

A Tool Set for the Genome-Wide Analysis of *Neurospora crassa* by RT-PCR

Jennifer M. Hurley,^{*1} Arko Dasgupta,^{*1} Peter Andrews,[†] Alexander M. Crowell,^{*} Carol Ringelberg,^{*} Jennifer J. Loros,[‡] and Jay C. Dunlap^{*2}

^{*}Department of Genetics and [†]Department of Biochemistry, Geisel School of Medicine, Dartmouth College, Hanover, New Hampshire 03755, and [‡]Institute for Biomedical Informatics, Perelman School of Medicine, Philadelphia, Pennsylvania 19104

ABSTRACT *Neurospora crassa* is an important model organism for filamentous fungi as well as for circadian biology and photobiology. Although the community-accumulated tool set for the molecular analysis of *Neurospora* is extensive, two components are missing: (1) dependable reference genes whose level of expression are relatively constant across light/dark cycles and as a function of time of day and (2) a catalog of primers specifically designed for real-time PCR (RT-PCR). To address the first of these we have identified genes that are optimal for use as reference genes in RT-PCR across a wide range of expression levels; the mRNA/transcripts from these genes have potential for use as reference noncycling transcripts outside of *Neurospora*. In addition, we have generated a genome-wide set of RT-PCR primers, thereby streamlining the analysis of gene expression. In validation studies these primers successfully identified target mRNAs arising from 70% (34 of 49) of all tested genes and from all (28) of the moderately to highly expressed tested genes.

KEYWORDS

Neurospora crassa
reference genes
circadian
light induction
quinic acid
induction

Profiling gene expression at the mRNA level via real-time PCR (RT-PCR) is a rapid means of assessing an organism's response to changing environmental conditions. In comparison to other techniques that also track the steady-state levels of mRNA such as Northern blot analysis, microarrays, and RNA-Seq, RT-PCR is cheaper, faster, requires less mRNA, and produces large amounts of data with limited effort. Two key pieces contributing to the success of this technique are (1) the efficient design of optimal primers that are used in the reaction and (2) reference genes to serve as internal controls for template inputs. RT-PCR primer design must take into account sequence position, exon position, product size, melting temperature (T_m), secondary structure, and GC content, as well as the terminal nucleotides to create an optimal primer. To achieve this, programs to aid in primer design have been developed (Rozen and Skaletsky 2000; Untergasser *et al.* 2012) as well

as Internet-based primer design resources (*e.g.*, Perfect Primer; Invitrogen, Waltham, MA). However, existing Internet-based primer design web sites are not high-throughput, leaving the user to provide individually not only the gene but also the genomic sequence of interest. Designing individual primers for target genes is a time-consuming process, and very often multiple primers for a single gene are required as part of the validation process; designing multiple RT-PCR primers per gene for multiple genes can quickly become tedious, and a genome-scale data set is optimal for large-scale efforts (Cui *et al.* 2007).

Neurospora has long been a salient model for nonyeast fungal species, including important plant and animal pathogens, and has been defined as a model organism by the NIH (<http://www.nih.gov/science/models/>). Its 43-Mb genome contains approximately 11,000 predicted genes (Galagan *et al.* 2003; Stajich *et al.* 2012) (<http://fungidb.org/fungidb/>). Haploid progeny, regulatable promoters (Hurley *et al.* 2012; Lamb *et al.* 2013; Giles *et al.* 1985), efficient and simple transformations (Chakraborty *et al.* 1991; Margolin *et al.* 1997), as well as an extensive culture collection (Fungal Genetics Stock Center) (McCluskey *et al.* 2010) and knockout library (Colot *et al.* 2006) (Dunlap *et al.* 2007) contribute to the utility of *Neurospora* for the study of a wide variety of biological processes, including many that are not fungal-specific. For example, in addition to being an established model organism for the study of circadian clock, *Neurospora* is commonly used to study fungal biomass deconstruction, epigenetics, fungal infection model, cell morphology, sexual development, gene silencing, and photobiology (Aramayo and Selker 2013; Riquelme *et al.* 2011; Znameroski and Glass

Copyright © 2015 Hurley *et al.*

doi: 10.1534/g3.115.019141

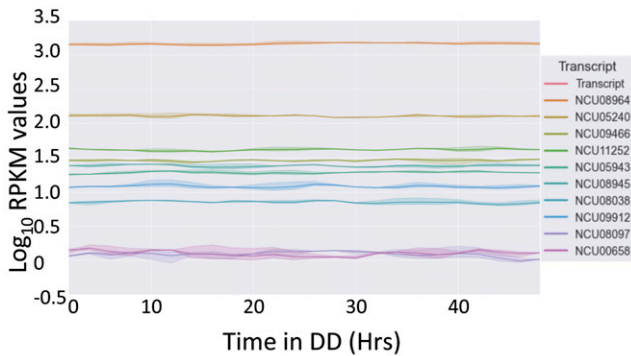
Manuscript received June 24, 2015; accepted for publication August 3, 2015; published Early Online August 6, 2015.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supporting information is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.115.019141/-/DC1

¹These authors contributed equally to this work.

²Corresponding author: Geisel School of Medicine at Dartmouth, HB 7400, Hanover NH 03755. E-mail: jay.c.dunlap@dartmouth.edu



NCU Number	Gene Symbol	Gene Product
NCU00658	<i>mb-3</i>	MALE BARREN-3
NCU05240	N/A	hypothetical protein
NCU05943	N/A	hypothetical protein
NCU08038	<i>apf-2</i>	CAS1
NCU08097	<i>gh55-5</i>	EXO-BETA-1,3-GLUCANASE
NCU08945	<i>exr-7</i>	RNA EXONUCLEASE 3
NCU08964	<i>crp-43</i>	60S RIBOSOMAL PROTEIN L10
NCU09466	N/A	hypothetical protein
NCU09912	N/A	MFS TRANSPORTER
NCU11252	<i>div-40</i>	CYCLIN PCH1

Figure 1 Optimal reference genes for RT-PCR in *Neurospora*. A graphical representation of the \log_{10} of FPKM values from the RNA-Seq data set for the 10 *Neurospora* genes (two reported from each quintile) that demonstrated the least variability in expression over 2 d in culture. The chart reports the gene name as well as the gene symbol for each of the NCUs reported. Gene symbols are from the *Neurospora* e-Compendium at Leeds (http://www.bioinf.leeds.ac.uk/~gen6ar/newgenelist/genes/gene_list.htm).

