Diverse stresses dramatically alter genome-wide p53 binding and transactivation landscape in human cancer cells

Daniel Menendez[1], Thuy-Ai Nguyen[1], Johannes M. Freudenberg[2], Viju J. Mathew[2,3], Carl W. Anderson[1,4], Raja Jothi[2,*] and Michael A. Resnick[1,*]

[1]Chromosome Stability Group, Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), Research Triangle Park, NC 27709, USA, [2]Systems Biology Group, Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), Research Triangle Park, NC 27709, USA, [3]William G. Enloe High School, Raleigh, NC 27610, USA and [4]Department of Biology, Brookhaven National Laboratory, Upton, NY 11973, USA

ABSTRACT

The effects of diverse stresses on promoter selectivity and transcription regulation by the tumor suppressor p53 are poorly understood. We have taken a comprehensive approach to characterizing the human p53 network that includes p53 levels, binding, expression and chromatin changes under diverse stresses. Human osteosarcoma U2OS cells treated with anti-cancer drugs Doxorubicin (DXR) or Nutlin-3 (Nutlin) led to strikingly different p53 gene binding patterns based on chromatin immunoprecipitation with high-throughput sequencing experiments. Although two contiguous RRRCWWGYYY decamers is the consensus binding motif, p53 can bind a single decamer and function *in vivo*. Although the number of sites bound by p53 was six times greater for Nutlin than DXR, expression changes induced by Nutlin were much less dramatic compared with DXR. Unexpectedly, the solvent dimethylsulphoxide (DMSO) alone induced p53 binding to many sites common to DXR; however, this binding had no effect on target gene expression. Together, these data imply a two-stage mechanism for p53 transactivation where p53 binding only constitutes the first stage. Furthermore, both p53 binding and transactivation were associated with increased active histone modification histone H3 lysine 4 trimethylation. We discovered 149 putative new p53 target genes including several that are relevant to tumor suppression, revealing potential new targets for cancer therapy and expanding our understanding of the p53 regulatory network.

INTRODUCTION

Gene transcription is modulated by dynamic interactions between cis-regulatory elements and regulatory proteins that include DNA-binding transcription factors (TFs). Genome-wide identification of cis-elements *in vivo* is essential for a better understanding of the relationship between DNA-binding proteins and global gene expression programs. We have used the combination of chromatin immunoprecipitation (ChIP) with high-throughput sequencing (ChIP-seq) and functional analysis to characterize the 'cistrome' of the tumor suppressor p53 under various stress conditions.

In response to a variety of cellular stresses, the p53 TF binds to target sequences to regulate expression of hundreds of genes that can affect a variety of cellular processes including growth arrest, senescence and apoptosis (1–4). Widely known as the 'guardian of the genome', it is mutated in over half of all cancers, with a majority of the mutations in the DNA-binding domain of p53. Following cellular stress, stabilized p53 translocates into the nucleus, where it has been proposed, based largely on *in vitro* data, to bind as a dimer of dimers to the consensus sequence motif $RRRCWWGYYY(N_{0-13})RRRCWWGYYY$ (where R is a purine, W is an adenine or thymine, Y is a pyrimidine and N is any base) (5,6). Our

recent studies demonstrated that p53 also can bind non-canonical response elements (REs) (including half and three-quarter consensus sequences) to mediate transactivation of several genes involved in angiogenesis, DNA repair and innate immune responses (7–9).

Many of the factors expected to influence p53-induced changes in gene expression are poorly understood. Included are p53 activation, specificity of p53 binding to various targets, relationship between binding and transactivation and the impact of different stresses that can induce p53. Various genome-wide approaches including ChIP-seq have been used to map sites of actual p53 binding among the hundreds of thousands of potential sites in the human genome and to identify candidate p53 target genes (10–14). Although the identification of non-canonical REs is expected to increase the universe of p53-transcribed regions in the human genome, the overall importance of non-canonical sites in the p53 regulatory network remains unclear. As p53 plays a key role in maintaining genome integrity and cellular homeostasis, genome-wide studies to identify all of its binding sites and/or gene targets would provide a better understanding of its transcriptional regulation functions in various biological processes.

Here, we take a comprehensive approach to characterizing the human p53 network that takes into account p53 levels, binding, expression and chromatin changes under diverse stress conditions. U2OS osteosarcoma cancer cells, expressing wild-type p53, had strikingly different p53-binding patterns and transcriptional responses following exposure to the DNA-damaging agent Doxorubicin (DXR) versus the small molecule Nutlin-3 (Nutlin), considered a non-genotoxic activator of p53. Although two RRRCWWGYYY decamers separated by a zero spacer is the consensus p53-binding motif, we also found that a single decamer (half-site) is sufficient for binding and function *in vivo*. Surprisingly, the vehicle control dimethylsulfoxide (DMSO) led to binding of many sites without changes in gene expression despite changes in chromatin modifications, suggesting an intermediate state that separates site-specific binding from transcriptional activation. In addition, using appropriate genetic controls, we have validated new p53 target genes involved in key biological processes including nervous system development, lipid metabolism and DNA repair. Overall, we identified 149 new putative direct p53 target genes including several that have recurrent tumor-associated mutations.

MATERIALS AND METHODS

Cell culture

Human cancer cell lines U2OS (HTB-96), A549 and H1299 were obtained from ATCC (Manassas, VA). Human colon carcinoma HCT116 $p53^{+/+}$ and its isogenic derivative HCT116 $p53^{-/-}$ were provided by B. Vogelstein (The Johns Hopkins Kimmel Cancer Center, Baltimore, MD). Human breast adenocarcinoma cells MCF7 stably expressing shRNA to p53 from the pSUPER vector, designated as 'MCF7-p53i', or only carrying pSUPER vector as a control ('MCF7-vector') were kindly provided by R. Agami (The Netherlands Cancer Institute, Amsterdam). All cells were cultured in McCoy's 5A or Dulbecco's modified Eagle's medium medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin (Invitrogen, Carlsbad, CA) and maintained at 37°C with 5% CO_2.

Chemicals

DMSO, DXR hydrochloride and Nutlin were purchased from Sigma-Aldrich (St Louis, MO). DXR was dissolved in water and used at 0.6 μg/ml, whereas Nutlin was dissolved in DMSO and used at 10 μM. DMSO was used at 0.1%.

Immunoblot analysis

Whole-cell lysates were prepared by lysing cells in radioimmunoprecipitation assay (RIPA) lysis buffer (Thermo Scientific Pierce, Rockford, IL) supplemented with 1× Halt Protease Inhibitor Cocktail. Nuclear protein lysates were obtained using the nuclear and cytoplasmic extraction reagent (NE-PER) protein extraction kit (Thermo Scientific Pierce) following the manufacturer's recommendations. Equal protein amounts were separated on Bis-Tris NuPAGE (4–12%) gels, transferred to polyvinylidene difluoride membranes (Invitrogen) and incubated with primary antibodies against the following proteins: p53 (DO-1), actin (C-11) and lamin B (C-20) from Santa Cruz Biotechnology (Santa Cruz, CA); Hsp90 (S88) antibody from Abcam (Cambridge, MA). Horseradish peroxidase linked goat anti-mouse, goat anti-rabbit and donkey anti-goat (Santa Cruz Biotechnology) were used as secondary antibodies. Protein was detected using SuperSignal Chemiluminescent Substrate (Thermo Scientific Pierce).

