En masse nascent transcription analysis to elucidate regulatory transcription factors

Jinshui Fan, Ming Zhan¹, Jikui Shen², Jennifer L. Martindale, Xiaoling Yang, Tomoko Kawai and Myriam Gorospe*

Laboratory of Cellular and Molecular Biology and ¹Research Resources Branch, National Institute on Aging-Intramural Research Program, National Institutes of Health, Baltimore, MD 21224, USA and ²Department of Ophthalmology, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

Received December 22, 2005; Revised and Accepted January 30, 2006

ABSTRACT
Despite exhaustively informing about steady-state mRNA abundance, DNA microarrays have been used with limited success to identify regulatory transcription factors (TFs). The main limitation of this approach is that altered mRNA stability also strongly governs the patterns of expressed genes. Here, we used nuclear run-on assays and microarrays to systematically interrogate changes in nascent transcription in cells treated with the topoisomerase inhibitor camptothecin (CPT). Analysis of the promoters of coordinately transcribed genes after CPT treatment suggested the involvement of TFs c-Myb and Rfx1. The predicted CPT-dependent associations were subsequently confirmed by chromatin immunoprecipitation assays. Importantly, after RNAi-mediated knockdown of each TF, the CPT-elicited induction of c-Myb- and/or Rfx1-regulated mRNAs was diminished and the overall cellular response was impaired. The strategies described here permit the successful identification of the TFs responsible for implementing adaptive gene expression programs in response to cellular stimulation.

INTRODUCTION
In mammalian cells, damaging stimuli trigger the stress response, which is characterized by the coordinate expression of subsets of genes through both transcriptional and post-transcriptional mechanisms. Knowledge of stress-triggered transcriptional regulation has increased vastly in recent years through the elucidation of the signaling pathways, chromatin alterations and transcription factors (TFs) involved. Our understanding of the ensuing changes in expressed mRNAs has also expanded spectacularly through the utilization of the microarray technology. However, a systematic identification of the links between transcriptional regulatory events and the subsets of expressed transcripts has remained elusive due to two major obstacles. First, the combinatorial nature of TF function upon gene promoters. Since several TFs often bind to the promoters of eukaryotic genes in order to activate transcription, studying their individual and joint regulation is generally quite complex (1,2). Second, the strong contribution of altered mRNA stability in determining the patterns of expressed genes. Given that changes in mRNA half-life potently control the collections of expressed mRNAs, alterations in mRNA abundance may reflect changes in mRNA turnover rates instead of transcription (3–5).

Here, we focus on the analysis of newly transcribed (nascent) mRNAs in an effort to identify shared regulatory promoter elements and hence the TFs responsible for coordinating gene expression. Using HeLa cells treated with the topoisomerase I inhibitor camptothecin (CPT) as model system, the transcription of thousands of genes was assessed simultaneously using the nuclear run-on (NRO) assay and cDNA arrays. Comparison of the promoters present in the genes whose transcription was most robustly induced revealed highly conserved TF-binding sites, including those for c-Myb and Rfx1. This approach successfully identified TFs that were pivotal for the cell’s response to genotoxic stress, as supported by additional studies demonstrating (i) the CPT-dependent association of c-Myb and Rfx1 with the promoters of predicted target genes, (ii) the requirement of c-Myb and Rfx1 for their transcriptional activation and (iii) the critical influence of Rfx1 and c-Myb on cell proliferation and survival after CPT treatment.

*To whom correspondence should be addressed at Box 12, LCMB, NIA-IRP, NIH, 5600 Nathan Shock Drive, Baltimore, MD 21224, USA. Tel: +1 410 558 8443; Fax: +1 410 558 8386; Email: myriam-gorospe@nih.gov
Correspondence may also be addressed to Ming Zhan, RRB, NIA-IRP, NIH, 5600 Nathan Shock Drive, Baltimore, MD 21224, USA. Tel: +1 410 558 8373; Fax: +1 410 558 8674; Email: zhanmi@mail.nih.gov

© The Author 2006. Published by Oxford University Press. All rights reserved.
MATERIALS AND METHODS

Cell culture, small interfering (siRNA) transfection, cell toxicity measurements

Human cervical carcinoma HeLa cells were cultured in DMEM (Gibco) supplemented with 5% fetal bovine serum. Cells were transfected twice sequentially using Oligofectamine™ and 200 nM siRNA (Qiagen), treated with CPT (500 nM) for the times indicated and collected for the analysis of RNA, DNA or protein. c-Myb siRNA, AAAGGGUGGAUCUCAACUG; Rfx-1 siRNA, AAGAGUGGUGUUUGUGUG; Ctrl siRNA, AAUUCUCUGA-ACGUGCUCAGU (targeting genes expressed in the fungus Ustilago maydis and the bacterium Thermotoga maritima).