2013; Baker *et al.* 2012; Chen *et al.* 2009; Kuo *et al.* 2014; Lehr *et al.* 2014; Wang *et al.* 2014b; Fulci and Macino 2007).

Recently, extensive analyses of mRNA output via deep sequencing of mRNA levels as well as microarray analyses in *Neurospora* have thoroughly described circadian gene expression, as well as the organism's response to light and to addition of quinic acid (Hurley *et al.* 2014; Logan *et al.* 2007; Wu *et al.* 2014; Sancar *et al.* 2015). By probing these data sets, we have identified genes that can potentially serve as stably expressed reference genes when tracking changes in mRNA levels under various experimental conditions. The identified genes have a broad range of expression levels, providing a variety of comparable and appropriate normalizing genes for any target gene of interest. In addition, we report the development of a genome-wide data set of RT-PCR primers to streamline RT-PCR analyses. To create this resource, we designed and implemented a program to generate a list of five optimal RT-PCR primers for each gene in the *Neurospora* genome. We used the industry-standard program Primer3 (Untergasser *et al.* 2012; Rozen and Skaletsky 2000) to identify multiple primer pairs that are statistically likely to generate reliable gene detection via RT-PCR studies and validated a selection of the primer pairs (0.5%) by testing their utility on target genes selected to span the range of gene expression levels.

MATERIALS AND METHODS

Primer design

A software application was written in Python to retrieve regions from each ORF (designated from Assembly 12 of the *Neurospora* genome; www.broadinstitute.org/annotation/genome/neurospora/GenomeDescriptions.html#NC12; <http://fungidb.org/fungidb/>; <http://www.ncbi.nlm.nih.gov/>

genome/?term=txid367110) (Stajich *et al.* 2012) (Supporting Information, File S1) and pass them to Primer3 (version 2.3.6; http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi), which automatically generated a list of candidate primers with designated parameters (Table S1). Three to five primer pairs were selected and primers were ordered for 49 genes and supplied in 96-well plates (Illumina, San Diego, CA), diluted, and mixed by a Biomek NX robot (Beckman, Danvers, MA). The 49 genes were selected because they were potentially circadianly regulated genes based on previous work (Hurley *et al.* 2014).

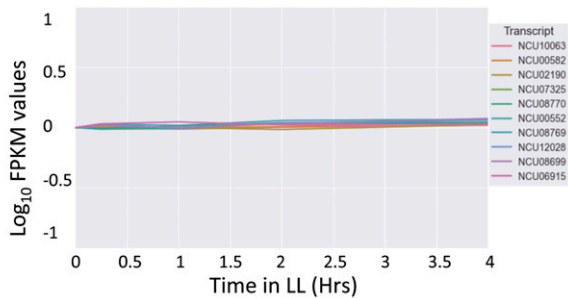
RT-PCR protocols

One-thousand five-hundred ng of RNA [a mix of the 10- and 30-hr time points from the third RNA-Seq time course published in the work by Hurley *et al.* (2014) chosen to encompass maximal expression of genes at two different clock phases] was used to prepare cDNA using the SuperScript III First-Strand synthesis kit (Invitrogen, Waltham, MA). This was followed by RT-PCR using the Fast SYBR green master mix kit in an ABI 7500 real-time cycler (Applied Biosystems, Waltham, MA). The primer combinations used are listed in Table S1 and the final concentration of primers in the reaction mix was 0.5 μ M. The following cycling parameters were used: step 1: 95° for 5 min and step 2: 95° for 10 sec and 60° for 30 sec for 40 cycles. C_t values were calculated using software provided by the instrument manufacturer. 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 three-fold. Only a single biological replicate and primer concentration was tested.

Data analysis

The functionality of each of the 49 selected primer pairs was tested. Each unique primer pair was assigned a C_t value with respect to its target by the ABI 7500 Fast Real-Time PCR software (Applied Biosystems, Waltham, MA) using the auto C_t command. Each C_t value represents a single technical replicate from the mix of the cDNA sample.

To determine which genes showed the least variation across time and would be optimal for use as reference genes over circadian time, we first rejected extremely low-expressing and nondetected genes with a log-transformed fragments per kilobase per million (FPKM) value cutoff of -2 from our RNA-Seq data. Next, we used our Jonckheere-Terpstra-Kendall (JTK) cycle (Hughes *et al.* 2010) results from Hurley *et al.* (2014), accounting for replicates, to reject genes called with a high confidence as being circadianly regulated (Benjamini Hochberg corrected $q < 0.05$). We then performed an ANOVA on the \log_{10} FPKM values for the remaining genes, comparing within time point variation against between time point variation for each gene and rejected genes for which $P < 0.05$. Because we chose genes to reject, we did not adjust for multiple comparisons because that would introduce more type I errors while reducing type II errors; our goal is to screen out genes with strong variation and the cutoff is itself arbitrary, making such an adjustment unnecessary. This left us with a set of expressing, noncircadian, ANOVA selected (ENCAS) genes. In the last step of our analysis, linear regressions were performed on the same \log_{10} FPKM values used for all previous calculations for each ENCAS gene using the scikit learn package (a set of data mining and machine learning tools) in python. These linear regressions were used to create prediction intervals (95%) for the expected mean at each time point for each gene. Genes were then scored based on the sum of the absolute value of the differences of the upper and lower bounds of these prediction intervals with respect to the overall mean expression of the gene at each time point. Under this



NCU Number	Gene Symbol	Gene Product
NCU00545	N/A	hypothetical protein
NCU01920	N/A	hypothetical protein
NCU02318	N/A	hypothetical protein
NCU04413	N/A	UPF0132 DOMAIN-CONTAINING PROTEIN
NCU06857	N/A	hypothetical protein
NCU06929	<i>trf-3</i>	TRANSCRIPTION ELONGATION FACTOR SPT-5
NCU07113	N/A	hypothetical protein
NCU07497	N/A	hypothetical protein
NCU08941	<i>mic-32</i>	CALCIUM-BINDING MITOCHONDRIAL CARRIER PROTEIN ARALAR2
NCU09492	N/A	hypothetical protein