Cell cycle analysis

For cell cycle analysis, cells were trypsinized, washed with cold phosphate-buffered saline and fixed in 70% ethanol at 4°C overnight. In all, 1×10^6 cells were stained with propidium iodide (50 μg/ml) supplemented with RNase A (100 μg/ml) at 37°C in the dark for 1 h. Flow cytometry analysis was performed on FACS Scan Flow Cytometer (Becton Dickinson). Data were analyzed using CellQuest and Modfit softwares (Becton Dickinson, San Jose, CA).

ChIP and sequencing

Cells were grown to 80% confluency and treated with 0.6 μg/ml DXR, 10 μM Nutlin or 0.1% DMSO for 24 h or mock treated. Sixty million cells were used per each ChIP-seq experiment. Proteins and DNA were cross-linked with 1% formaldehyde for 10 min. Cross-linking was stopped by adding 125 mM glycine. Cells were scraped from plates in Lysis buffer [5 mM PIPES (pH 8.0), 85 mM KCl, 0.5% NP-40 supplemented with protease inhibitors (Roche, Nutley NJ)], then lysed with ChIP RIPA Buffer (1× PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS supplemented with protease inhibitors), and nuclei were collected. Chromatin was

sheared to 100–300 bp using a Bioruptor sonicator (Diagenode, Denville, NJ) for 30 min at high power, 30 s ON, 30 s OFF. p53 (DO-1) antibody (12 μg) or normal mouse immunoglobulin G (IgG) (3 μg) (Santa Cruz) was bound to 200 μl Dynabeads Protein G magnetic beads (Invitrogen) by incubation at 4°C in 5 mg/ml of bovine serum albumin for 6 h. Five percent of the chromatin was saved as input. Forty million cells were used for the p53 IP, and 15 million cells were used for the IgG IP. Chromatin was added to the antibody-coupled beads, and DNA–protein complexes were immunoprecipitated overnight at 4°C. Antibody chromatin complexes were washed five times with LiCl IP Wash Buffer [100 mM Tris (pH 7.5), 500 mM LiCl, 1% NP-40, 1% sodium deoxycholate] and eluted in 1% SDS, 0.1 M NaHCO$_3$. Cross-links in the eluted protein–DNA complex and the Input DNA were reversed by incubation at 65°C overnight, and DNA was purified with phenol–chloroform extraction followed by a PCR clean-up column (QIAGEN, Valencia, CA). DNA was quantified using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE). Libraries were prepared using ChIP-seq DNA Sample Prep kit as per manufacturer's instructions and were sequenced on Illumina GAII to yield 36-bp reads (Illumina, Inc., San Diego, CA). For histone modification analyses, 15 million cells were used for IP with histone H3 lysine 4 trimethylation (H3K4me3), histone H3 lysine 27 trimethylation (H3K27me3) (EMD Millipore, Billerica, MA) or normal rabbit IgG (Santa Cruz), and the resulting DNA fragments were quantified by quantitative real time PCR (qPCR).

ChIP-seq analysis

Sequences were aligned to the reference genome (human NCBI36/hg18 assembly) using Bowtie 0.12.2 (15) with default settings, aligning all 36 bases and allowing for two mismatches. Only reads that mapped to unique genomic locations were retained for analysis, and p53-binding sites were determined using the Site Identification from Short Sequence Reads (SISSRs) peak finder tool (16) (default settings) with sequenced reads from Input DNA used as a control ($P < 0.001$). Genes were defined using Entrez gene IDs in the UCSC Genome Browser's known Genes table (human NCBI36/hg18 assembly). Genes were considered bound by p53 if there was a p53-binding site within a region starting 5 kb upstream of that gene's transcription start site (TSS) and covering the gene body (i.e. end of the transcribed region).

To determine 'sites' (i.e. sites identified using the SISSRs) common between treatments, we used the following procedure: First, sites shorter than 200 bp were extended both up- and downstream from the site's center by 100 bp. A site in one treatment is considered to be common in another treatment if the corresponding sites for the two treatments overlapped by 1 bp. ChIP-over-Input fold enrichment was determined by SISSRs (16), and TSS, exons and introns were used as defined in the known genes table (human NCBI36/hg18 assembly).

Motif search

Sequences from the reference genome (human NCBI36/hg18 assembly) were extracted and extended to have at least 200 bp length where applicable as described earlier in the text. A *de novo* motif search was performed using MEME (17) with zero or one motif occurrence per sequence, and E-value threshold 0.1. Owing to the computational complexity, the search was limited to the top 1000 binding site sequences with the highest ChIP-over-Input fold enrichment. Consensus motif sequences were extracted from MEME results and visualized. Search for known motifs was performed using the Motif Alignment and Search Tool (18) with default options. Position weight matrices (PWMs) were used directly from MEME output for 'spacer search' (inserting spacers with 0 weight for each base, i.e. each base was assumed to be equally likely). PWMs were computed from frequency matrices from the Transfac database (19), by computing the log odds matrix, given base frequencies for a given position in the motif.

ChIP-qPCR

A total of 42 binding peaks were selected for validation using qPCR. The primers used are listed in Supplementary Table S4. Two biological replicates were performed for each experiment. Real-time PCR and melting curve analysis was performed using the SYBR® Green (Invitrogen) dye detection method on the ABI PRISM 7900 HT Sequence Detection System under default conditions. The comparative Ct method was used for quantification, and enrichment of the ChIP material was calculated as recovery over an unspecific control glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) for p53 and H3K27me3 ChIP; myogenic differentiation 1 (*MyoD1*) for H3K4me3 ChIP). Enrichment in the ChIP samples at specific targets was calculated as a fraction of the Input (%).

Gene ontology pathway and biological process analyses

Functional enrichment analysis to determine enriched Gene Ontology pathways and biological processes was done using the CLEAN software package (20) and DAVID (21).

RNA isolation and gene expression profiling and microarray analysis

RNA was isolated using the RNeasy Midi Kit (QIAGEN) with DNAse treatment according to the manufacturer's protocol. The quantity and purity of the extracted RNA was evaluated using a NanoDrop (Nanodrop Technologies), and its integrity was measured using an Agilent Bioanalyzer (Agilent Biotechnologies, Santa Clara, CA). The RNA was prepared for microarray expression analysis by the NIEHS Microarray Core using the GeneChip 3' IVT Express kit protocol as per the manufacturer's instructions and the Affymetrix Human Genome U133 Plus 2.0 GeneChip® arrays (Affymetrix, Santa Clara, CA). Arrays were scanned in an Affymetrix Scanner 3000, and data were obtained using the GeneChip® Command Console Software (AGCC;