Hoechst 33342 (1 µg/ml) was added directly to the cell culture medium and nuclei were visualized and scored 10 min later from >1000 cells; three independent experiments were performed. To monitor [3H]thymidine incorporation, cells were seeded in duplicate 6-well cluster plates (in each experiment, one cluster was used to measure [3H]thymidine incorporation, the other for counting cells and calculating DNA concentration), cultured for 20 h and pulsed with 2 µCi/ml [3H]thymidine at 37°C for 1 h. On ice, cultures were washed once with ice-cold KRB buffer (118 mM NaCl, 25 mM NaHCO3, 5 mM KCl, 1.28 mM CaCl2, 1.18 mM MgCl2 and 1.17 mM K2HPO4) and precipitated by the addition of 2 ml/well ice-cold KRB containing 5% trichloroacetic acid (TCA) for 30 min. Samples were rinsed twice with ice-cold KRB and, following the addition of 1 ml 0.5 M NaOH/0.5% SDS, lysates were collected into scintillation vials. [3H]thymidine incorporation was calculated as c.p.m./106 cells (c.p.m./µg DNA yielded similar results).

Nuclear run-on array analysis

NRO RNA was prepared and analyzed as described elsewhere (3,5) with some modifications. Fifty million cells were lysed in 20 ml Tris–HCl (pH 7.5), 20 mM NaCl, 5 mM MgCl2, 0.25% [v/v] NP-40; the pelleted nuclei were lifted in suspension buffer [20 mM Tris–HCl (pH 7.5), 20 mM NaCl, 5 mM MgCl2], layered onto a sucrose cushion [20 mM Tris–HCl (pH 7.5), 20 mM NaCl, 5 mM MgCl2 and 1 M sucrose], spun at 600 g (4°C, 30 min) and resuspended in storage buffer [50 mM Tris–HCl (pH 8.3), 5 mM MgCl2, 0.1 mM EDTA (pH 8.0) and 45% [v/v] glycerol]. For the NRO reaction, thawed nuclei (200 µl aliquots) were mixed with 200 µl of 2x NRO reaction buffer [10 mM Tris–HCl (pH 8.0), 300 mM KCl, 5 mM MgCl2, 5 mM DTT, 0.5 mM of each rATP, rUTP and rGTP, and 1.2% sarcosyl] plus 500 µCi of [α-35P]UTP (3000 Ci/mmol, 10 mCi/ml) and incubated for 30 min at 30°C with shaking, after which they were digested with DNase I (RNase-free) for 30 min at 37°C and then with proteinase K (1 µg/µl) for 1 h at 37°C. Nascent RNA, purified by filtration using Sephadex G-50 columns (Pharmacia), typically yielded ~3 × 106 c.p.m. MGC arrays (Mammalian Genome Collection, [http://www.ggc.nia.nih.gov/branches/rrb/dna/array.htm], containing 9600 genes (6385 unique) spotted as full-length cDNAs http://mgc.ncbi.nih.gov/), were prehybridized for 2 h in Invitrogen Micro-Hyb™ buffer containing 10 µg Cot DNA and 8 µg poly(A), then the nascent radiolabeled RNA was added and hybridized for 48 h at 55°C. Following washes (2× SSC/0.1% SDS, 2× SSC/0.1% SDS and 1× SSC/0.1% SDS at 55°C), signals on the array filters were detected using a PhosphorImager (Pharmacia) and analyzed using the ArrayPro software (MediaCybernetics, Silver Spring, MD).

For data analysis, the raw intensities were transformed to log10, then used for the calculation of Z-scores, as described elsewhere (6). Significant changes in gene expression were calculated in the form of Z-ratios and Z-test values (7). All of the gene expression changes were assessed through comparison with untreated cells (time 0). The significance thresholds used in this study were Z-ratio values of ±1.50 and Z-test value of P < 0.01.

Computational analysis of promoters

Proximal promoter sequences for the 58 genes upregulated transcriptionally were available from the Promoter database (8). From each gene, 1.2 kb promoter sequences (1.0 kb upstream and 0.2 kb downstream of the transcription start site) were studied. The potential binding sites of TFs in each promoter were detected by scanning against the Transfac Professional database (version 9.2) using the software Match for TF-binding site identification (www.biobase.com). Since TF-binding sites are short in sequence and many are degenerate, random, false positive hits appear frequently. To avoid this problem, we employed a comparative genomics approach based on the notion that bona fide promoters/enhancers exhibit conserved core functional domains and locations. For all of the up-regulated human genes, the mouse homologous counterparts were retrieved from the homologene database of NCBI. In total, 58 mouse homologous genes were obtained and the corresponding proximal promoters searched for TF-binding sites. Using similarity criteria of 0.95 for the core and 0.85 for the matrix (both ranging from 0 to 1), the top 10% most frequent TFs were chosen from human and mouse gene lists and those common between the two species were selected for further analysis.

