fragment containing the *frq* locus (the negative strand of contig 7.10; nt 270,000 to 280,000) in a customized genome browser to integrate ChIP-seq results with gene annotation data (http://gb.fungalgenomes.org/gb/gbrowse/neurospora_crassa_OR74A_7/). The y axis has a maximum value of 20 reads per base, and regions with >10 reads are shown in black, with others in gray. Gene annotations are based on known or predicted cDNAs (UTRs are shown in light green, coding regions in dark green, and introns as thin lines). (D) WCC binding results in light induction of *frq*, antisense *frq*, and NCU02264 transcripts and is dependent on WC-2. Quantitative RT-PCR was performed on RNA isolated from control (black bars) and Δ*wc-2* (white bars) strains grown in the dark (D) or following a 15-min light induction (L). The values shown are the averages of two replicates, normalized to the level of actin expression (y axis). The error bars indicate standard deviations.

In previous studies, only the *frq*, *vvd*, *al-3*, *fl*, and *sub-1* promoters had been identified as direct targets for the WCC (6, 18, 24, 36). By ChIP-seq, we identified >400 significant regions of WCC enrichment, with >200 falling in known or predicted promoters of at least one gene. Here, we focus on two groups: (i) the most significant peaks located in promoters and (ii) peaks in promoters of TF genes (Table 1). We also compiled a complete list of regions with significant reads, including genes involved in the circadian clock, chromatin function, kinases, phosphatases, cell cycle, DNA replication, DNA repair, and metabolism (see Table S1 in the supplemental material). Target genes were summarized according to their functional categories. While almost half of all target genes encode unclassified proteins, we noticed an enrichment for

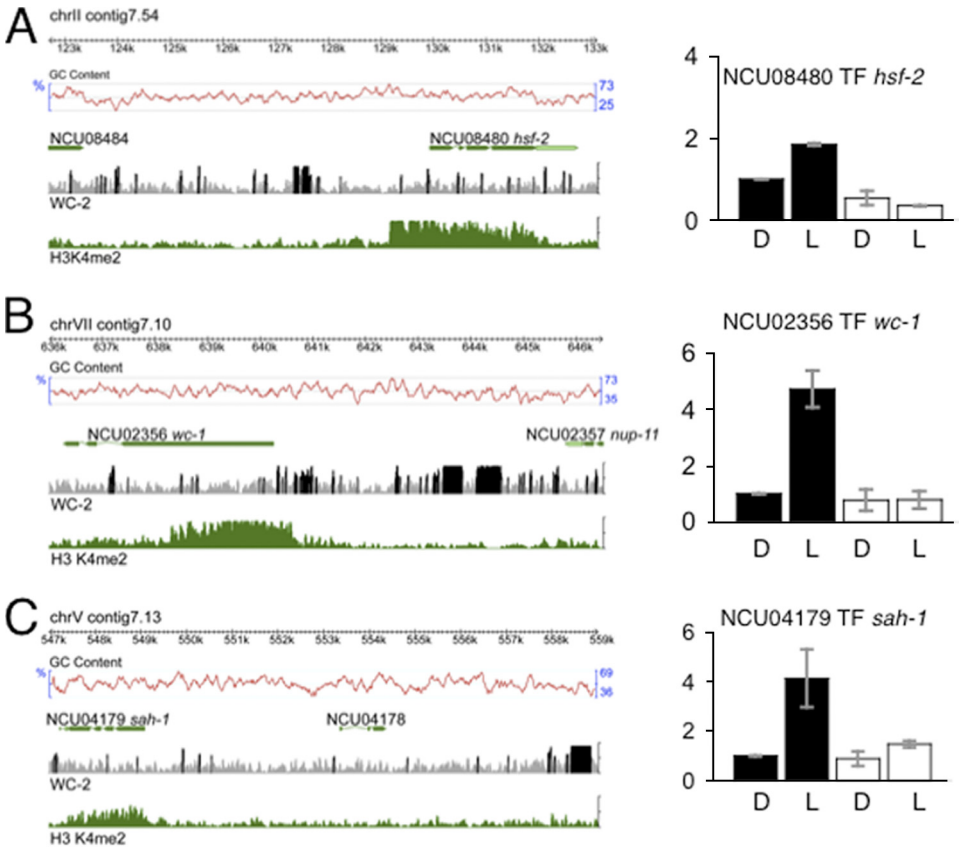

FIG. 2. ChIP-seq reveals new targets of the WCC in the promoters of TF genes, and these genes are light inducible. (A) The *hsf-2* gene has a single WCC binding site and is light induced in a WC-2-dependent manner. (B) The *wc-1* gene, an early light-induced gene, has three upstream WCC binding sites, one close to and two 3 to 4 kb upstream of the transcriptional start site. (C) The *sah-1* gene has a single WCC binding site ~9 kb upstream of its predicted transcriptional start site. Tracks, colors, and conditions for qRT-PCR are the same as for Fig. 1C and D. The error bars indicate standard deviations.

genes with functions in the cell cycle, transcription, protein binding, response to the environment, and cellular components when we compared the best 109 targets (z score > 5; $P < 2.9e{-}7$) to all 584 targets (z score > 3.09; $P < 0.001$) (Fig. 1A; see Fig. S3 in the supplemental material).