Version 1.1). Raw data files ('CEL files', Affymetrix array type Human Genome U133 Plus 2.0) were processed using appropriate R/Bioconductor packages. Specifically, Robust Multichip Average (RMA) (22) was performed using Entrez gene-based custom chip definition files (CDFs), version 13, (23), http://brainarray.mbni.med. umich.edu/Brainarray/Database/CustomCDF/genomic_ curated_CDF.asp, resulting in a single expression intensity measure per Entrez gene ID. All subsequent analyses were carried out on the $\log_2$ scale. An intensity filter was applied removing probe sets with intensity values constitutively below the first quartile in each array. This filter typically removes ~20% of the probe sets. To determine differentially expressed genes, a family-wise moderated *t*-test (24) (treated versus control) was performed followed by a Benjamini and Hochberg multiple testing correction procedure to control the false discovery rate (FDR). Unless otherwise noted, genes were considered differentially expressed if they had an FDR of ≤0.1 and were at least 2-fold up- or downregulated. Cluster analysis of per-gene normalized expression levels was performed using the CLEAN software package (20) where genes were normalized by subtracting the per-gene arithmetic mean of the respective control/untreated samples. The Smeenk *et al.* (12) data set was downloaded from the Gene Expression Omnibus (GEO; accession GSE22186) and re-processed as described earlier in the text. The microarray platform in this case was Human Exon 1.0 ST Array, but probes were summarized on per-gene basis again using Entrez gene based custom CDFs, version 13, (23), http://brainarray.mbni.med.umich.edu/Brainarray/ Database/CustomCDF/genomic_curated_CDF.asp.

Entrez gene IDs were used to map mRNA expression data ('probesets') and genes in the vicinity of ChIP-seq defined TF-binding sites (as defined earlier in the text). If multiple TF-binding sites were near a gene with a probeset on the array, each was mapped to that probeset.

Gene expression by qPCR

Total RNA was extracted from cells using the RNeasy Mini Kit including DNAse treatment (QIAGEN). Complementary DNAs were generated from 1 µg of purified RNA using TaqMan reverse transcription reagents from Applied Biosystems (Foster City, CA). qPCR was performed on an HT7900 system (Applied Biosystems) using pre-validated primers for the indicated genes, as previously described (8). Relative quantification values were calculated based on the 2-ΔΔCt method and using expression from β-2-microglobulin or *GAPDH* for normalization.

Lentiviral shRNA

Lentiviral constructs for stable depletion of human p53 mRNA were obtained from Sigma-Aldrich (MISSION shRNA); their integrity was confirmed by restriction analysis. Lentiviruses were produced by the NIEHS Viral Vector Core and packaged in HEK293T/17 cells (ATCC # CRL-11268) as previously described (25). All titers were determined by qPCR. U2OS cells were transduced with lentiviral particles, and stable clones were recovered after puromycin selection (2 µg/ml). Clones were characterized by qPCR and western blot analysis. Clones containing one of two different shRNAs against p53 were used: p53 shRNA-3755 (TRCN 0000003755) and p53shRNA-3756 (TRCN0000003756).

Cancer gene mutation analysis

The Cancer Genome Atlas data
Interpreted Cancer data from five cancers (glioblastoma, colon adenocarcinoma, ovarian cancer, sarcoma and prostate cancers) were downloaded from the Cancer Genome Atlas (TCGA) website (https://tcga-data.nci. nih.gov/tcga) using the CDGS-R package (http://www. cbioportal.org/public-portal/cgds_r.jsp), with querying performed using human Entrez gene identifiers as of 17 April 2012. The detailed description of TCGA data is available on the TCGA website.

Data processing
Genes with no data (N/As) across all tumor samples as well as tumor samples with no data (N/As) for all genes, were excluded from the analysis. A matrix of data was constructed in which rows and columns represent the genes and the tumor samples used in the analysis, respectively, and the cells are the measured values. For mutation data, each cell represents if a mutation was found in a gene.

Mutation
Background Gene Mutation Rate was calculated as the total number of mutations detected in sequenced tumor samples divided by the total number of genes sequenced. Based on Background Gene Mutation Rate and an assumed Poisson distribution, the *P*-value was then calculated and adjusted by the Benjamini and Hochberg FDR procedure.

Data availability

All ChIP-seq and microarray data generated for this study have been deposited in NCBI's GEO and are accessible through GEO Series accession number GSE46642.

RESULTS

Genome-wide mapping and characterization of p53 binding following drug treatments

Treatment of U2OS cells with DXR (0.6 µg/ml) or Nutlin (10 µM) resulted in the stabilization and increase of nuclear p53 protein levels (Figure 1A and Supplementary Figure 1A) after 24 h (a time point commonly used in studies of p53 activating drugs; see 'Discussion' section). Nutlin treatment induced a clear G1 and G2 cell cycle arrest, whereas DXR-treated cells were arrested at G2. DMSO treatment had no effect on cell cycle progression when compared with the No Treatment (NT) control. There was no substantial induction of apoptosis (measured as a sub-G1 fraction) for any of the treatment conditions at 24 h (Supplementary Figure S1B). ChIP-seq analysis using the p53 specific antibody DO-1 revealed considerable difference in the number and pattern of
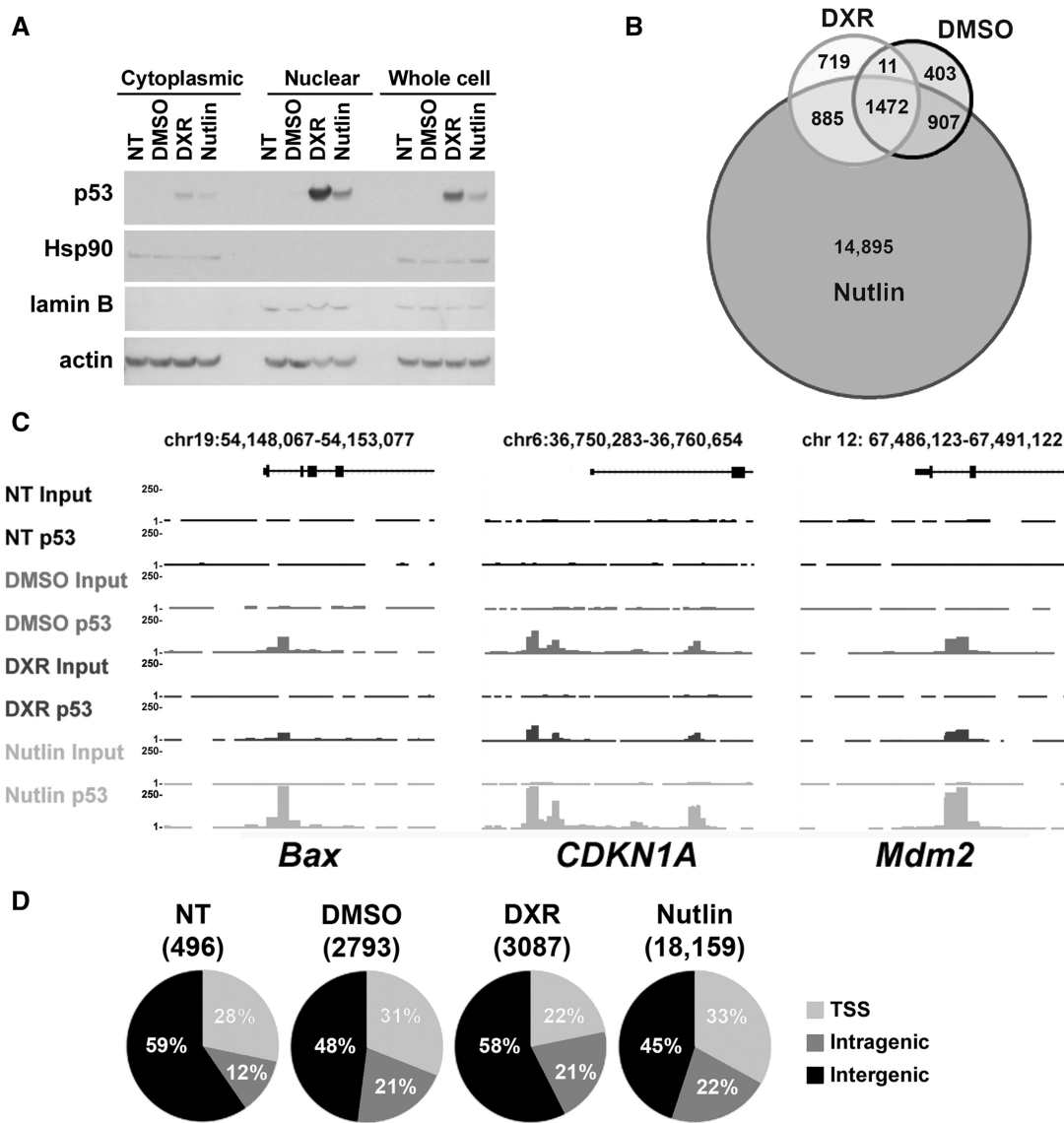
Figure 1. Genome-wide p53 binding after various treatments. (**A**) Representative western blot of p53 levels in U2OS cytoplasmic, nuclear or whole-cell extracts from untreated (NT) cells, or after 24 h treatment with 0.1% DMSO, 0.6 µg/ml DXR or 10 µM Nutlin. Lamin B and actin Hsp90, lamin B and actin are loading controls. (**B**) Number of p53 binding sites identified in U2OS 24 h after treatment with DMSO, DXR or Nutlin. (**C**) UCSC Genome Browser shots summarizing p53 ChIP-Seq and Input control data for known p53 target genes after various treatments. The Y axis corresponds to the number of reads. The same scale is used for all examples presented. (**D**) Genomic distribution of p53 binding sites relative to genic regions after the various treatments: TSS, 5 kb upstream through the first intron; Intragenic, second exon through 3′UTR; and Intergenic, everything else.