Chromatin immunoprecipitation (ChIP) assay and quantitative PCR (Q-PCR)

Crosslinking of cells (~5 × 106 per treatment group) was performed in 1% formaldehyde for 10 min at 25°C and was stopped by adding 0.125 M glycine. Cells were washed twice with ice-cold PBS, then with ice-cold buffer I (0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA and 10 mM HEPES, pH 7.0) and buffer II (200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA and 10 mM HEPES, pH 7.0) and were lifted in 300 µl Lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris, pH 8.0 and 1× Roche Protease Inhibitor Cocktail). DNA was sheared to ~500 bp average size fragments using a Fisher Scientific Sonic Dismembrator (FS 100 Model, 60% output, 5 pulses, 5 s each). After sonication, insoluble cell debris was removed by centrifugation (21 000 g, 4°C, 10 min) and supernatants were transferred to fresh 1.5 ml tubes. After assessing DNA concentration, DNA was diluted to 2 U (A260 units)/ml using dilution buffer (0.01% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris–HCl, pH 8.0 and 1× Roche protease inhibitor cocktail). The diluted samples (500 µl) were precleared for 4 h at 4°C with agitation by adding 50 µl of pretreated Protein A/G–Sepharose mixture 50% slurry...
(Amersham). Beads were gently pelleted and the supernatants collected; after setting aside 1/10 vol of precleared lysate as input DNA, the remainder was divided equally into two parts for specific-antibody IP and control IP reactions.

IP reactions were carried out for 12 h at 4°C using 5 μg of specific antibodies (all from Santa Cruz Biotechnology). For IP of c-Myb, a mixture of rabbit anti-human-c-Myb (sc-7874x/H-141) and mouse anti-human-c-Myb (sc-8412x/C-2) antibodies was used; for IP of Rfx1, a mixture of two goat anti-human-Rfx1 antibodies (sc-10650x/D-19 and sc-10652x/I-19) was used; specific control IgGs (from mouse, goat, rabbit) were also from Santa Cruz Biotech.

Following incubation with primary antibodies, 50 μl of pretreated Protein A/G–Sepharose slurry, 10 μg of sheared salmon sperm DNA and 50 μg of BSA were added and incubated for an additional 1 h. Precipitates were washed for 10 min each with TSE I (low-salt buffer, containing 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl, pH 8.1 and 500 mM NaCl) and buffer III (LiCl buffer, containing 0.25 M LiCl, 5% NP-40, 1% deoxycholate, 2 mM EDTA, 1% Triton X-100, 2 mM EDTA, 1 mM Tris–HCl, pH 8.1 and 500 mM NaCl) and buffer III (LiCl buffer, containing 0.25 M LiCl, 5% NP-40, 1% deoxycholate, 2 mM EDTA, 1% Triton X-100, 2 mM EDTA, 1 mM Tris–HCl, pH 8.1). Precipitates were washed twice with TE buffer and the DNA eluted by incubating beads in 300 μl elution buffer (1% SDS, 0.1 M NaHCO3, 0.5 M NaCl) at 65°C for 15 min; after centrifugation at 4500 g (1 min), the eluted DNA was transferred to fresh tubes. Formaldehyde crosslinking was reversed by addition of NaCl (0.3 M final concentration) and 20 μg RNase A (Sigma), then heating at 65°C for 6 h. To purify the DNA fragments, eluates were digested with 50 μg/ml proteinase K (50°C, 2 h), then extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitated with glycerogen (20 μg) as carrier. The resulting DNA, dissolved in 100 μl TE, was used for Q-PCR analysis.

For Q-PCR analysis, all of the fragments were first amplified by regular PCR, purified from agarose gels, serially diluted (10^-10^-10) and used in Q-PCRs to prepare standard curves from which target gene fragment numbers were calculated in both IP DNA and ‘Input’ DNA (for normalization). Q-PCRs were performed using SYBR® Green, the oligomers listed below and the MJ Research Chrom4 thermal Cycler System (MJ Research Inc., Waltham, MA). Quality-control tests for the Q-PCR products were routinely performed by monitoring melting curves and the amplification of single DNA bands (Supplementary Data).

Oligomers for Q-PCR analysis after ChIP (for detailed promoter sequence information see Supplementary Data) FANCG, (for both c-Myb and Rfx1 IP) CCGGCTCTGC-GAAGCTCTGGGCT (forward) and GTTGTGGCCAGCGAG- GAAGGCGC (reverse), yielding a 258 bp fragment; PANX1, (for both c-Myb and Rfx1 IP) CAAGGCTCTGGATTGGaat-GGCAG and GCAAGCGACTTCATGGAAG, yielding a 293 bp fragment; POLD2, (for c-Myb IP) CAAGGCTCTGGATTGGGAATGATC and GCTGCTGTATCGGA- AAGATGG, yielding a 276 bp fragment. For the detection of c-Myb, Rfx1, cleaved PARP, GAPDH and β-tubulin, whole-cell lysates were prepared in RIPA buffer and 10 μg aliquots were size-fractionated by SDS–PAGE for western blot analysis. Primary antibodies recognized c-Myb (Upstate Biotechnology), Rfx1 (Santa Cruz Biotechnology), GAPDH (Abcam), cleaved PARP (Cell Signaling Technology) or β-tubulin (Santa Cruz Biotechnology). Following secondary antibody incubations, signals were detected by enhanced chemiluminescence.