Figure 2 Optimal reference genes for RT-PCR under light induction in *Neurospora*. A graphical representation of the \log_{10} of FPKM values from the RNA-Seq data set for the 10 *Neurospora* genes that demonstrated the least variability following light induction. Note that the light-induction data are normalized to the initial time point for each gene. The chart reports the gene name as well as gene symbol for each of the NCUs reported. Gene symbols are from the *Neurospora e-Compendium* at Leeds.

scoring metric, which we have termed the prediction interval ranking score (PIRS), genes that are more stable across circadian time will have lower scores. For both the light-induction and quinic acid-induction (QA) time courses, neither the JTK cycle nor the ANOVA-based screening methods were applicable because the time courses were not circadian in nature and did not contain replicates. It is important to note that the light-induction data were all normalized to the initial time point for each gene; because of this, we could not examine absolute expression levels and identify the best genes from each expression quintile for the light-induction data set.

In addition, as an independent analysis, normalized, \log -transformed FPKM values (Hurley *et al.* 2014) were averaged across each of the three reported circadian time courses as well as averaged across an average of the three time courses, and the SD of the \log_{10} FPKM values for each gene over circadian time was determined. Genes where the SD of the \log_{10} FPKM values was 0.5% or less of the \log_{10} FPKM averages were selected as candidates. The list of genes was compiled by selecting genes that appeared in the averaged time course as well as at least one of the three individual time courses (Hurley *et al.* 2014) (Figure S1).

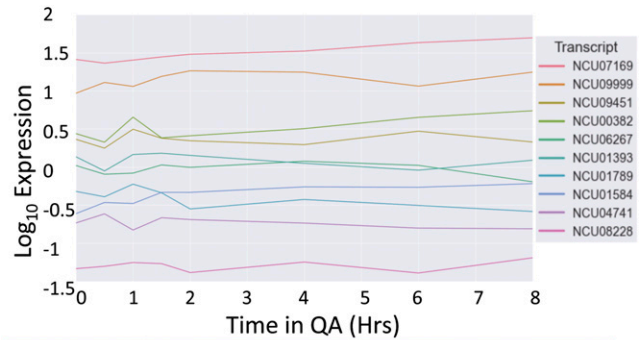
Data availability

All strain, data analysis programs and ancillary information are either freely available or available upon request. Table S1 and File S1 contain all necessary information regarding the generated RT-PCR primers.

RESULTS

Identifying a constitutively expressed reference gene

To identify genes with near-constant expression in *Neurospora*, we examined the global gene expression profiles in *Neurospora* informed by previous RNA-Seq and microarray experiments/studies (Hurley



NCU Number	Gene Symbol	Gene Product
NCU00382	N/A	hypothetical protein
NCU01393	<i>gh18-10</i>	CLASS III CHITINASE
NCU01584	N/A	hypothetical protein
NCU01789	<i>ham-5</i>	HYPHAL ANASTOMOSIS-5
NCU04741	<i>dash-7</i>	hypothetical protein
NCU06267	<i>apg-15</i>	AUTOPHAGY PROTEIN 16
NCU07169	<i>sut-23</i>	MFS GLUCOSE TRANSPORTER
NCU08228	<i>gld-2</i>	UDP-GLUCOSE DEHYDROGENASE
NCU09451	N/A	hypothetical protein
NCU09999	N/A	hypothetical protein

Figure 3 Optimal reference genes for RT-PCR under quinic acid (QA) induction. A graphic representation of the \log_{10} of FPKM values from the RNA-Seq data set for the 10 *Neurospora* genes (two reported from each quintile) that demonstrated the least variability after the addition of QA. The chart reports the gene name as well as gene symbol for each of the NCUs reported. Gene symbols are from the *Neurospora e-Compendium* at Leeds.

et al. 2014; Wu *et al.* 2014; Logan *et al.* 2007). Using our analysis (*Materials and Methods*) we assigned a prediction interval ranking score (PIRS) to each gene. The resulting PIRS was minimized for genes that exhibit both flat expression as a function of time and low variance in observed expression at each time point (Figure 1). The gene ranks produced by our scoring metric were compared to rankings based on SD and relative SD (RSD) of expression using the Jaccard similarity coefficient (Levandowsky and Winter 1971) [Figure S1 and Figure S2 (Hintze and Nelson 1998)]. The rankings produced by our metric (PIRS) corresponded very well with SD and were noticeably more similar to RSD than random ranking. The majority of the difference between our method and RSD likely resulted from RSD's preference for highly expressing genes, because the similarity score markedly improved when comparing the highest expressing quintile of genes under our scoring metric with the RSD results. A similar analysis was applied to the light-induced and QA-induced data sets with a few differences (Figure 2 and Figure 3). The selection criteria for nonlight-induced genes and non-QA-induced genes were identical to those for noncircadian genes except that the removal of noncircadian genes and the ANOVA were omitted (due to lack of replicate datasets).

For both the circadian and QA induction datasets, we selected the top two genes in each quintile of expression (Figure 1 and Figure 3). Because only relative expression data were available for the light-induction time course, we have reported the top 10 scoring genes overall (Figure 2). The top ranking genes in each quintile outperformed the average gene, whereas the top three quintiles were markedly higher than the average gene [Figure S3 (Scott 1992)]. In all cases, the genes selected by our metric show flat expression with low variation, whereas genes that rank poorly show high variation both within and between time points (compare Figure 1, Figure 2, and Figure 3 to Figure S4). The