The Weeder (37) and SCOPE (4) algorithms were used to derive a consensus binding site for WCC based on 1-kb regions centered on the WCC ChIP-seq peak at 29 genes that were confirmed to be light induced (Fig. 1B). Both algorithms identified a common consensus binding site, GATCGA (with variability in the first and last bases), which extends the most recently published consensus (GATC) that was derived from studies with expression arrays (6). Every 1-kb fragment used to generate the consensus contained at least one, but often two or more, copies of this motif interrupted by a variable number of nucleotides.

For a known WCC target gene, *frq*, we observed peaks in two regions of the *frq* promoter that corresponded exactly to the previously identified pLRE and dLRE sites (18, 24) (Fig. 1C). Enrichment at the *frq* dLRE (or C box), which deviates from the consensus WCC binding motif, was less pronounced than at the pLRE. WCC binding at the pLRE correlated with increased *frq* transcription in response to light (Fig. 1D, black bars). This induction was absent in a Δ*wc-2* strain (Fig. 1D, white bars). We found an additional peak downstream of the *frq* coding sequence, which we named the "antisense LRE" (aLRE) because we predicted it would control light-induced expression of the antisense *frq* transcript (13).

The aLRE also falls in the promoter of NCU02264, encoding a predicted protein with a prefoldin chaperone domain. Binding of WCC to the aLRE is associated with induction of both the antisense *frq* transcript and the divergently transcribed NCU02264 (Fig. 1D). Approximately 15% of all WCC binding sites in promoters occur between two divergently transcribed genes, and most often, both genes are regulated by one shared binding site (Table 1). We measured the light induction of transcription by quantitative PCR (qPCR) and compared our results to previous results from microarray experiments (6) performed under similar conditions (Table 1). This list included known clock- and light-regulated genes and several novel genes. A replicate WC-2 ChIP was used to validate the ChIP-seq results by using duplex ChIP-PCRs for numerous regions. In all cases, enrichment shown by ChIP-PCR validated the WC-2 ChIP-seq results.

Of the other known WCC binding sites, both *vvd* and *sub-1*, listed in Table 1, have highly significant WCC binding sites in their promoters. Both were also light induced (Table 1; see Fig. S2A in the supplemental material for *sub-1*). The

TABLE 1. Regions of WCC enrichment after 8 min of light induction

Contig[a]	Start nt	No. of reads	*z* score[b]	Gene(s)[c]	Array induction[d]	qPCR[e]
Largest peaks[f]						
7.10	277601	738	45.66	NCU02265 *frq*	2.35 (10)	13
7.12	422401	696	42.93	NCU03967 *vvd*	4.4 (15)	130
7.2	298501	657	40.40	NCU00582 *cry*	4.8 (15)	269
7.57	14601	415	24.68	NCU08699 *bli-4*	4.9 (15)	ND
7.9	872451	387	22.87	NCU03071 *os-4*/NCU03072	1.6 (10)/2.3 (15)	4.2/7.1
7.10	272251	378	22.28	NCU02264/NCU02265 *frqAS*	2.3 (15)/*frqAS*[g]	2.6/13.4
7.48	193601	321	18.58	NCU10063	6.5 (10)	ND
7.12	623601	298	17.09	NCU04021	NF	2.7
7.22	197301	262	14.75	NCU05594	8 (15)	ND
7.7	782201	253	14.17	NCU02800/NCU02801	NF/3.8 (15)	2.4/ND
7.45	138451	235	13.00	NCU07541	2.8 (30)	ND
7.2	196651	230	12.67	NCU00552 *al-1*	14.3 (15)	ND
7.12	426151	226	12.41	NCU03968	1.8 (30)	ND
7.21	6451	210	11.37	NCU11300/NCU06017	NF	4.7/1.7
7.2	305951	200	10.72	NCU00584/NCU00585 *al-2*	3.7 (15)/5.1 (45)	ND
Peaks close to transcription factor genes[h]						
7.13	558401	340	19.81	NCU04179 *sah-1* TF	3.4 (30)	4.1
7.10	644051	245	13.65	NCU02356 *wc-1* TF	3.1 (10)	4.7
7.13	125751	236	13.06	NCU04295 TF	1.4 (10)	1.2
7.81	66601	203	10.92	NCU09615 TF	NF	7.1
7.59	47101	182	9.56	NCU09068 *nit-2* TF	NF	NI
7.4	336501	171	8.84	NCU05964 TF	NF	2.2
7.88	31301	170	8.78	NCU09829/NCU09830 TF	NF/11.6 (15)	7.2
7.15	93601	162	8.27	NCU01242/NCU01243 TF	NF/1.8 (10)	2.4/1.3
7.66	66151	158	8.00	NCU01871 TF/NCU01873	NF	NI/1.4[i]
7.54	127201	150	7.48	NCU08480 *hsf-2* TF	NF	1.8
7.6	191451	124	5.79	NCU02713 *csp-1* TF	4.4 (10)	16.4
7.60	44551	122	5.66	NCU08807 *cre-1* TF/NCU08806 *rhp-55*	NF	1.3/2.1
7.38	129051	115	5.20	NCU07392 *adv-1* TF	NF	3.8
7.1	389051	115	5.20	NCU00097 *bek-1*TF	NF	5.5
7.47	109401	107	4.69	NCU07705 TF	NF	1.2
7.15	391951	96	3.97	NCU01154 *sub-1* TF	1.8 (15)	8.1
7.28	101201	92	3.71	NCU06534 TF/NCU06536	NF	NI/NI
7.16	202151	91	3.61	NCU04731 *sah-2* TF	NF	1.2
7.47	25201	90	3.58	NCU07728 *sre* TF	NF	3.1
7.4	173001	89	3.52	NCU07846 TF	NF	2
7.1	1039001	87	3.39	NCU00275 TF/NCU00276 *mip-1*	NF/1.7 (30)	2.5
7.3	949151	86	3.32	NCU02094 *vad-2* TF/NCU02095	NF	2.3/NI
7.4	200401	86	3.32	NCU05994 TF	NF	NI
7.9	175701	85	3.26	NCU03273 TF/NCU03271	NF	NI/NI
7.9	498501	85	3.26	NCU03184 TF	NF	1.7
7.51	99701	84	3.19	NCU08000 TF	NF	NI
7.52	161651	83	3.13	NCU08159 TF	NF	1.9
7.21	260701	83	3.13	NCU06095 TF	1.8 (10)	6.2