binding sites across the genome after DXR or Nutlin treatment. Genome-wide analysis of the ChIP-seq tags using the SISSRs peak-finding tool (16) identified 496 binding sites (peaks) for the NT control, 3087 sites after DXR treatment and nearly 6-fold more sites (18 159) in cells treated with Nutlin (Figure 1B, Supplementary Figure S1C and Supplementary Table S1). Surprisingly, the solvent DMSO (0.1%) alone caused substantial p53 binding (2793 sites). No relationship between the number of binding sites and the level of p53 expression was seen: DXR led to the largest induction of p53; Nutlin treatment induced less p53; there was little, if any, p53 induction after DMSO incubation (Figure 1A and Supplementary Figure S1A). Specific binding to the REs of the

well-characterized p53 target genes *BAX*, *CDKN1A* (p21) and *MDM2* attest to the quality of ChIP-Seq data (Figure 1C and Supplementary Figure S1D). To ensure the reliability of our ChIP-Seq data, we validated 36 of the identified p53 sites by ChIP-qPCR (described in Supplementary Table S2 and Supplementary Figure S1E and S1F). Nearly half of the sites bound by p53 after DXR treatment were also bound by p53 under DMSO and Nutlin treatments, suggesting common features to these presumably different modes of induced binding. However, nearly 80% of the Nutlin-induced p53-binding sites were unique (Figure 1B and Supplementary Figure S1C), with only seven sites in common between all three treatments and the NT control (Supplementary

Figure S1C), suggesting a different mechanism of p53 binding in the basal condition without any treatment.

For all treatment conditions, a considerable enrichment of p53-binding sites, relative to the genome composition, was observed in the surrounding region of a TSS, extending from 5 kb upstream through the first intron of a gene (Figure 1D). Nearly 30% of p53 binding was observed in these regions, which represent <5% of the genome; an additional ~20% was found in the remainder of intragenic regions (from second exon through 3′UTR). These data establish a strong specificity for p53 binding in regions of potential transcriptional regulation, regardless of the level of p53 induction, even after exposure to the solvent DMSO.

Redefining the p53 binding motif

A *de novo* motif analysis performed using the top 1000 p53 sites revealed a consensus motif (Figure 2A and Supplementary Data Set 1) that was nearly identical for all three treatments. This consensus sequence motif contained two copies of previously established decamers (half-sites) but with no spacer sequence in between them. Overall, this consensus is comparable with motifs derived from validated p53 targets (4) and from other genome-wide studies (10,12–14). Regions bound by p53 (200 nt long) were enriched for motifs known to be bound by the basic leucine zipper protein family of TFs, including AP-1 and FRA1, and other TFs, such as GCN4, LBP9 and SMAD (E-value < 0.1) (Supplementary Table S3). The enrichment of potential co-factor sites was at least 2-fold higher than in the corresponding randomly selected control sites. However, these associations differed between DXR and Nutlin treatments. No p53 binding motif was identified for the NT condition, reinforcing the suggestion that the low constitutive binding in U2OS cells may either be non-specific or indirect binding through protein–protein interactions.

Using the p53 consensus motif defined in this study, we queried all the p53 sites to categorize the types of p53 target sequences that were bound (Figure 2B). Among the binding sites, 68% after DXR, 28% after Nutlin and 57% after DMSO treatment had the p53 canonical motif consisting of two decamers with a zero spacer. An additional 7–11% of all binding sites had 1–15 base spacers between the two RRRCWWGYYY decamers, regardless of treatment. Importantly, we also found that 3–4% of binding sites contained just one decamer (p53 half site), reinforcing our previous observations that half sites might play a role in the p53 regulatory network (3,26). Only 7% of binding sites (35 of 496) in the NT control contained a sequence resembling at least a p53 half site. A majority of Nutlin induced p53-binding sites contained neither the consensus full site nor the half site. For those p53-binding sites where no p53 motif was found, we validated nine by ChIP-PCR (Supplementary Figure S1G and Supplementary Table S2). For NT and Nutlin conditions where most of the p53 peaks have no p53 motif, we identified using TRANSFAC that AP-1, BACH1, BACH2, FRA1 and/or GCN4 motifs were associated with p53-binding peaks (E-value < 0.1). Furthermore,

after a *de novo* motif search of the top 1000 p53 peaks that do not harbor any p53 decamer (Supplementary Data Set 2), we found the same TF motifs enriched among the DMSO and Nutlin 'no motif' peaks as described earlier in the text, in addition to Sp1, Sp3 and PITX2. The common TF profile between DMSO and Nutlin suggests that the p53-binding signature for Nutlin is largely due to the DMSO solvent used with Nutlin. On the other hand for NT and DXR, the *de novo* identified motifs were not associated with any TF other than HSF. p53 has been reported to bind other non-canonical motifs, such as the TGYCC microsatellite sequence described by Contente *et al.* (27) in the promoter of the *PIG3* gene (also known as *TP53I3*). The motif analysis by MEME identified a p53-like motif within the 300 bp ChIP region, which also contains the TGCYY microsatellite sequence. When we masked out the p53 REs and performed a *de novo* motif, we did not identify any additional (TGYCC)n related motifs. To complement this analysis, we also searched the genome for segments with greater than five copies of the TGYCC motif and asked whether p53 bound these regions. Only 7 of 87 (TGYCC) that had greater than five repeat segments (consecutive/no gaps) were bound by p53, which includes the *TP53I3/PIG3* TGCYY microsatellite region that contains 15 repeats and six regions with less than nine repeats. Most of the bound sites were in the intergenic region, and only the *TP53I3/PIG3* gene was induced after p53 activation.