RESULTS AND DISCUSSION

En masse identification of newly transcribed genes by NRO and cDNA arrays

We present an approach (Figure 1) to identify specific TFs mediating the transcriptional control of gene expression based on the systematic analysis of de novo (nascent) transcription. The assessment of newly synthesized mRNA (using NRO) was chosen in order to circumvent the problematic influence of mRNA turnover on the pools of expressed mRNAs, since NRO analysis would allow the direct identification of bona fide transcriptionally regulated genes. The subsequent comparison of the corresponding promoter regions of these genes would then be used to elucidate the putative TFs involved. To test the validity of this approach, HeLa (human cervical carcinoma) cells were treated with the topoisomerase I inhibitor CPT [reviewed in Ref. (9)], a drug that elicited rapid, robust and consistent changes in gene transcription. The ensuing
transcriptional changes were systematically assessed by isolating nuclei from each cell population and carrying out NRO analysis, whereby individual RNA polymerase II molecules were allowed to resume on-going synthesis of endogenous transcripts in the presence of $[^{33}P]$UTP. The resulting radiolabeled nascent RNA products were then studied through hybridization of cDNA arrays, using methodologies described previously (3,6,7,10) and in Materials and Methods. For most genes, transcription remained unaltered following CPT treatment, as determined by monitoring NRO signals on the arrays, but a distinct subset of genes was transcriptionally upregulated and another subset was transcriptionally downregulated after CPT treatment (Figure 2). Control hybridizations conducted using samples from cells that had been treated with actinomycin D (an inhibitor of RNA polymerase II) indicated that the signals on the arrays were indeed derived from RNA polymerase II activity (Supplementary Data). Among the 85 genes that were transcriptionally upregulated, the proximal promoters (arbitrarily set at 1.0 kb upstream, 0.2 kb downstream of the transcription initiation site) were available for 58 genes and were thus chosen for further study.

**Shared TF-binding sites in the promoters of transcriptionally upregulated genes**

Table 1 lists the genes whose nascent transcription was upregulated by CPT; $Z$-ratio >1.5 [empirically found to correspond to >3-fold differences in signal intensity (data not shown)] and $P < 0.01$ values were chosen. There was some overlap between these genes and those identified as CPT-regulated in other high-throughput studies [including CDKN1A and DDB2 (11,12)], despite differences in cell lines employed, treatment conditions, array platforms and parameters measured (nascent RNA versus steady-state mRNA). The promoters corresponding to the genes in Table 1 were then analyzed using the PromoSer database (8). By considering similarity scores for TF-binding sites in these 58 promoters (with similarity values $\geq 0.95$ and $\geq 0.85$ for the core and matrix of each TF site, respectively), the presence (+) or absence (−) of binding sites on each promoter region for each TF were examined. Several TF sites that were either
Table 1. Genes transcriptionally upregulated by treatment with CPT