[a] Assembly 7 of the neurospora genome (http://www.broadinstitute.org/annotation/genome/neurospora/MultiHome.html).
[b] *z* score calculated as described in Materials and Methods.
[c] If binding sites were located between two divergently transcribed genes, both locus names are listed.
[d] Level of induction (*x*-fold) calculated from raw data (6) at the time in minutes shown in parentheses. If two genes were near WC-2 binding sites, the fold induction levels are separated by a slash. NF, not found in array experiments.
[e] Expression level changes determined in this study by qPCR. If genes were known to be light induced from earlier studies and reference 6, qPCR was usually not done. NI, not induced; ND, not determined because there was previous evidence for light induction.
[f] Only regions with 200 or more reads/500-bp sliding window are listed (all peaks are shown in Table S1 in the supplemental material).
[g] The antisense *frq* transcript was not tested in the array experiments.
[h] Transcription factors are named or identified by their locus numbers and indicated by "TF." The WCC is enriched in promoter regions of 28 TFs, 24 of which are induced by light and/or their regulation is altered in a *wc-2* deletion mutant (Fig. 2; see Fig. S2 in the supplemental material).
[i] Large standard deviation.

vvd and *al-3* promoter fragments were used in ChIP-PCRs to check the quality of the ChIP libraries prior to sequencing (see Fig. S1 in the supplemental material), and both were enriched in the library. Even though *al-3* promoter sequence was present in the library, the number of reads sequenced from this region fell below our cutoff for significance. The *fl* promoter was recently shown to be a target of WCC (36), but we did not find enrichment at this promoter by ChIP-seq. The *al-3* PCR product band was much fainter than the *frq* and *vvd* bands (see Fig. S1 in the supplemental material).

These disparities may be explained by weak and/or transient binding in these regions.

Identification of novel WCC binding sites in promoters of transcription factors. We chose to focus on WCC binding near TF genes as the first step in unraveling transcriptional networks that respond to light and are controlled by the circadian clock. We initially predicted that light signaling by the WCC would involve the activation of a few key downstream TFs that would, in turn, control a network of target genes. Instead, we identified 28 known or putative TF genes with significant ($P < 0.001$) WCC binding in their promoters (Table 1; see Fig. S2 in the supplemental material). To discover if these TF genes were true targets for the WCC, we analyzed transcript levels by quantitative RT-PCR in response to a 15-min light exposure in *wc-2*$^+$ (*bd*) and *wc-2* deletion (*bd* Δ*wc-2*) strains because we expected that direct targets of WCC would be induced by light and that light induction requires WC-2, an obligatory subunit of the WCC (9, 18, 24, 44).

We found that of the 28 TF genes with significant WCC binding sites upstream of their transcriptional start sites, 21 were induced by light in a WC-2-dependent fashion (Table 1; see Fig. S2A in the supplemental material). Three genes (*nit-2*, NCU06534, and NCU08000) were not regulated by light under our conditions, but expression at wild-type levels was dependent on WC-2 (see Fig. S2B in the supplemental material). Four TF genes (NCU05994, NCU03273, NCU01871, and *cre-1*) were excluded as WCC targets, at least under our conditions (see Fig. S2C in the supplemental material). No target has been identified for the peaks in the promoters of NCU05994 and NCU03273, but for NCU01871 and *cre-1*, the neighboring genes (NCU01873 and *rhp55*, respectively) were light inducible in a WC-2-dependent manner, suggesting that they are the actual targets for the WCC.