Interestingly, p53 target sequences with a spacer of one or more bases were roughly 20- to 30-fold less abundant as compared with sites with a zero spacer (Figure 2C) for each of the three treatments. In addition, the average ChIP-over-Input fold enrichment for the p53 sites containing the consensus motif with a zero spacer is roughly 2-fold higher compared with those sites that contain a motif with spacers or those that contain half sites (Figure 2D). The total number of p53 bound sites with motifs that contain spacers or just half sites amount to ~10–15% of all the binding sites. ChIP-qPCR experiments confirmed p53 binding at these sites, which were included among the validated sites in Supplementary Figure S1E. Thus, although the consensus sequence consists of two decamers with no intervening spacer, the minimal *in vivo* binding unit appears to be a single decamer (see 'Discussion' section).

p53 binding is necessary, but not sufficient for transactivation

Global gene expression was examined using microarray analysis of the NT and the 24 h DMSO, Nutlin or DXR-treated cells. To establish clear differences in expression, we identified genes for which changes in mRNA levels were ≥2-fold and had an FDR of ≤0.1. As shown in Figure 3A, 1717 genes were differentially expressed between DXR and NT, whereas for Nutlin versus DMSO, only 416 genes were differentially expressed; only 220 genes had altered gene expression in both DXR and Nutlin treatments. In Figure 3B, we present a heat map of all the differentially expressed genes and a cluster
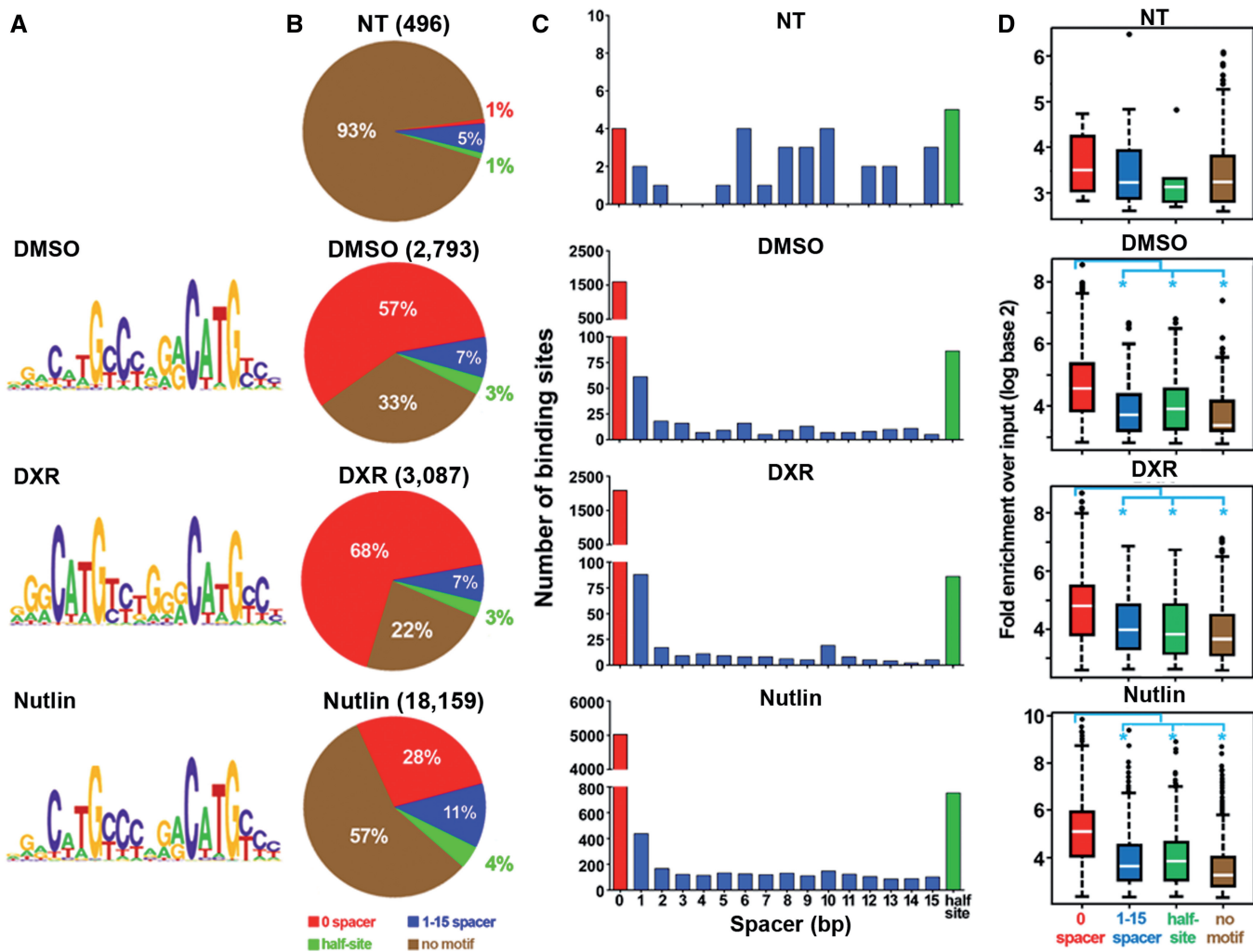
Figure 2. Motif analysis and distribution of p53-binding sites. (**A**) Consensus p53-binding motifs identified from the top 1000 p53-binding sites for each treatment. (**B**) Characterization of p53-binding sites based on the type of sequence bound: zero spacer, consensus motif containing two decamers with no spacer sequence as shown in (**A**); 1–15 spacer, two decamers with a spacer sequence of length 1–15 nt; half site, one decamer; no motif, no evidence of zero spacer, 1–15 spacer, or a half site. (**C**) Frequency of p53-binding sites containing zero spacer (red), 1–15 spacer (blue), and half sites (green). (**D**) Box-plot depicting distribution of ChIP-over-Input fold enrichment for p53-binding sites containing zero spacer, 1–15 spacer, half site and no motif. Asterisk indicates statistically significant ($P < 10^{-7}$).

functional analysis. Cell cycle related, p53 signaling and apoptosis categories were enriched based on a combination of GO and KEGG analyses (the complete cluster functional enrichment analysis is available as supplemental information (Supplementary Data Set 4). Further KEGG and GO analysis of all the differentially expressed genes revealed that the global gene expression pathways activated by DXR and Nutlin includes processes commonly associated with p53 functions (Figure 3C and Supplementary Data Set 4) such as cell cycle checkpoints and regulation of cell proliferation, consistent with the cell cycle arrest induced by these two drugs (Supplementary Figure S1B). Remarkably, little change in gene expression was observed following incubation of cells with DMSO versus NT, even though DMSO induces nearly 3000 binding sites, about half of which also are induced during DXR treatments (Figures 3B and 1B). Additionally, although the number of sites bound by p53

were six times greater for Nutlin than for DXR (Figure 1B), expression changes induced by Nutlin were much less dramatic compared with those induced by DXR (Figure 3B). Taken together, these data support the view that p53 binding does not predict transactivation.