| Symbol  | Z-ratio | P-value | RefSeq  |
|---------|---------|---------|---------|
| ABCF2   | 1.55    | 0.0001  | NM_007189 |
| AMPD2   | 1.75    | 0.0000  | NM_004037 |
| AP2M1   | 1.51    | 0.0006  | NM_004068 |
| APM2C1  | 2.17    | 0.0000  | NM_021203 |
| ASC     | 4.25    | 0.0000  | NM_013258 |
| BC-2    | 1.51    | 0.0008  | NM_014453 |
| CDA4    | 1.72    | 0.0007  | NM_017955 |
| CDN1A   | 4.68    | 0.0000  | NM_078467 |
| CDKN2D  | 1.58    | 0.0000  | NM_001800 |
| CHRNA3  | 1.59    | 0.0001  | NM_007043 |
| CIZ1    | 1.56    | 0.0012  | NM_012127 |
| CTX4    | 1.77    | 0.0032  | NM_014312 |
| CDCA4   | 2.33    | 0.0000  | NM_014003 |
| DGLC8   | 1.73    | 0.0000  | NM_022720 |
| EGR-RS  | 1.55    | 0.0000  | NM_022450 |
| EIF3S4  | 3.11    | 0.0000  | NM_003755 |
| FANC3   | 1.81    | 0.0001  | NM_004629 |
| FCR2    | 3.02    | 0.0000  | NM_002002 |
| GBA     | 1.77    | 0.0000  | NM_000157 |
| GEMIN4  | 2.01    | 0.0000  | NM_015721 |
| GPA1    | 1.94    | 0.0045  | NM_003801 |
| GPRC1   | 2.55    | 0.0000  | NM_014373 |
| GPR30   | 1.54    | 0.0000  | NM_015105 |
| GRN     | 1.71    | 0.0000  | NM_020687 |
| HARS    | 1.60    | 0.0004  | NM_002109 |
| HS179303| 1.60    | 0.0001  | NM_013301 |
| HYAL3   | 1.66    | 0.0001  | NM_003549 |
| ISG20   | 2.45    | 0.0000  | NM_022010 |
| LOC51193| 1.59    | 0.0000  | NM_016331 |
| MADH5   | 1.62    | 0.0000  | NM_005903 |
| MARS    | 1.52    | 0.0000  | NM_004990 |
| MAPT    | 3.50    | 0.0000  | NM_016180 |
| MRPL10  | 1.77    | 0.0106  | NM_148887 |
| NCK1    | 2.33    | 0.0000  | NM_006153 |
| NEUGRIN | 1.83    | 0.0000  | NM_016645 |
| NIC1    | 1.85    | 0.0000  | NM_023216 |
| NPPC1   | 2.00    | 0.0000  | NM_026273 |
| Nup37   | 1.65    | 0.0003  | NM_024057 |
| OGG1    | 1.83    | 0.0002  | NM_016819 |
| PANX1   | 1.93    | 0.0000  | NM_015368 |
| PB1     | 2.39    | 0.0000  | NM_018165 |
| POLD2   | 1.67    | 0.0009  | NM_006230 |
| POLG2   | 1.39    | 0.0000  | NM_007215 |
| POP1CA  | 1.95    | 0.0001  | NM_002078 |
| PVR     | 2.94    | 0.0000  | NM_006505 |
| RNAHP   | 3.69    | 0.0000  | NM_007372 |
| RPS5    | 1.73    | 0.0000  | NM_001009 |
| RRM2    | 1.67    | 0.0004  | NM_001034 |
| SH3GLB2 | 1.59    | 0.0000  | NM_020145 |
| SNX5    | 1.67    | 0.0000  | NM_015227 |
| SRP     | 3.49    | 0.0000  | NM_033199 |
| STK16   | 1.55    | 0.0014  | NM_003691 |
| TAPBP-R | 2.77    | 0.0000  | NM_018009 |
| TCEA2   | 1.68    | 0.0087  | NM_003195 |
| TPIP1   | 1.85    | 0.0048  | NM_012143 |
| TNFRSF7 | 3.96    | 0.0000  | NM_012142 |
| UQCR    | 1.51    | 0.0000  | NM_006830 |
| VAMP3   | 1.99    | 0.0000  | NM_004781 |

Following treatment of HeLa cells with CPT for 2 h at 500 nM, nuclei were isolated for NRO analysis to detect nascent transcription using cDNA arrays (Materials and Methods). Promoter sequences were available for 58 transcriptionally upregulated genes (those genes showing Z-ratios >1.5 when comparing CPT-treated with untreated groups, and P-values <0.01); the analysis of these promoters revealed the presence of common binding sites for various TFs present at high, medium and low abundance. A TF-binding site was positive if the promoter sequence matched the consensus sequence with score values >0.95 for the core and >0.85 for the matrix of each TF. Plus, one or several TF-binding sites predicted; minus, no TF-binding sites predicted.
extensively shared (e.g. c-Myb through Lmo2), moderately shared (e.g. NF-κB, c-Rel, USF) or minimally shared (e.g. AHR, Pax-3, E2F) among the corresponding promoters are indicated. The potential binding sites of TFs in these promoters were also identified in the mouse orthologous genes (the complete lists for both species are available from the authors), showing extensive sharing of many TFs for these 58 genes, suggesting their evolutionary conservation and their possible co-regulatory function upon CPT-induced genes.

To test whether this approach could adequately identify shared regulatory TFs, we further assessed the influence of TFs Rfx1 and c-Myb [described in Refs (13–15)], for which shared regulatory TFs, we further assessed the influence of possible co-regulatory function upon CPT-induced genes. Further verification of this regulatory scheme was undertaken by ChIP followed by real-time Q-PCR analysis to examine the transcriptional increase after CPT treatment was unchanged despite the silencing of c-Myb or Rfx1, suggesting that their transcriptional increase after CPT treatment was regulated by other TFs. Taken together, these findings suggest that c-Myb, Rfx1 or both proteins were required for maximal transcriptional induction of a sizeable group of genes (19 out of 58) which had the corresponding TF-binding sites in their promoters (Table 1).

The identified TFs bind to the promoters of putative target genes and regulate mRNA abundance

Further verification of this regulatory scheme was undertaken by ChIP followed by real-time Q-PCR analysis to examine whether c-Myb, Rfx1 or both proteins physically associated with the promoters of target genes. Cells were crosslinked with formaldehyde to preserve the association of proteins with target DNA sequences, whereupon nascent transcripts were prepared by NRO and assessed by array analysis (Materials and Methods). Shown are the Z-ratios of the nascent transcript signals (CPT-treated versus untreated) in each siRNA group. The genes on the graph have significant Z-ratios in the Ctrl. siRNA transfection group (>1.5) and lower Z-ratios in either the c-Myb siRNA transfection group, the Rfx1 siRNA transfection group or both.