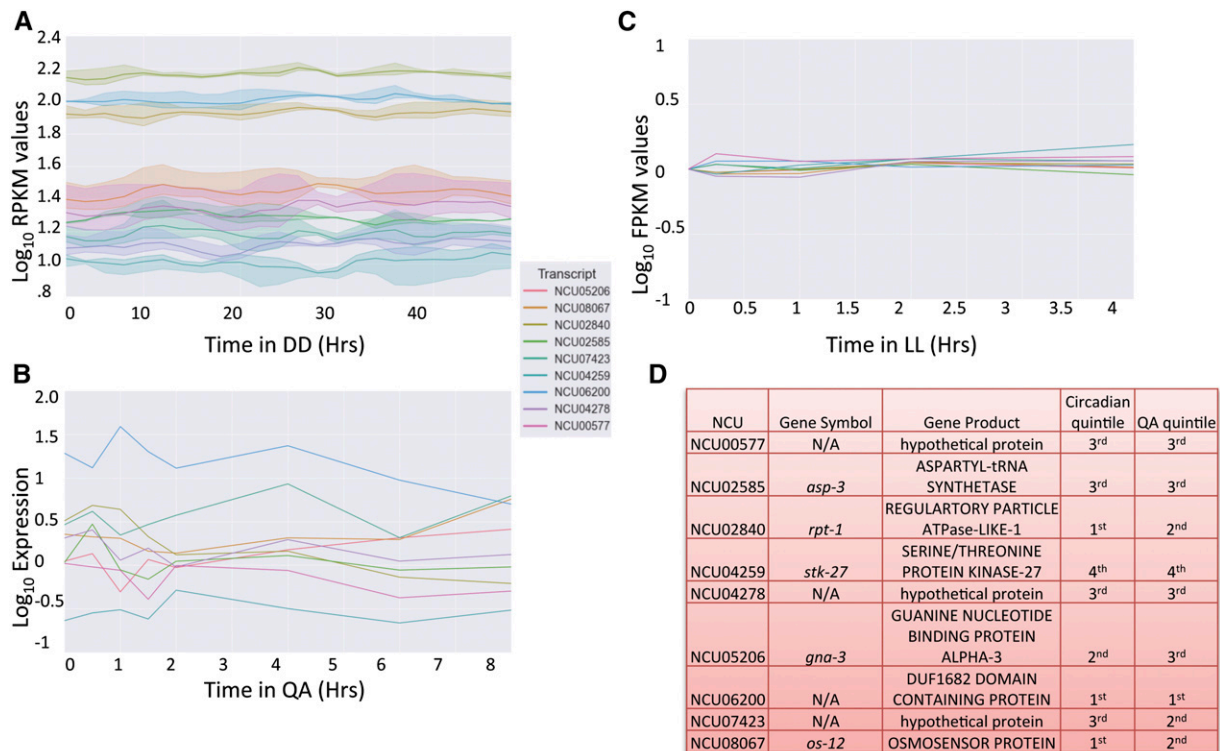


Figure 4 Least variably expressed genes for circadian RT-PCR in *Neurospora*. (A–C) A graphic representation of the \log_{10} expression values from the (A) circadian, (B) quinic acid, and (C) light-induction data sets for the nine *Neurospora* genes in each category that were reported as the most stably expressed according to pareto optimization and normalized summed PIRS. (D) The chart reports the NCUs plotted in (A–C) with the quintile of expression associated with each gene from the circadian and QA time courses as well as the gene symbol and name for each of the NCUs reported. Gene symbols are from the *Neurospora* e-Compendium at Leeds.

average FPKM values of the noncircadian genes ranged from \log_{10} 0.087 to \log_{10} 3.16, whereas the non-QA-induced genes ranged from \log_{10} -1.30 to \log_{10} 1.49, giving a wide variety of expression levels from which to choose from when selecting the appropriate reference gene. Not surprisingly, of the genes with predicted function, classifications included ribosomal, cytoskeletal, or protease genes, all proteins that could logically be expected to be housekeeping genes.

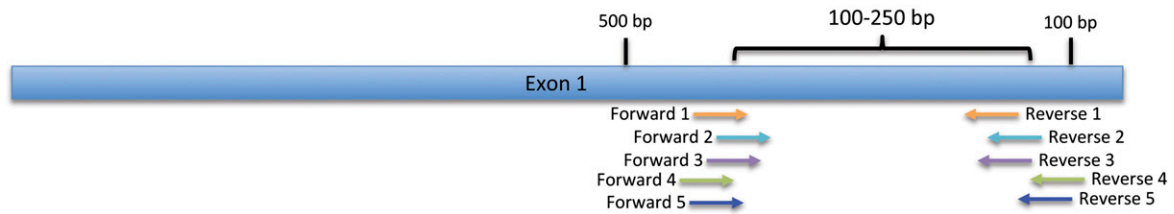
As a way to identify genes that could be used as reference genes, we subjected all of the genes that were in our ENCAS set to pareto optimization to select nondominated genes across the three data sets. Pareto optimization is a technique for multi-objective optimization that aims to find the set of all choices that might be considered optimal under any of the possible combinations of weightings of individual scores (Pareto 1906). The algorithm rejects any choice for which another choice is available with superior scores under every metric. Genes for which there is an alternative gene with superior scores under all available metrics are termed "dominated" and genes for which any scoring metric exists under which they are optimal are termed "nondominated." This produced a list of 27 genes from the three data sets. Because it was possible that some of the genes on the pareto plane represent an optimal combination of PIRS for one experiment individually as opposed to all experiments in combination, we then compared these 27 genes to the summed normalized PIRS scores (Figure 4). We selected the overlap between the top 10 summed normalized scores and the pareto optimized set, leaving us with nine genes as good candidates for reference. These nine genes spanned the top four quintiles of expression (Figure 4), suggesting genes buffered against variation are not specific to a given level of expression except at the limits of detection (Figure S3).

Generation of a genome-wide RT-PCR primer data set

Although there are many ways to analyze and validate mRNA levels and consistency between any two methods constitutes validation, it has become common practice in *Neurospora* to test mRNA steady-state levels using RT-PCR (Hurley *et al.* 2012, 2014; Wang *et al.* 2014a; Sancar *et al.* 2012; Hong *et al.* 2014; Hutchison *et al.* 2009). To analyze mRNA, RT-PCR primers are generally designed for each gene by hand, a time-consuming process when done on a large scale. To streamline this process, a catalog of RT-PCR primers was created for each gene designated in *Neurospora* by the Broad database. To achieve this we created a program that would design five primer sets for each gene based on the standard RT-PCR primer criteria (see *Materials and Methods*). In addition, each gene should have five primer pair sets that are not identical to one another. In the case of mono-exonic genes, the primer pairs were all selected from sequence in the last 500 base pairs of the gene (Figure 5). In the case of multi-exonic genes with a single intron, the primer pairs were designed to encompass the intron; in multi-exonic genes with multiple introns, the primer pairs were designed to encompass the last intron in the gene (Figure 5).