To explore more closely the relationship between p53 binding and gene expression outcome, we examined genes that were bound by p53 near the TSSs and in the intragenic region (Figure 3D and E). Among the genes with p53-binding sites near a TSS, 25% after DXR and only 5% after Nutlin treatment exhibited ≥2-fold change in expression (Figure 3D), whereas for the intragenic region, only 14 and 3.8% of sites have associated gene expression, respectively (Figure 3E), suggesting that the association between p53 and binding and expression may be more frequent for p53 bound near the TSS. Only 48 genes close to TSS were bound by p53 and differentially expressed after both the DXR and the Nutlin
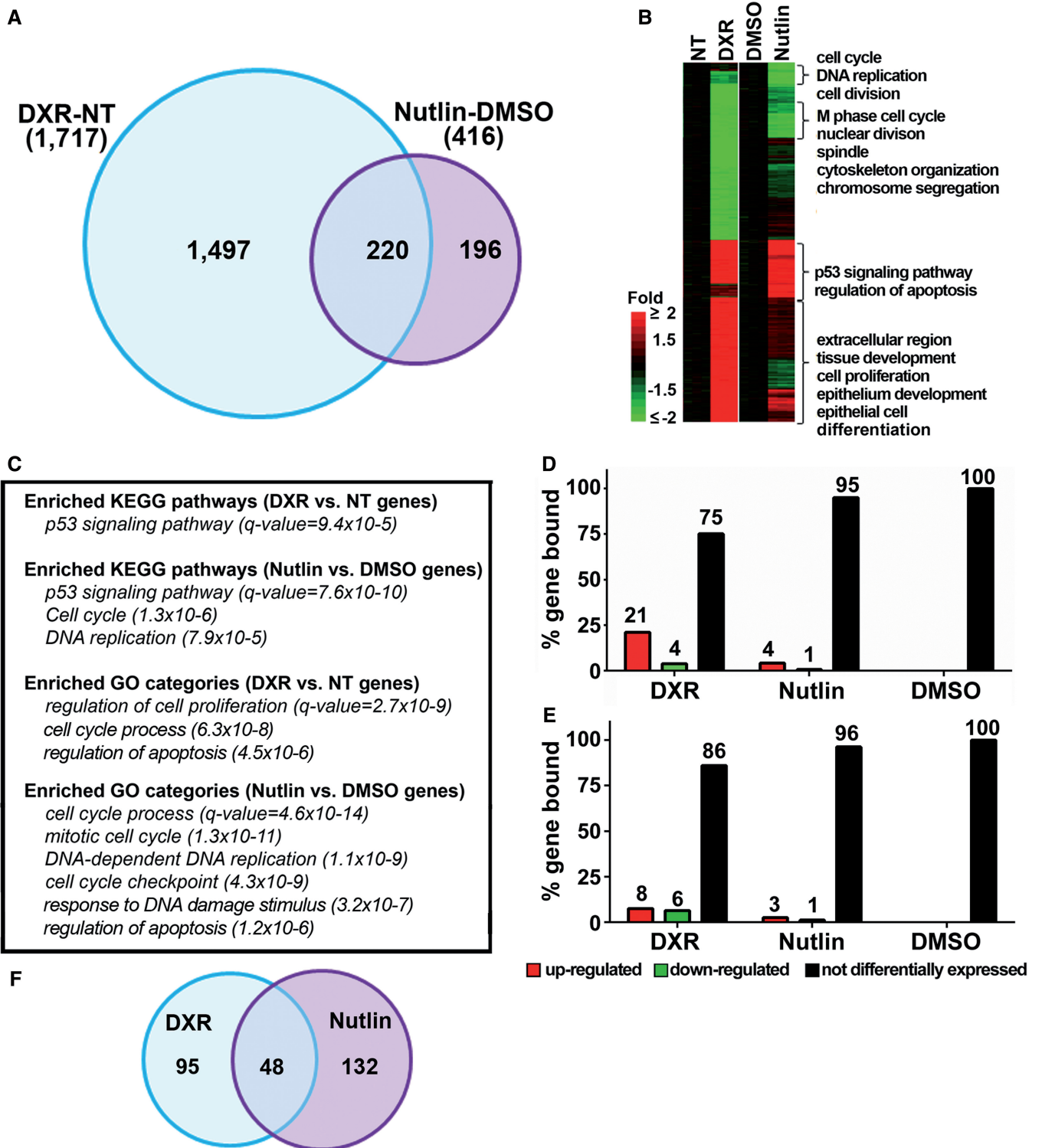
Figure 3. Gene expression analysis following various treatments. (**A**) Number of genes with altered expression following DXR or Nutlin treatment as compared with the control NT or DMSO controls, respectively. (**B**) Heat map showing gene expression changes based on microarray analysis of cells after 24 h of DXR, DMSO or Nutlin treatment. DXR and DMSO changes were assessed in relation to NT, and Nutlin changes were assessed in relation to DMSO. The KEGG/GO terms shown are non-redundant terms manually selected among the most significantly enriched categories for each of the bracketed clusters. The full lists of terms are available in Supplementary Data Set 5. (**C**) KEGG pathways and GO biological processes enriched among all the differentially expressed genes in (**A**) after various treatments. (**D**) Breakdown of expression changes of genes with a p53-binding site near the TSS or (**E**) in the intragenic region. (**F**) Number of genes that were bound by p53 near a TSS and were differentially expressed after DXR or Nutlin treatment.

treatments (Figure 3F). Thus, binding to most sites near TSSs did not lead to expression changes of the associated genes.

Among those genes with altered expression, there was a strong bias towards activation versus repression, with 83 and 86% of the genes, respectively, showing increased expression after DXR or Nutlin treatment (Figure 3D). p53 has been reported to repress more genes than it activates (28); however, relatively few examples of repression by direct p53 binding are known. For the set of genes that bound p53 near their TSS and showed changes in expression after DXR or Nutlin, <20% were downregulated. We did not find any evidence for a repression specific p53 binding motif, contrary to a previous report (29); instead, ∼60% contained a p53 consensus full site or half site. As in most cases DXR- or Nutlin-induced p53 binding near a TSS did not alter associated gene expression, we explored the possibility of chromatin modifications playing a role in the selective transcriptional responses to the different treatments (Figure 4 and Supplementary Figure S2). We examined H3K4me3 and H3K27me3, active and repressive chromatin marks, respectively, at several p53 target sites. Among the 36 genes examined for p53 binding (Supplementary Figure S1E), a vast majority of those that had a ≥2-fold increase in gene expression also had a ≥2-fold increase in H3K4me3 levels at their promoters: 21/23 for DXR and 21/27 for Nutlin treatments (Figure 4A and B and Supplementary Figure S2A–C). Increases in expression generally were associated with decreases in H3K27me3 following Nutlin treatment, consistent with upregulation of the genes. However, for DXR treatment few upregulated genes showed a reduction in H3K27me3 levels. For DMSO treatment, despite the lack of gene expression changes, there were p53 binding-associated changes in both histone modifications (Figure 4C and Supplementary Figure S2B and C), with nearly all showing increased H3K4me3 levels. Thus, although changes in chromatin landscape between treatments suggest a role for histone modifications in p53-mediated gene expression changes, it is clear that binding at a p53 motif *per se* is associated with increased H3K4me3. Together, these data imply a two-stage mechanism for p53 transactivation in which p53 binding-only constitutes the requisite first stage.