The identified TFs bind to the promoters of putative target genes and regulate mRNA abundance

Further verification of this regulatory scheme was undertaken by ChIP followed by real-time Q-PCR analysis to examine whether c-Myb, Rfx1 or both proteins physically associated with the promoters of target genes. Cells were crosslinked with formaldehyde to preserve the association of proteins with target DNA sequences, whereupon nascent transcripts were prepared by NRO and assessed by array analysis (Materials and Methods). Shown are the Z-ratios of the nascent transcript signals (CPT-treated versus untreated) in each siRNA group. The genes on the graph have significant Z-ratios in the Ctrl. siRNA transfection group (>1.5) and lower Z-ratios in either the c-Myb siRNA transfection group, the Rfx1 siRNA transfection group or both.
transcripts can hybridize to the double-stranded cDNAs spotted on the arrays, while oligomer arrays might theoretically be used to circumvent this problem, it is technically impossible to achieve the sensitivity and specificity required with the tools presently available. These issues must be taken into consideration when interpreting NRO array data, since all of them have thus far been generated using cDNA arrays (3,5,16–18).

Taken together, these findings reveal that the analysis of nascent transcripts uniquely permits the identification of TFs which coordinate the expression of subsets of genes. Rfx1 and c-Myb, elucidated by this approach, were found to play critical roles in the cellular response to the genotoxic agent CPT. They were shown to associate with the promoters of the predicted target genes, binding was influenced by CPT treatment and knockdown of Rfx1 or c-Myb suppressed the CPT-induced expression of target genes. The varying degree of inhibition for the transcripts tested (Figure 4) likely depends on the influence of additional TFs independently acting upon the promoters and also on the possible effects of c-Myb and/or Rfx1-interacting proteins (19–21). It is important to note that the NRO methodology has significant limitations, including the facts that NRO data are semiquantitative and can have low sensitivity, specificity or both (whether derived from arrays or from traditional spotted nucleic acids). In addition, endogenous antisense

![Figure 5. Quantitative ChIP analysis. ChIP analysis was performed essentially as described elsewhere (36). Following transfection (with Ctrl., c-Myb, Rfx1 siRNAs) cells were treated with CPT (2 h, 500 nM) and collected for analysis. Immunoprecipitations were carried out using anti-c-Myb, anti-Rfx1, or control IgG, whereupon Q-PCR was performed to amplify FANCG, PANX1, POLD2, TFIP11 or VAMP3 promoter regions (details in Materials and Methods). The absolute number of molecules amplified was calculated by carrying out parallel amplification curves with known substrate input quantities. Data are shown as the means from three independent biological triplicates and the SEMs.](image)

![Figure 6. Quantitative mRNA analysis. Total RNA was extracted from the indicated siRNA groups at the times indicated following CPT treatment and mRNA levels calculated by Q-PCR using sequence-specific oligomer pairs and normalized to 18S rRNA levels (details in Materials and Methods). The absolute number of molecules amplified was calculated by carrying out parallel parallel amplification curves with known substrate input quantities. Data are shown as the means from three independent biological triplicates. Following c-Myb or Rfx1 silencing, the basal levels of these genes were either unchanged or modestly altered (up or down by 2-fold or less): FANCG mRNA was unchanged by 2-fold (both siRNA interventions), PANX1 mRNA was slightly down and 2-fold up (after silencing c-Myb and Rfx1, respectively), POLD2 was unchanged and 2-fold down (after silencing c-Myb and Rfx1, respectively), TFIP11 mRNA was slightly down and 2-fold up (after silencing c-Myb and Rfx1, respectively), and VAMP3 was slightly elevated after silencing Rfx1.](image)
Essential roles of c-Myb and Rfx1 in the cellular response to CPT

Finally, the influence of c-Myb and Rfx1 on the overall cellular response to CPT was evaluated by monitoring changes in cellular division, DNA replication and apoptosis of each population. As shown, c-Myb appeared to suppress proliferation slightly, since knocking down c-Myb resulted in increased cell numbers; however, silencing either c-Myb or Rfx1 significantly reduced cell numbers following CPT treatment (Figure 7A). Similarly, populations in which c-Myb or Rfx1 were knocked down incorporated significantly less [\(^{3}\)H]thymidine, suggesting that c-Myb and Rfx1 contributed to maintaining elevated levels of DNA replication in response to CPT treatment (Figure 7B). Moreover, apoptosis increased significantly after Rfx1 knockdown, even in cultures that were left without additional treatment (U populations) and was further elevated after CPT treatment in both c-Myb and Rfx1 siRNA groups, as monitored both by counting apoptotic (condensed or fragmented) nuclei and by assessing the levels of cleaved PARP (Figure 7C). Together, these findings suggest that Rfx1 and c-Myb exert a growth-regulatory, anti-apoptotic influence on HeLa cells exposed to CPT. The finding that c-Myb and Rfx1 are required for the optimal survival of CPT-treated HeLa cells is in keeping with the documented proto-oncogenic effects of c-Myb (24–27) and Rfx1 (28–31).