Of the 21 TF genes we identified here as WCC targets, only 4 (*wc-1*, *sub-1*, *csp-1*, and *sah-1*) had been previously identified as light regulated (6), and *sub-1*, *csp-1*, and *sah-1* were not known to be involved in circadian output pathways. Conversely, two TF genes identified as light induced in the microarray study lacked high-confidence WCC binding sites in their promoters in our study (NCU06407 *vad-3* and NCU03643, the gene encoding cutinase TF-1β). However, both genes had a single perfect match to the WCC consensus site (GATCGA) and several sites with a single mismatch (GATCCA) within 1 kb of their predicted initiation codons. Thus, while the two studies were largely congruent, ChIP-seq identified more direct first-tier targets and was more suitable for detecting regulation of TF genes, which are often expressed at low levels, or whose expression levels are changed within a narrow range.

Examples of genes regulated by single or multiple WCC binding sites are shown in Fig. 2, and data for the most significant peaks are summarized in Table 1. We found a single WCC binding site upstream of the gene for heat shock factor 2, *hsf-2*, and transcription of *hsf-2* was light inducible and dependent on WC-2 (Fig. 2A). We found three WCC binding sites upstream of the *wc-1* gene (Fig. 2B), which encodes the limiting subunit of the WCC. Transcription of *wc-1* was induced by light and dependent on WC-2, in agreement with previous reports (1). Our data show that *wc-1* is directly regulated by the WCC in a positive-feedback loop, as had been suggested by previous work (27). Transcriptional start sites of the *wc-1* gene have been mapped at −924 and −1,222 bp (27). One WCC binding site is close to the transcriptional start site at −924. The other WCC binding sites at −3 and −4 kb are, however, located far upstream of the mapped transcriptional start sites.

WCC binds to sites far upstream of predicted transcriptional start sites in additional TF genes, e.g., *sah-1*, *sub-1*, *nit-2*, *bek-1*, and NCU05964 (see Fig. S2A in the supplemental material). In particular, a single strong WCC binding site is located about 10 kb upstream of the open reading frame (ORF) of *sah-1* (Fig. 2C), which is light inducible in a WC-2-dependent manner. Thus, WCC binding appears to be capable of exerting long-range effects in the relatively compact neurospora genome, where promoters are typically short and sometimes overlapping (20). The precise mechanism of the enhancer-like effects of WCC merits further investigation.

Developmental stages that have been associated with light regulation, i.e., asexual sporulation (conidiation) and sexual development, are represented by nine TFs whose mutation results in developmental phenotypes (*hsf-2*, *adv-1*, *sub-1*, *sah-1*, *sah-2*, *bek-1*, *vad-2*, *csp-1*, and *ghh*) (11, 46). All nine genes are induced by light, and this response is abolished in a *wc-2* deletion strain (Table 1; see Fig. S2A in the supplemental material).

WCC binding sites were associated with six TF genes involved in metabolism or stress response (*nit-2*, NCU05994, *sre* [regulator of iron uptake] [48], *cre-1*, NCU08000, and NCU05964 [homolog of *Aspergillus vosA*] [35]). Three (*cre-1*, *vosA*, and *sre*) were light induced, and this response was abolished in the Δ*wc-2* mutant, but the other three genes showed no light induction under our conditions (see Fig. S2C and D in the supplemental material). The expression levels of these genes in the Δ*wc-2* strain were decreased. These results suggest that WCC directly impinges on metabolic pathways.

The downstream targets of 11 TFs identified here remain completely unknown. These putative TF genes (NCU00275, NCU01243, NCU03184, NCU03273, NCU04295, NCU06534, NCU07705, NCU07846, NCU08159, NCU09615, and NCU09829) have no previously described function, but they encode motifs that match well-studied DNA binding domains (11). NCU00275, NCU01243, NCU03184, NCU04295, NCU07705, NCU07846, NCU08159, NCU09615, and NCU09829 were light induced, a response abolished in the Δ*wc-2* mutant (see Fig. S2A in the supplemental material).

The WCC activates an output pathway required for rhythmic spore development. To investigate the link between light and circadian clock pathways regulated by the WCC, we assayed available knockout mutants of the WCC target TFs for changes in circadian rhythms in development (data not shown). The most striking phenotype was observed in the Δ*adv-1* mutant, so we investigated this mutant further for clock defects. Loss of *adv-1* had only minor effects on the development of spores, but development was no longer under the control of the clock (Fig. 3A).