Using our program we were able to successfully design primers for 10,798 (99.9%) of the 10,802 ORFs predicted by Assembly 12 of the *Neurospora* genome (www.broadinstitute.org/annotation/genome/neurospora/GenomeDescriptions.html#NC12; <http://fungidb.org/fungidb/>; <http://www.ncbi.nlm.nih.gov/genome/?term=txid367110>) (Stajich *et al.* 2012) (File S1). The outline of the database can be seen in Table 1. For each transcript, the number of primer pairs that was designed is noted. In a few cases (five total), fewer than five unique primer pairs were designed; in theory, this is due to the limitations of

Case 1 – Mono-exonic genes



Case 2 – Multi-exonic genes

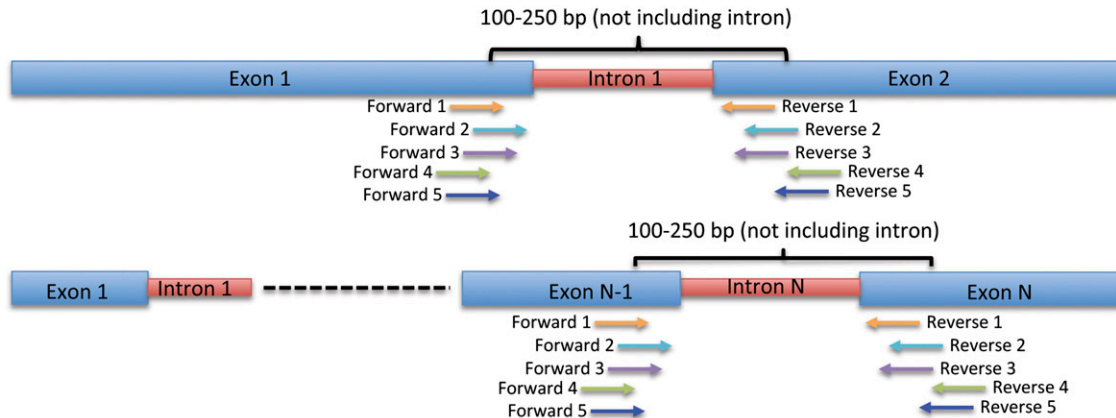


Figure 5 Criteria for the selection of RT-PCR primers. A diagrammatic representation of the rules used for primer design. In all cases, the primers are designed to create fragments between 100 and 250 bp in length. In the case of genes with no introns (mono-exonic genes), five unique primer pairs are designed to fall between 500 and 100 bp from the end of the gene. In the case of a gene with one intron, the primer pairs are designed to exclude the intron. In the case of genes with more than one intron, the primer pairs are designed to exclude the last intron in the gene, no matter how many introns there are.

the gene sequence and the parameters that were set. Each primer pair has a penalty score associated with it and that score, assigned by Primer3, represents how close the primer pair comes to the optimum conditions set by the program. The lower the score, the better the primer pair. The pair with the lowest score is ordered first, the second lowest score is second, and this pattern continues through to the fifth pair.

Analyzing the generated RT-PCR primers

To confirm that the primers generated by our program were capable of producing valid RT-PCR products, we tested 49 genes with three or more of the generated primer pairs (Table 2). In our work, transcripts with low abundance generally give rise to highly variable RT-PCR estimates (Table 2). As a standard, we decided C_t values had to be below 30 to consider the cDNA target as being reliably detected by RT-PCR. We noted that genes that we classified as significantly amplified had an average Primer3 primer score of 0.4, whereas genes that were classified as not significantly amplified had an average Primer3 primer score of 0.9, suggesting that there is a correlation between Primer3 primer score and the successful amplification of PCR products.

Of the 49 genes analyzed, 34 (69%) returned a C_t value under 30 in at least one of the primer pairs (Table 2), meaning that these primer pairs were viable for use in RT-PCR studies. For medium to highly expressed genes with an average C_t value under 28, all of the primer pairs (28) were successfully able to return C_t values under 30. The remaining 15 (31%) genes had C_t values over 30 across the three primer pairs, indicating that they were very poorly expressed. For five of the 15 genes that had C_t values over 30, we tested the remaining two primer sets to see if we could obtain a reliable signal. None of the genes tested were able to generate meaningful C_t values, suggesting that genes that were

not detected by our designed primers are either in extremely low abundance under the conditions tested or poor candidates for RT-PCR analysis rather than being an inherent program primer design failure. Furthermore, when comparing the average C_t values of all the tested primer pairs for genes with C_t values below 30 to the average C_t values of test primer pairs for genes with C_t values above 30, the SD of the average C_t value was strikingly lower in genes with C_t values below 30 than those genes with C_t values over 30 (Table 2) (*i.e.*, compare NCU06651 to NCU6799). The average SD for genes where all primer pairs reported a C_t values below 30 was 0.52 (this calculation does not include genes where some of the primer pairs had C_t values above 30), whereas the average SD of those genes with a C_t value above 30 (this calculations includes only the genes where a SD could be determined) was 2.45.