Perspectives

Most of our current understanding of the dynamics of expressed mRNA in mammalian systems comes from studies looking at steady-state mRNA abundance. However, a heightened interest in elucidating the TFs that regulate gene expression is driving the development of genome-wide methodologies to directly and systematically assess gene transcription and TF function. Indeed, the past few years have seen the development of powerful array-based and other high-throughput methods for ChIP and gene transcription analyses (16–18,32–35). The high-throughput, ChIP-derived strategies have the advantage of identifying all bound DNA at a genomic level, including previously unknown target regulatory regions, but have two significant limitations: they require prior knowledge of the involvement of a specific TF in a given response, and they provide little information on the functional consequences of the TF acting upon the particular genes (i.e. whether transcription is elevated, reduced or unaffected). The NRO array methodologies provide genome-wide information on transcriptional changes and do not require prior knowledge of the TFs involved, but fail to identify the factors that orchestrate the altered gene transcription programs. Here, we have reported a strategy which employs elements from both of these approaches. By obtaining en masse transcription data as a starting point, and coupling it to the analysis of the corresponding promoters, the TFs that drive transcriptional changes in response to a given stimulus can be systematically identified. We propose that the approaches described here can be widely applied to the identification of other TFs that orchestrate adaptive gene expression programs and shape the ensuing cellular responses.

Essential roles of c-Myb and Rfx1 in the cellular response to CPT

Finally, the influence of c-Myb and Rfx1 on the overall cellular response to CPT was evaluated by monitoring changes in cellular division, DNA replication and apoptosis of each population. As shown, c-Myb appeared to suppress proliferation slightly, since knocking down c-Myb resulted in increased cell numbers; however, silencing either c-Myb or Rfx1 significantly reduced cell numbers following CPT treatment (Figure 7A). Similarly, populations in which c-Myb or Rfx1 were knocked down incorporated significantly less [\(^{3}\)H]thymidine, suggesting that c-Myb and Rfx1 contributed to maintaining elevated levels of DNA replication in response to CPT treatment (Figure 7B). Moreover, apoptosis increased significantly after Rfx1 knockdown, even in cultures that were left without additional treatment (U populations) and was further elevated after CPT treatment in both c-Myb and Rfx1 siRNA groups, as monitored both by counting apoptotic (condensed or fragmented) nuclei and by assessing the levels of cleaved PARP (Figure 7C). Together, these findings suggest that Rfx1 and c-Myb exert a growth-regulatory, anti-apoptotic influence on HeLa cells exposed to CPT. The finding that c-Myb and Rfx1 are required for the optimal survival of CPT-treated HeLa cells is in keeping with the documented proto-oncogenic effects of c-Myb (24–27) and Rfx1 (28–31).

Perspectives

Most of our current understanding of the dynamics of expressed mRNA in mammalian systems comes from studies looking at steady-state mRNA abundance. However, a heightened interest in elucidating the TFs that regulate gene expression is driving the development of genome-wide methodologies to directly and systematically assess gene transcription and TF function. Indeed, the past few years have seen the development of powerful array-based and other high-throughput methods for ChIP and gene transcription analyses (16–18,32–35). The high-throughput, ChIP-derived strategies have the advantage of identifying all bound DNA at a genomic level, including previously unknown target regulatory regions, but have two significant limitations: they require prior knowledge of the involvement of a specific TF in a given response, and they provide little information on the functional consequences of the TF acting upon the particular genes (i.e. whether transcription is elevated, reduced or unaffected). The NRO array methodologies provide genome-wide information on transcriptional changes and do not require prior knowledge of the TFs involved, but fail to identify the factors that orchestrate the altered gene transcription programs. Here, we have reported a strategy which employs elements from both of these approaches. By obtaining en masse transcription data as a starting point, and coupling it to the analysis of the corresponding promoters, the TFs that drive transcriptional changes in response to a given stimulus can be systematically identified. We propose that the approaches described here can be widely applied to the identification of other TFs that orchestrate adaptive gene expression programs and shape the ensuing cellular responses.

ACKNOWLEDGEMENTS

We are grateful to our colleagues A. Jani, M. Fann, K. G. Becker, C. Cheadle and W. H. Wood, III for reagents, advice and assistance throughout this investigation. This work was supported in part by the Intramural Research Program of...
the National Institute on Aging, National Institutes of Health. Funding to pay the Open Access publication charges for this article was provided by the National Institute on Aging-Intramural Research Program, NIH.

Conflict of interest statement. None declared.

REFERENCES

1. Chen,L. (1999) Combinatorial gene regulation by eukaryotic transcription factors. *Curr. Opin. Struct. Biol.*, 9, 48–55.

2. Ogata,K., Sato,K. and Tahirov,T.H. (2003) Eukaryotic transcriptional regulatory complexes: cooperativity from near and afar. *Curr. Opin. Struct. Biol.*, 13, 40–48.

3. Fan,J., Yang,X., Wang,W., Wood,W.H.III, Becker,K.G. and Gorospe,M. (2002) Global analysis of stress-regulated mRNA turnover by using cDNA arrays. *Proc. Natl Acad. Sci. USA.*, 99, 10611–10616.