To determine if arrhythmicity of the Δ*adv-1* mutant was due to an effect on the FRQ/WCC oscillator, we assayed rhythms in the accumulation of the FRQ protein in control and Δ*adv-1* strains over the course of 2 days in cultures grown in the dark (Fig. 3B and data not shown). No differences were observed in

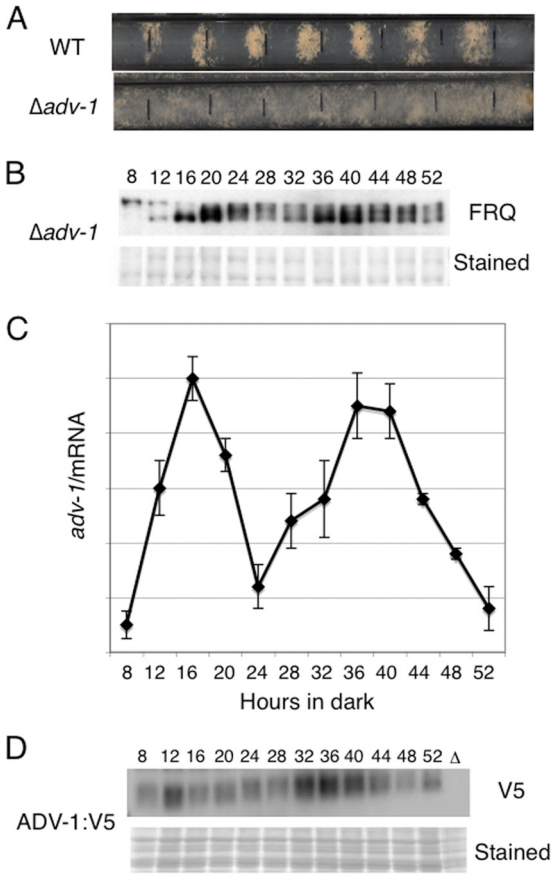

FIG. 3. *adv-1* is a clock-controlled gene required for circadian rhythms in development. (A) Δ*adv-1* strains are arrhythmic. Shown is a race tube assay for rhythmic development in wild-type clock (WT) and Δ*adv-1* strains. Each strain was inoculated and grown in constant light for 24 h before transfer to constant darkness at 25°C, after which the growth front was marked every 24 h (black lines). (B) FRQ protein remains rhythmic in strains lacking ADV-1. Shown is a Western blot of FRQ protein in the indicated strain. The numbers of hours in constant darkness are shown above the blots. (C) *adv-1* mRNA accumulates with a circadian rhythm. Northern blots of *adv-1* mRNA levels were performed three times, and the average and SD for each time point were calculated and plotted. For each experiment, the *adv-1* mRNA from each time point was normalized to rRNA, and the lowest point was set to 1. The number of hours in constant darkness are shown on the *x* axis. (D) ADV-1 protein accumulates with a circadian rhythm. Shown is a Western blot of ADV-1–V5 probed with antibody to V5 from cultures harvested at the indicated times in the dark. Protein from the *adv-1* deletion strain (Δ) harvested after 16 h in the dark was used as a negative control to demonstrate the specificity of the antibody. Amido black-stained membranes are shown as loading controls in the Western blots in panels B and D (Stained). Each experiment was repeated at least two times with similar results.

the FRQ rhythms in the mutant strains, suggesting that ADV-1 functions downstream of the oscillator within an output pathway from the clock. In support of this idea and consistent with previous microarray results (15), we observed that the accumulation of *adv-1* mRNA is rhythmic, peaking in the subjective morning (Fig. 3C). This is similar to the time of peak WCC activity in the circadian clock (41). Furthermore, we found that ADV-1 protein tagged with a V5 epitope accumulated with a circadian rhythm (Fig. 3D). Future efforts will determine if the WCC consensus binding site present in the *adv-1* promoter (at −735 from the predicted start of transcription) is essential for circadian rhythmicity of *adv-1* mRNA accumulation and for the developmental rhythm.

Conclusions. In summary, we have identified a large number of direct targets of the WCC, the key regulator of the circadian clock. These targets include 24 TFs that have the potential to control downstream target genes on a second hierarchical level. Genomewide identification of target genes for these second-tier TFs, together with genomewide expression studies by RNA-seq, will allow us to build a detailed network of the early and late light responses, as well as circadian clock output pathways. Substantial effort over the past several years has been spent on identifying key components of circadian output pathways in neurospora and other organisms, with only limited success (7, 43, 47). By applying WCC ChIP-seq to this effort, we have now uncovered excellent candidates for output pathway components, such as ADV-1, that lie directly downstream of the oscillator and that regulate distinct overt rhythms. In addition to gene products that may mediate posttranslational clock and light effects, we also found many WCC binding sites in promoters of metabolic genes that act in various anabolic and catabolic pathways.

Most, but not all, of the genes that are direct targets of the WCC were found in our experiments to be light induced. This may not be surprising for genes that are subject to complex regulation. For example, a gene may be repressed under most growth conditions and be activated by light and the bound WCC only when these conditions are not met and repression is removed. Similarly, while we expect most, if not all, genes directly regulated by the WCC to be rhythmically transcribed, other regulatory elements may take precedence under certain growth conditions.