DISCUSSION

Finding a functional and truly constitutive “housekeeping” gene to serve as a reference gene in RT-PCR has often been difficult because changing cell or experimental conditions have been shown to have a large influence on gene expression levels (Schmittgen and Zakrajsek 2000), *i.e.*, reference genes tend to change expression levels under different conditions. This problem is augmented in circadian biology because many commonly used RT-PCR reference genes have been shown to be regulated by the circadian clock (*e.g.*, GAPDH) (Shinohara *et al.* 1998). Here, we report a partially validated genome-wide set of RT-PCR primers for *Neurospora* that reliably reports mRNA levels for all moderately to highly expressed genes and for 70% of all genes. There are many reasons for a less than perfect reporter rate, the most likely being that we only investigated the mRNA levels in senescent hyphal growth in

■ Table 1 Setup of the RT-PCR primer data set

Gene Transcript	# of Primers Pairs	Set 1 Penalty	Set 1 Left Primer	Set 1 Right Primer	Set 2 Penalty	Set 2 Left Primer	Set 2 Right Primer	Set 3 Penalty	Set 3 Left Primer	Set 3 Right Primer	Set 4 Penalty	Set 4 Left Primer	Set 4 Right Primer	Set 5 Penalty	Set 5 Left Primer	Set 5 Right Primer
NCU00003T0	5	0.209787	TTTCAG..	TGAAAA..	0.496616	CCTCCGC..	TGAAAA..	0.634468	TTTCAG..	GAGAAC..	0.634468	TTTCAG..	GGAGAA..	0.780732	CCTCCG..	TCGCTG..
NCU00004T0	5	0.579651	GTGCGC..	TTCATCA..	0.696041	CGAAGA..	GGACAG..	0.731866	GAATCC..	TTTATC..	0.923985	CCAGGA..	GGACAG..	0.951922	TTCGTC..	TTCATC..

An example of the contents of the RT-PCR primer data set. Each transcript in Neurospora is listed with five primer pairs. Each primer pair is assigned a penalty score by Primer3 that represents the strength of the primer pair: the lower the score, the closer the primer pair is to the required/enforced criteria.

a single type of growth medium. If we tracked mRNA extracted from a number of different stages of the Neurospora life cycle in a variety of different nutrient conditions, it seems likely that a greater percentage of gene targets would be more highly expressed and could be identified by our primers. In addition, we did not discern if the mRNAs investigated by this study were known to be poor candidates for RT-PCR for some unknown reason (*i.e.*, high levels of secondary structure). To develop manually a single RT-PCR primer sequence that meets all of the standards set forth in our criteria takes approximately 5–10 min, suggesting that these predesigned primers may afford considerable time-savings.

Circadian rhythms present a unique challenge for those investigating mRNA levels via RT-PCR. Because as much as 40% of the genome could be rhythmic at the mRNA level (Hurley *et al.* 2014), finding genes that show little change in expression over the circadian day to serve as reference genes has been difficult. Even well-known housekeeping genes that are used regularly as RT-PCR reference genes have been shown to cycle over circadian time (Shinohara *et al.* 1998). In this work, we used an in-depth investigation into mRNA expression levels to identify genes that demonstrated little change at the steady-state mRNA level over circadian time, light induction, and QA response (Hurley *et al.* 2014). We identified nine potential candidates that could

■ Table 2 Designed RT-PCR primers detect mRNAs in Neurospora

Gene	Set 1	Set 2	Set 3	Set 4	Set 5	Average CT value	Standard Deviation	#s	Significant Amplification CT<30	AVE=0.5
NCU00019	35.84708 (0.5)	32.21396 (1.2)	29.7343 (1.2)	NT	NT	32.59845	3.07447			
NCU00275	NA (1.5)	39.21742 (1.6)	NA (1.7) (0.7)	NT	NT	39.21742	***			
NCU00285	22.43273 (0.5)	21.74351 (0.6)	22.69039 (0.7)	NT	NT	22.28888	0.48956			