4. Khodursky,A.B. and Bernstein,J.A. (2003) Life after transcription—revisiting the fate of messenger RNA. *Trends Genet.*, 19, 113–115.

5. Cheadle,C., Fan,J., Cho-Chung,Y.S., Werner,T., Ray,J., Do,L., Gorospe,M. and Becker,K.G. (2005) Control of gene expression during T cell activation: alternate regulation of mRNA transcription and mRNA stability. *BMCC Genomics*, 6, 75.

6. Cheadle,C., Vawter,M.P., Freed,W.J. and Becker,K.G. (2003) Analysis of microarray data using Z score transformation. *J. Mol. Diagn.*, 5, 73–81.

7. Nadon,R., Woody,E., Shi,P., Rghei,N., Hubscble,H., Susko,E. and Ramm,P. (2002) Statistical inference in array genomics. In Geschwind,D. and Gregg,J. (eds), *Microarrays for the Neurosciences*. Cambridge, MIT Press.

8. Halees,A.S., Leyfer,D. and Weng,Z. (2003) PromoSer: a large-scale mammalian promoter and transcription start site identification service. *Nucleic. Acids. Res.*, 31, 3554–3559.

9. Thomas,C.J., Rahier,N.J. and Hecht,S.M. (2004) Camptothecin: current perspectives. *Bioorg. Med. Chem.*, 12, 1585–1604.

10. Kawai,T., Fan,J., Mazan-Mamczarz,K. and Gorospe,M. (2004) Global mRNA stabilization preferentially linked to translational repression during the endoplasmic reticulum stress response. *Mol. Cell. Biol.*, 24, 6773–6787.

11. Zhou,Y., Gwadry,F.G., Reinhold,W.C., Miller,L.D., Smith,L.H., Oasland,R., White-Cooper,H., Dyson,N. and Brehm,A. (2004) Native ribonomic profiling with autoimmune sera. *Genome-wide regulatory analysis using en masse nuclear run-ons and immunoprecipitation with autoimmune sera. Gene*, 317, 79–87.

12. Afroze,T. and Husain,M. (2000) c-Myb-binding sites mediate G(1)/S-associated repression of the plasma membrane Ca(2+)-ATPase 1 promoter. *J. Biol. Chem.*, 275, 9062–9069.

13. Sengupta,P.K., Fardo,J. and Smith,B.D. (2002) The RFX family interacts at the collagen (COL1A2) start site and represses transcription. *J. Biol. Chem.*, 277, 24926–24937.

14. Fan,J., Durand,B., Emery,P., David,E., Hearing,G., Mach,B. and Reith,W. (1993) RFX1 is identical to enhancer factor C and functions as a transactivator of the hepatitis B virus enhancer. *Cell.*, 76, 21–60.

15. Gonda,T.J. (1998) The c-Myb oncoprotein. *Cancer Res.*, 58, 473–515.

16. Introna,M. and Golay,J. (1999) How can oncogenic transcription factors cause cancer: a critical review of the myb story. *Leukemia*, 13, 1301–1306.

17. Weston,K. (1999) Reassessing the role of C-MYB in tumorigenesis. *Oncogene*, 18, 3034–3038.

18. Ostapchuk,P., Scheirle,G. and Hearing,G. (1989) Binding of nuclear factor EF-C to a functional domain of the hepatitis B virus enhancer region. *J. Biol. Chem.*, 264, 2787–2797.

19. Sengupta,P.K., Fardo,J. and Smith,B.D. (2002) The RFX family interacts at the collagen (COL1A2) start site and represses transcription. *J. Biol. Chem.*, 277, 24926–24937.

20. Gajiwala,K.S., Chen,H., Cornille,F., Roques,B.P., Reith,W., Mach,B. and Burke,S.K. (2000) Structure of the winged-helix protein RFX1 reveals a new mode of DNA binding. *Nature*, 403, 916–921.

21. Gonda,T.J., Durand,B., Emery,P., David,E., Hearing,G., Mach,B. and Reith,W. (1993) RFX1 is identical to enhancer factor C and functions as a transactivator of the hepatitis B virus enhancer. *Cell.*, 76, 21–60.

22. Labrie,C., Lee,B.H. and Mathews,M.B. (1995) Transcription factors RFX/EF-C and ATF-1 associate with the adenovirus E1A-responsive element of the human proliferating cell nuclear antigen promoter. *Nucleic Acids Res.*, 23, 3732–3741.

23. Maijgren,S., Sur,J., Nilsson,M. and Toftgard,R. (2001) The transcriptional program of a human B cell line in response to TCR stimulation. *Cancer Res.*, 61, 9062–9069.

24. Afridi,T. and Husain,M. (2000) c-Myb-binding sites mediate G(1)/S-associated repression of the plasma membrane Ca(2+)-ATPase 1 promoter. *J. Biol. Chem.*, 275, 9062–9069.