While many of the downstream genes regulated by the WCC are not well studied or are uncharacterized, most of them have homologs in plants and mammals. Thus, our work provides key information to refine the growing network of light- and clock-regulated genes in a genetically and biochemically tractable model organism.

ACKNOWLEDGMENTS

We thank Mark Dasenko, Steve Drake, and Scott Givan at the OSU CGRB core facility for Illumina sequencing; Zachary Lewis for helpful discussions and comments on the manuscript; and Noah Fahlgren for sharing code. We thank Jonathan Arnold for supplying files with expression microarray data; Matt Sachs for sharing data from RNA sequencing; and Dave Jacobson, Don Natvig, Louise Glass, and John Taylor for use of data from the *N. discreta* and *N. tetrasperma* genome projects. The V5-tagged ADV-1 strain was generated and generously provided by Jay Dunlap's laboratory in the context of the Neurospora Functional Genomics and Systems Biology Program Project (NIH 2P01 GM068087).

J.E.S. was supported by a fellowship from the Miller Institute for Basic Research in Science. This work was supported by grants from the American Cancer Society (RSG-08-030-01-CCG to M.F.), NIH (GM58529 and NS39546 to D.B.-P.; AI43288 to J.C.C.; 2P01 GM068087 to D.B.-P and M.F.), NSF (MCB0618433 to J.C.C.), and DFG (BR 1375-1 and SFB 638 to M.B.).

We have no conflicting interests.

REFERENCES

1. **Ballario, P., P. Vittorioso, A. Magrelli, C. Talora, A. Cabibbo, and G. Macino.** 1996. WHITE COLLAR-1, a central regulator of blue light responses in *Neurospora*, is a zinc finger protein. EMBO J. **15:**1650–1657.
2. **Belden, W. J., J. J. Loros, and J. C. Dunlap.** 2007. Execution of the circadian