NCU00289	21.58568 (0.1)	21.90976 (0.1)	22.18193 (0.2)	NT	NT	21.89246	0.29850			
NCU00340	NA (1.7)	NA (2.1)	36.64867 (2.1)	NT	NT	36.64867	***			
NCU00399	38.97393 (0.2)	33.12985 (0.3)	35.47003 (0.6)	NT	NT	35.85794	2.94129			
NCU00401	36.53366 (0.5)	28.89523 (0.9)	31.16346 (1.2)	NT	NT	32.19745	3.92279			
NCU00499	18.11522 (0.1)	18.05121 (0.4)	18.5345 (0.6)	NT	NT	18.23364	0.26251			
NCU00552	34.82531 (0.5)	NA (0.6)	34.4582 (0.8)	NT	NT	34.64176	***			
NCU00582	25.8184 (0.6)	26.04333 (1.3)	25.31097 (1.4)	NT	NT	25.72423	0.37515			
NCU00855	20.34 (0.1)	20.44 (0.1)	20.28 (0.3)	NT	NT	20.35333	0.08083			
NCU01118	36.02 (0.1)	NA (0.5)	NA (0.5)	NT	NT	36.02000	***			
NCU01119	36.22 (0.2)	34.09 (0.3)	34.89 (0.6)	NT	NT	35.06667	1.07593			
NCU01233	36.3 (0.4)	36.63 (0.4)	39.43 (0.6)	NT	NT	37.45333	1.71978			
NCU01348	37.73 (0.9)	N/A (1.4)	34.66 (1.4)	NT	NT	36.19500	***			
NCU01418	29.67219 (0.3)	36.39851 (0.3)	34.23722 (0.6)	NA (0.6)	37.42 (0.7)	34.43198	3.43943			
NCU01528	13.65319 (0.6)	14.25128 (0.8)	13.216 (1)	NT	NT	13.70682	0.51972			
NCU01810	21.51 (0.3)	21.01 (0.4)	21.6 (0.4)	NT	NT	21.37333	0.31786			
NCU01898	21.08712 (0.3)	21.8134 (0.3)	21.15356 (0.5)	NT	NT	21.35136	0.40151			
NCU02084	24.26303 (0.2)	24.15344 (0.5)	20.91727 (0.5)	NT	NT	23.11125	1.90083			
NCU02193	14.86928 (0.6)	15.06402 (0.8)	19.34261 (1)	NT	NT	16.42530	2.52834			
NCU02238	26.15509 (0.2)	26.57688 (0.3)	26.31321 (0.6)	NT	NT	26.34839	0.21308			
NCU02265	28.23747 (0.4)	27.64946 (0.5)	27.13745 (0.6)	NT	NT	27.67479	0.55045			
NCU02408	37.26475 (0.4)	34.58207 (0.7)	NA (0.8)	NT	NT	35.92341	***			
NCU02500	39.07831 (1.7)	NA (1.8)	35.49304 (1.8)	36.08 (1.8)	31.61 (1.9)	35.56534	3.06897			
NCU02505	29.82169 (0.8)	32.92965 (1.1)	33.17355 (1.1)	NT	NT	31.97496	1.86877			
NCU03362	19.42573 (0.3)	19.45164 (0.3)	19.50507 (0.3)	NT	NT	19.46081	0.04046			
NCU03489	28.43918 (0.4)	34.50522 (0.4)	31.61792 (0.4)	NT	NT	31.52077	3.03419			
NCU03501	32.6744 (0.3)	39.14352 (0.4)	36.7199 (0.6)	NT	NT	36.17927	3.26827			
NCU03753	39.62009 (0.3)	26.58978 (0.7)	33.75635 (0.9)	32.05 (0.9)	35.92 (0.9)	33.58724	4.82683			
NCU04041	23.85297 (0.3)	23.99246 (0.3)	23.97424 (0.4)	NT	NT	23.93989	0.07582			
NCU04923	25.18983 (0.3)	24.83653 (0.3)	25.50805 (0.3)	NT	NT	25.17814	0.33591			
NCU06001	33.3679 (0.6)	37.4062 (0.7)	37.1078 (0.8)	35.79 (0.8)	NA (0.8)	35.91798	1.83936			
NCU06651	30.88221 (0.4)	30.48683 (0.5)	30.63587 (0.6)	36.43 (0.8)	33.39 (0.9)	32.36498	2.56373			
NCU06799	25.40175 (0.3)	25.42869 (0.3)	25.71959 (0.3)	NT	NT	25.51668	0.17624			
NCU06961	22.11948 (0.2)	21.92295 (0.3)	22.11205 (0.3)	NT	NT	22.05149	0.11138			
NCU07139	21.03585 (0.3)	21.90745 (0.6)	20.85369 (0.8)	NT	NT	21.26566	0.56322			
NCU07318	31.23103 (0.9)	NA (1.2)	NA (1.2)	37.3 (1.5)	NA (1.5)	34.26552	***			
NCU07787	14.63946 (0.6)	14.40931 (0.7)	14.88278 (1.1)	NT	NT	14.64385	0.23677			
NCU07817	14.37452 (0.1)	15.2668 (0.1)	20.02864 (0.3)	NT	NT	16.55665	3.03975			
NCU08141	27.65954 (0.5)	27.46938 (0.6)	27.97097 (0.6)	NT	NT	27.69996	0.25323			
NCU08457	20.53625 (0.6)	20.49179 (0.7)	21.03932 (0.7)	NT	NT	20.68912	0.30410			
NCU08769	25.10282 (0.8)	25.646 (0.8)	25.59032 (0.9)	NT	NT	25.44638	0.29883			
NCU08791	18.92834 (0.1)	18.99877 (0.1)	19.12129 (0.2)	NT	NT	19.01613	0.09764			
NCU08907	14.76518 (0.6)	15.03193 (1)	14.43113 (1)	NT	NT	14.74275	0.30103			
NCU08936	33.90786 (1.3)	35.65217 (1.3)	NA (1.3)	37.51 (1.3)	30.06 (1.5)	34.28251	3.17609	#s	Significant Amplification CT<30	AVE=0.5
NCU09559	19.74438 (0.1)	20.38487 (0.2)	20.03251 (0.2)	NT	NT	20.05392	0.32078	#s	No Significant Amplification CT>30	
NCU09686	22.98228 (0.5)	23.15264 (0.5)	22.67511 (0.5)	NT	NT	22.93668	0.24201	NA	No Amplification	AVE=0.9
NCU09976	20.58221 (0.6)	21.0014 (0.7)	20.67222 (0.8)	NT	NT	20.75194	0.22067	NT	Not Tested	