Novel p53 target genes

Most experimentally validated p53-regulated genes contain a p53 motif within either their promoter or the first intron (4). The list of 275 genes that were both bound by p53 and differentially expressed (Supplementary Data Set 3) includes 219 potential new p53 targets as well as 56 previously established p53 responsive genes (4,14). Comparing the set of 219 targets with putative p53 targets identified in other p53 genome-wide studies (11,12) (Supplementary Data Set 3), we identify 149 potential new p53 target genes. Functional clustering of gene annotations using DAVID (21) revealed, in addition to the usual p53 characterized functions (cell cycle arrest, apoptosis, DNA repair), an enrichment for genes involved in a broad variety of biological process including differentiation, metabolism, protein transport, nucleotide binding, hydrolase and deaminase activities and signal transduction. Based on nucleotide identity, ∼58% of the sequences of p53-binding sites associated with these genes were conserved in rodents (Supplementary Data Set 3), suggesting conservation of p53 binding and function across species.

Of the 149 potential new p53 targets, 11 candidate genes were characterized further for responses to DMSO, DXR or Nutlin (Figure 5). Included were genes with p53 REs having zero spacer (*CPEB4, APOBEC3C, ARHGEF3, ACYP2, SYTL1, DYRK3, NF2*), a 9 nt spacer (*LASS5*) or no p53 motif in the region near the TSS (*INPP1, EBI3, CDC27*). We also examined three candidate genes that were bound by p53 but showed no expression changes: *APOBEC3H*, zero spacer; *CYP3A7* and *EML2*, half sites. Quantitative RT-PCR analysis confirmed that among the 14 genes examined, 12 were upregulated while 2 were downregulated by DXR or Nutlin treatment (Figure 5A). As previously observed in the microarray analysis, no significant changes were observed in DMSO-treated cells. Although *APOBEC3H, CYP3A7* and *EML2* showed no differential expression using microarray analysis, expression changes could be detected with qRT-PCR.

The p53-dependent gene expression changes were verified using two U2OS cell lines infected with lentiviral vectors harboring small-hairpin RNAs that target p53 mRNA (termed p53shRNA 3755 and 3756) (Supplementary Figure S3A). The levels of p53 protein after drug treatments were reduced in the p53-silenced U2OS cells (Supplementary Figure S3B). p53-dependent gene expression changes for 14 genes examined in control (parental or scrambled) cells (Figure 5A and B) were completely lost in p53-depleted cells confirming that these are genuine p53 targets (Figure 5C and D). p53 dependency following DXR or Nutlin treatment was further verified for seven genes using the breast carcinoma MCF7 cell line stably expressing p53shRNAi and its vector control, the isogenic colorectal carcinoma HCT116 $p53^{+/+}$ and $p53^{-/-}$ cells, as well as the lung carcinoma cell lines A549 ($p53+$) and H1299 ($p53^-$) (Supplementary Figure S4A–F). Extrapolating from these results, most of the 219 differentially expressed genes with an associated p53-binding site near the TSS are likely to be direct p53 targets.

Novel p53 targets are associated with cancer

Given that p53 is a prominent tumor suppressor and that its target genes might be important in cancer etiologies, we interrogated recent cancer genome mutation data from TCGA (http://www.cancer.gov) to determine whether mutations were increased among the 275 differentially expressed genes associated with p53-binding sites. Although data are available for several types of tumors, only a few have data for healthy matched control tissue and/or sufficient samples to provide statistical confidence. For three tumor types (ovarian, glioblastoma and colon) with numerous samples, we applied an FDR of $P < 0.001$ as described in the 'Materials and Methods' section to
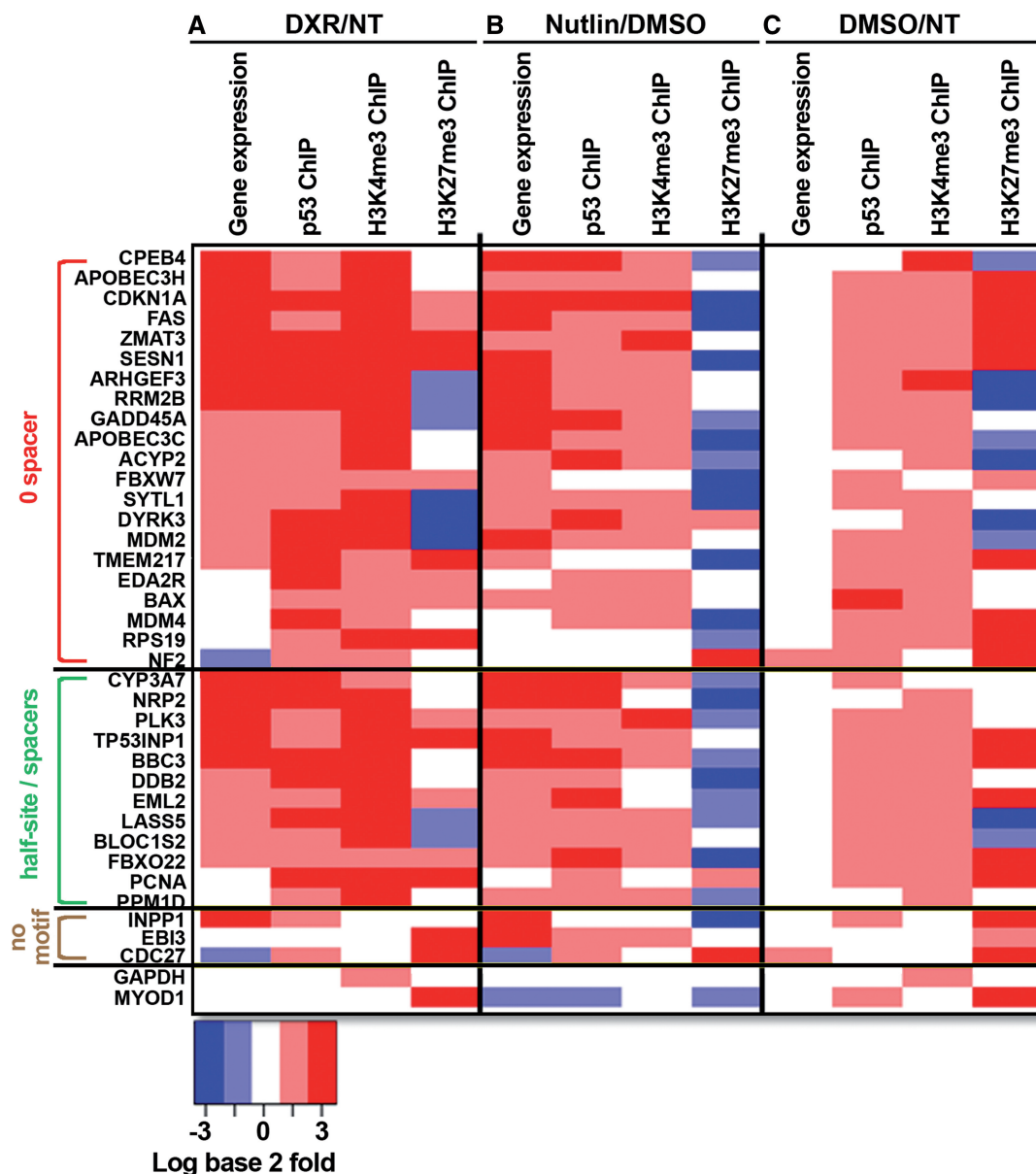
Figure 4. Correlation between p53 binding, changes in gene expression and histone modifications H3K4me3 and H3K27me3 at 24 h after DXR, Nutlin or DMSO treatments. The heat map represents fold changes in mRNA levels quantified using qRT-PCR, as well as fold changes in p53, H3K4me3 and H3K27me3 occupancy quantified using ChIP-qPCR (see also Supplementary Figures S1D and S2A–C). Presented are results for 36 genes that were bound by p53 and two genes (*GAPDH* and *MYOD1*) that were not bound by p53. Genes are clustered into groups depending on whether the p53 target site contains a consensus p53-binding motif containing two decamers with no spacer (zero spacer), two decamers with a spacer sequence (spacers), or one decamer (half site) or no motif. Fold changes for gene expression were normalized to β-2-microglobulin, whereas p53, H3K4me3 and H3K27me3 ChIP-qPCRs were normalized to the negative controls GAPDH, MYOD1 and GAPDH, respectively. (**A**) DXR/NT, (**B**) Nutlin/DMSO and (**C**) DMSO/NT.