negative feedback loop in *Neurospora* requires the ATP-dependent chromatin-remodeling enzyme CLOCKSWITCH. Mol. Cell **25:**587–600.
3. **Campbell, M. A., B. J. Haas, J. P. Hamilton, S. M. Mount, and C. R. Buell.** 2006. Comprehensive analysis of alternative splicing in rice and comparative analyses with *Arabidopsis*. BMC Genomics **7:**327.
4. **Carlson, J. M., A. Chakravarty, C. E. DeZiel, and R. H. Gross.** 2007. SCOPE: a web server for practical *de novo* motif discovery. Nucleic Acids Res. **35**:W259–W264.
5. **Cermakian, N., and P. Sassone-Corsi.** 2002. Environmental stimulus perception and control of circadian clocks. Curr. Opin. Neurobiol. **12:**359–365.
6. **Chen, C. H., C. S. Ringelberg, R. H. Gross, J. C. Dunlap, and J. J. Loros.** 2009. Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in *Neurospora*. EMBO J. **28:**1029–1042.
7. **Cheng, M. Y., C. M. Bullock, C. Li, A. G. Lee, J. C. Bermak, J. Belluzzi, D. R. Weaver, F. M. Leslie, and Q. Y. Zhou.** 2002. Prokineticin 2 transmits the behavioural circadian rhythm of the suprachiasmatic nucleus. Nature **417:**405–410.
8. **Cheng, P., Q. He, L. Wang, and Y. Liu.** 2005. Regulation of the *Neurospora* circadian clock by an RNA helicase. Genes Dev. **19:**234–241.
9. **Cheng, P., Y. Yang, K. H. Gardner, and Y. Liu.** 2002. PAS domain-mediated WC-1/WC-2 interaction is essential for maintaining the steady-state level of WC-1 and the function of both proteins in circadian clock and light responses of *Neurospora*. Mol. Cell. Biol. **22:**517–524.
10. **Collett, M. A., J. C. Dunlap, and J. J. Loros.** 2001. Circadian clock-specific roles for the light response protein WHITE COLLAR-2. Mol. Cell. Biol. **21:**2619–2628.
11. **Colot, H. V., G. Park, G. E. Turner, C. Ringelberg, C. M. Crew, L. Litvinkova, R. L. Weiss, K. A. Borkovich, and J. C. Dunlap.** 2006. A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. Proc. Natl. Acad. Sci. U. S. A. **103:**10352–10357.
12. **Correa, A., Z. A. Lewis, A. V. Greene, I. J. March, R. H. Gomer, and D. Bell-Pedersen.** 2003. Multiple oscillators regulate circadian gene expression in *Neurospora*. Proc. Natl. Acad. Sci. U. S. A. **100:**13597–13602.
13. **Crosthwaite, S. K., J. J. Loros, and J. C. Dunlap.** 1995. Light-induced resetting of a circadian clock is mediated by a rapid increase in *frequency* transcript. Cell **81:**1003–1012.
14. **Davis, R. H.** 2000. Neurospora: contributions of a model organism. Oxford University Press, Oxford, United Kingdom.
15. **Dong, W., X. Tang, Y. Yu, R. Nilsen, R. Kim, J. Griffith, J. Arnold, and H. B. Schüttler.** 2008. Systems biology of the clock in *Neurospora crassa*. PLoS One **3:**e3105.
16. **Dunlap, J. C., J. J. Loros, H. V. Colot, A. Mehra, W. J. Belden, M. Shi, C. I. Hong, L. F. Larrondo, C. L. Baker, C. H. Chen, C. Schwerdtfeger, P. D. Collopy, J. J. Gamsby, and R. Lambreghts.** 2007. A circadian clock in *Neurospora*: how genes and proteins cooperate to produce a sustained, entrainable, and compensated biological oscillator with a period of about a day. Cold Spring Harbor Symp. Quant. Biol. **72:**57–68.
17. **Fahlgren, N., C. M. Sullivan, K. D. Kasschau, E. J. Chapman, J. S. Cumbie, T. A. Montgomery, S. D. Gilbert, M. Dasenko, S. A. Givan, and J. C. Carrington.** 2009. Computational and analytical framework for small RNA profiling by high-throughput sequencing. RNA **15:**992–1002.
18. **Froehlich, A. C., Y. Liu, J. J. Loros, and J. C. Dunlap.** 2002. WHITE COLLAR-1, a circadian blue light photoreceptor, binding to the *frequency* promoter. Science **297:**815–819.
19. **Froehlich, A. C., J. J. Loros, and J. C. Dunlap.** 2003. Rhythmic binding of a WHITE COLLAR-containing complex to the *frequency* promoter is inhibited by FREQUENCY. Proc. Natl. Acad. Sci. U. S. A. **100:**5914–5919.
20. **Galagan, J. E., S. E. Calvo, K. A. Borkovich, E. U. Selker, N. D. Read, D. Jaffe, W. FitzHugh, L. J. Ma, S. Smirnov, S. Purcell, B. Rehman, T. Elkins, R. Engels, S. Wang, C. B. Nielsen, J. Butler, M. Endrizzi, D. Qui, P. Ianakiev, D. Bell-Pedersen, M. A. Nelson, M. Werner-Washburne, C. P. Selitrennikoff, J. A. Kinsey, E. L. Braun, A. Zelter, U. Schulte, G. O. Kothe, G. Jedd, W. Mewes, C. Staben, E. Marcotte, D. Greenberg, A. Roy, K. Foley, J. Naylor, N. Stange-Thomann, R. Barrett, S. Gnerre, M. Kamal, M. Kamvysselis, E. Mauceli, C. Bielke, S. Rudd, D. Frishman, S. Krystofova, C. Rasmussen, R. L. Metzenberg, D. D. Perkins, S. Kroken, C. Cogoni, G. Macino, D. Catcheside, W. Li, R. J. Pratt, S. A. Osmani, C. P. DeSouza, L. Glass, M. J. Orbach, J. A. Berglund, R. Voelker, O. Yarden, M. Plamann, S. Seiler, J. Dunlap, A. Radford, R. Aramayo, D. O. Natvig, L. A. Alex, G. Mannhaupt, D. J. Ebbole, M. Freitag, I. Paulsen, M. S. Sachs, E. S. Lander, C. Nusbaum, and B. Birren.** 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. Nature **422:**859–868.
21. **Garceau, N. Y., Y. Liu, J. J. Loros, and J. C. Dunlap.** 1997. Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. Cell **89:**469–476.
22. **Görl, M., M. Merrow, B. Huttner, J. Johnson, T. Roenneberg, and M. Brunner.** 2001. A PEST-like element in FREQUENCY determines the length of the circadian period in *Neurospora crassa*. EMBO J. **20:**7074–7084.