A chart of C_t values from primer pairs from the RT-PCR primer data set tested using RT-PCR with the Primer3 primer scores noted next to the C_t values in parentheses. Scores represent a single biological replicate. Average C_t values are an average of all samples that had a readable C_t score, whereas SDs of C_t values are only presented for genes that had three or more readable C_t values in the analysis. The average value of the Primer3 primer scores of gene with or without significant amplification is noted next to the key. Blue boxed numbers represent significant amplification $C_t < 30$; red boxed numbers represent no significant amplification $C_t > 30$; red boxed NA, amplification; NT, not tested.

be used as controls for any experiments involving long-term culture, with mRNA levels that are minimally variable across multiple modes of regulation. Many of the genes identified have orthologs in higher eukaryotes (Figure 1, Figure 2, Figure 3, and Figure 4 and Figure S1) and therefore have the potential to be useful to other species that serve as model organisms.

ACKNOWLEDGMENTS

The authors thank the Fungal Genetics Stock Center at the Kansas State University for *Neurospora* strains. This work is supported by NIH-National Institute of General Medical Sciences Grants GM083336 (to J.J.L.), GM34985, and GM068087 (to J.C.D.).

LITERATURE CITED

- Aramayo, R., and E. U. Selker, 2013 *Neurospora crassa*, a model system for epigenetics research. *Cold Spring Harb. Perspect. Biol.* 5: a017921.
- Baker, C. L., J. J. Loros, and J. C. Dunlap, 2012 The circadian clock of *Neurospora crassa*. *FEMS Microbiol. Rev.* 36: 95–110.
- Chakraborty, B. N., N. A. Patterson, and M. Kapoor, 1991 An electroporation-based system for high-efficiency transformation of germinated conidia of filamentous fungi. *Can. J. Microbiol.* 37: 858–863.
- 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.
- Colot, H. V., G. Park, G. E. Turner, C. Ringelberg, C. M. Crew *et al.*, 2006 A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proc. Natl. Acad. Sci. USA* 103: 10352–10357.
- Cui, W., D. D. Taub, and K. Gardner, 2007 qPrimerDepot: a primer database for quantitative real time PCR. *Nucleic Acids Res.* 35(Database issue): D805–D809.
- Dunlap, J. C., K. A. Borkovich, M. R. Henn, G. E. Turner, M. S. Sachs *et al.*, 2007 Enabling a community to dissect an organism: overview of the *Neurospora* functional genomics project. *Adv. Genet.* 57: 49–96.
- Fulci, V., and G. Macino, 2007 Quelling: post-transcriptional gene silencing guided by small RNAs in *Neurospora crassa*. *Curr. Opin. Microbiol.* 10: 199–203.
- Galagan, J. E., S. E. Calvo, K. A. Borkovich, E. U. Selker, N. D. Read *et al.*, 2003 The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422: 859–868.
- Giles, N. H., M. E. Case, J. Baum, R. Geever, L. Huiet *et al.*, 1985 Gene organization and regulation in the qa (quinic acid) gene cluster of *Neurospora crassa*. *Microbiol. Rev.* 49: 338–358.
- Hintze, J. L., and R. D. Nelson, 1998 Violin Plots: A Box Plot-Density Trace Synergism. *Am. Stat.* 52: 181–184.
- Hong, C. I., J. Zamborszky, M. Baek, L. Labiscsak, K. Ju *et al.*, 2014 Circadian rhythms synchronize mitosis in *Neurospora crassa*. *Proc. Natl. Acad. Sci. USA* 111: 1397–1402.
- Hughes, M. E., J. B. Hogenesch, and K. Kornacker, 2010 JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J. Biol. Rhythms* 25: 372–380.
- Hurley, J. M., C. H. Chen, J. J. Loros, and J. C. Dunlap, 2012 Light-inducible system for tunable protein expression in *Neurospora crassa*. *G3 (Bethesda)* 2: 1207–1212.
- Hurley, J. M., A. Dasgupta, J. M. Emerson, X. Zhou, C. S. Ringelberg *et al.*, 2014 Analysis of clock-regulated genes in *Neurospora* reveals widespread posttranscriptional control of metabolic potential. *Proc. Natl. Acad. Sci. USA* 111: 6995–17002.
- Hutchison, E., S. Brown, C. Tian, and N. L. Glass, 2009 Transcriptional profiling and functional analysis of heterokaryon incompatibility in *Neurospora crassa* reveals that reactive oxygen species, but not metacaspases, are associated with programmed cell death. *Microbiology* 155(Pt 12): 3957–3970.
- Kuo, H. C., S. Hui, J. Choi, F. O. Asiegbu, J. P. Valkonen *et al.*, 2014 Secret lifestyles of *Neurospora crassa*. *Sci. Rep.* 4: 5135.
- Lamb, T. M., J. Vickery, and D. Bell-Pedersen, 2013 Regulation of gene expression in *Neurospora crassa* with a copper responsive promoter. *G3 (Bethesda)* 3: 2273–2280.
- Lehr, N. A., Z. Wang, N. Li, D. A. Hewitt, F. Lopez-Giraldez *et al.*, 2014 Gene expression differences among three *Neurospora* species reveal genes required for sexual reproduction in *Neurospora crassa*. *PLoS One* 9: e110398.
- Levandosky, M., and D. Winter, 1971 Distance between sets. *Nature* 234: 34–35.
- Logan, D. A., A. L. Koch, W. Dong, J. Griffith, R. Nilsen *et al.*, 2007 Genome-wide expression analysis of genetic networks in *Neurospora crassa*. *Bioinformatics* 1: 390–395.
- Margolin, B. S., M. Freitag, and E. U. Selker, 1997 Improved plasmids for gene targeting at the his-3 locus of *Neurospora crassa* by electroporation. *Fungal Genet. Newsl.* 44: 34–36.
- McCluskey, K., A. Wiest, and M. Plamann, 2010 The Fungal Genetics Stock Center: a repository for 50 years of fungal genetics research. *J. Biosci.* 35: 119–126.
- Pareto, V., 1906 *Manuale di economia politica, con una introduzione alla scienza sociale*. Societ' a Editrice Libraria, Milan.
- Riquelme, M., O. Yarden, S. Bartnicki-Garcia, B. Bowman, E. Castro-Longoria *et al.*, 2011 Architecture and development of the *Neurospora crassa* hypha—a model cell for polarized growth. *Fungal Biol.* 115: 446–474.
- Rozen, S., and H. Skaletsky, 2000 Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132: 365–386.
- Sancar, G., C. Sancar, and M. Brunner, 2012 Metabolic compensation of the *Neurospora* clock by a glucose-dependent feedback of the circadian repressor CSP1 on the core oscillator. *Genes Dev.* 26: 2435–2442.
- Sancar, C., G. Sancar, N. Ha, F. Cesbron, and M. Brunner, 2015 Dawn- and dusk-phased circadian transcription rhythms coordinate anabolic and catabolic functions in *Neurospora*. *BMC Biol.* 13: 17.
- Schmittgen, T. D., and B. A. Zakrajsek, 2000 Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. *J. Biochem. Biophys. Methods* 46: 69–81.
- Scott, D. W., 1992 *Multivariate Density Estimation: Theory, Practice, and Visualization*. New York, Chichester: John Wiley & Sons.
- Shinohara, M. L., J. J. Loros, and J. C. Dunlap, 1998 Glyceraldehyde-3-phosphate dehydrogenase is regulated on a daily basis by the circadian clock. *J. Biol. Chem.* 273: 446–452.
- Stajich, J. E., T. Harris, B. P. Brunk, J. Brestelli, S. Fischer *et al.*, 2012 FungiDB: an integrated functional genomics database for fungi. *Nucleic Acids Res.* 40(Database issue): D675–D681.
- Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth *et al.*, 2012 Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40: e115.
- Wang, B., A. N. Kettenbach, S. A. Gerber, J. J. Loros, and J. C. Dunlap, 2014a *Neurospora* WC-1 recruits SWI/SNF to remodel frequency and initiate a circadian cycle. *PLoS Genet.* 10: e1004599.
- Wang, Z., F. Lopez-Giraldez, N. Lehr, M. Farre, R. Common *et al.*, 2014b Global gene expression and focused knockout analysis reveals genes associated with fungal fruiting body development in *Neurospora crassa*. *Eukaryot. Cell* 13: 154–169.
- Wu, C., F. Yang, K. M. Smith, M. Peterson, R. Dekhang *et al.*, 2014 Genome-wide characterization of light-regulated genes in *Neurospora crassa*. *G3 (Bethesda)* 4: 1731–1745.
- Znameroski, E. A., and N. L. Glass, 2013 Using a model filamentous fungus to unravel mechanisms of lignocellulose deconstruction. *Biotechnol. Biofuels* 6: 6.

Communicating editor: R. B. Brem