23. **He, Q., J. Cha, H. C. Lee, Y. Yang, and Y. Liu.** 2006. CKI and CKII mediate the FREQUENCY-dependent phosphorylation of the WHITE COLLAR COMPLEX to close the *Neurospora* circadian negative feedback loop. Genes Dev. **20:**2552–2565.
24. **He, Q., and Y. Liu.** 2005. Molecular mechanism of light responses in *Neurospora*: from light-induced transcription to photoadaptation. Genes Dev. **19:**2888–2899.
25. **Johnson, D. S., A. Mortazavi, R. M. Myers, and B. Wold.** 2007. Genome-wide mapping of *in vivo* protein-DNA interactions. Science **316:**1497–1502.
26. **Johnson, L., X. Cao, and S. Jacobsen.** 2002. Interplay between two epigenetic marks. DNA methylation and histone H3 lysine 9 methylation. Curr. Biol. **12:**1360–1367.
27. **Káldi, K., B. H. Gonzalez, and M. Brunner.** 2006. Transcriptional regulation of the *Neurospora* circadian clock gene *wc-1* affects the phase of circadian output. EMBO Rep. **7:**199–204.
28. **Lauter, F. R., V. E. Russo, and C. Yanofsky.** 1992. Developmental and light regulation of *eas*, the structural gene for the rodlet protein of *Neurospora*. Genes Dev. **6:**2373–2381.
29. **Lewis, Z. A., A. Correa, C. Schwerdtfeger, K. L. Link, X. Xie, R. H. Gomer, T. Thomas, D. J. Ebbole, and D. Bell-Pedersen.** 2002. Overexpression of WHITE COLLAR-1 (WC-1) activates circadian clock-associated genes, but is not sufficient to induce most light-regulated gene expression in *Neurospora crassa*. Mol. Microbiol. **45:**917–931.
30. **Li, R., Y. Li, K. Kristiansen, and J. Wang.** 2008. SOAP: short oligonucleotide alignment program. Bioinformatics **24:**713–714.
31. **Linden, H., and G. Macino.** 1997. WHITE COLLAR-2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in *Neurospora crassa*. EMBO J. **16:**98–109.
32. **Liu, Y., J. Loros, and J. C. Dunlap.** 2000. Phosphorylation of the *Neurospora* clock protein FREQUENCY determines its degradation rate and strongly influences the period length of the circadian clock. Proc. Natl. Acad. Sci. U. S. A. **97:**234–239.
33. **Luo, C., J. J. Loros, and J. C. Dunlap.** 1998. Nuclear localization is required for function of the essential clock protein FRQ. EMBO J. **17:**1228–1235.
34. **Mikkelsen, T. S., M. Ku, D. B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T. K. Kim, R. P. Koche, W. Lee, E. Mendenhall, A. O'Donovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch, C. Nusbaum, E. S. Lander, and B. E. Bernstein.** 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature **448:**553–560.
35. **Ni, M., and J. H. Yu.** 2007. A novel regulator couples sporogenesis and trehalose biogenesis in *Aspergillus nidulans*. PLoS One **2:**e970.
36. **Olmedo, M., C. Ruger-Herreros, and L. M. Corrochano.** 2010. Regulation by blue light of the *fluffy* gene encoding a major regulator of conidiation in *Neurospora crassa*. Genetics **184:**651–658.
37. **Pavesi, G., P. Mereghetti, G. Mauri, and G. Pesole.** 2004. Weeder Web: discovery of transcription factor binding sites in a set of sequences from co-regulated genes. Nucleic Acids Res. **32:**W199–W203.
38. **Pomraning, K. R., K. M. Smith, and M. Freitag.** 2009. Genome-wide high throughput analysis of DNA methylation in eukaryotes. Methods **47:**142–150.
39. **Sargent, M. L., W. R. Briggs, and D. O. Woodward.** 1966. Circadian nature of a rhythm expressed by an invertaseless strain of *Neurospora crassa*. Plant Physiol. **41:**1343–1349.
40. **Schafmeier, T., A. Diernfellner, A. Schafer, O. Dintsis, A. Neiss, and M. Brunner.** 2008. Circadian activity and abundance rhythms of the *Neurospora* clock transcription factor WCC associated with rapid nucleo-cytoplasmic shuttling. Genes Dev. **22:**3397–3402.
41. **Schafmeier, T., A. Haase, K. Kaldi, J. Scholz, M. Fuchs, and M. Brunner.** 2005. Transcriptional feedback of *Neurospora* circadian clock gene by phosphorylation-dependent inactivation of its transcription factor. Cell **122:**235–246.
42. **Stein, L. D., C. Mungall, S. Shu, M. Caudy, M. Mangone, A. Day, E. Nickerson, J. E. Stajich, T. W. Harris, A. Arva, and S. Lewis.** 2002. The generic genome browser: a building block for a model organism system database. Genome Res. **12:**1599–1610.
43. **Taghert, P. H., and O. T. Shafer.** 2006. Mechanisms of clock output in the *Drosophila* circadian pacemaker system. J. Biol. Rhythms **21:**445–457.
44. **Talora, C., L. Franchi, H. Linden, P. Ballario, and G. Macino.** 1999. Role of a WHITE COLLAR-1-WHITE COLLAR-2 complex in blue-light signal transduction. EMBO J. **18:**4961–4968.
45. **Tamaru, H., X. Zhang, D. McMillen, P. B. Singh, J. Nakayama, S. I. Grewal, C. D. Allis, X. Cheng, and E. U. Selker.** 2003. Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. Nat. Genet. **34:**75–79.
46. **Thompson, S., N. J. Croft, A. Sotiriou, H. D. Piggins, and S. K. Crosthwaite.** 2008. *Neurospora crassa heat shock factor-1* is an essential gene; a second heat shock factor-like gene, *hsf-2*, is required for asexual spore formation. Eukaryot. Cell **7:**1573–1581.
47. **Vitalini, M. W., R. M. de Paula, W. D. Park, and D. Bell-Pedersen.** 2006. The rhythms of life: circadian output pathways in *Neurospora*. J. Biol. Rhythms **21:**432–444.
48. **Zhou, L. W., H. Haas, and G. A. Marzluf.** 1998. Isolation and characterization of a new gene, *sre*, which encodes a GATA-type regulatory protein that controls iron transport in *Neurospora crassa*. Mol. Gen. Genet. **259:**532–540.