identify those for which there were significant frequency increases in mutations. The following seven genes were mutated with significant recurrence in human colon cancer (Supplementary Data Set 5): *ABCA12*, *ASCC3*, *CDC27*, *CDH10*, *EPHA5*, *FBXW7* and *SCN2A*. Except for *FBXW7*, all the remaining genes are newly identified p53 targets. No significantly recurring mutations within these genes were identified in the other two tumor types. Additionally, we found that the neurofibromin 2 (*NF2*) gene has been reported to be frequently mutated in cancers of tissues associated with the nervous system (30).

DISCUSSION

Genome-wide binding defines the potential for a TF to influence genomic expression and the corresponding TF 'cistrome'. As a master regulatory factor, p53 is important in the etiology of most cancers as well as their treatment. We establish in a well-studied human cancer cell line that p53 exhibits a wide drug-dependent variability in responses including binding and changes in gene expression. p53 transcriptional activity is highly complex and not simply dependent on the levels of p53, binding to specific sites or to changes in common chromatin
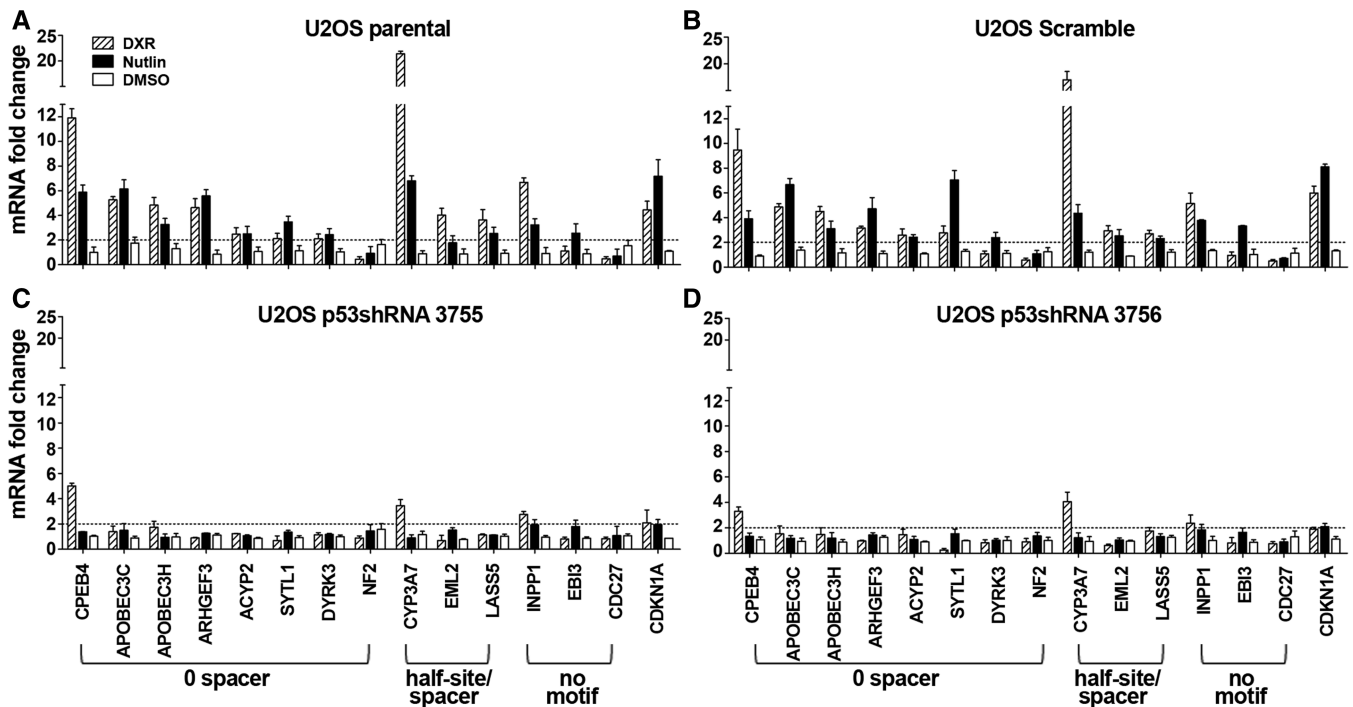
Figure 5. Validation of new p53 direct target genes. Real-time PCR validation of newly identified putative p53 targets in U2OS cells using shRNA-mediated silencing of p53 expression. Parental U2OS cells treated with (**A**) no shRNA, (**B**) scramble (control) shRNA, (**C**) p53shRNA 3755 or (**D**) p53shRNA 3756. mRNA levels were quantified in U2OS cells 24 h after NT, DMSO, DXR or Nutlin treatment using qRT-PCR and normalized to *GAPDH*. mRNA levels are shown as fold change after treatment relative to the NT control. Expression of the *CDKN1A* gene was used as a positive control. Shown are averages of three independent experiments. The dashed line corresponds to a 2-fold change in expression. Error bars represent standard deviation (SD).

modifications. Using a genome-wide approach, we identify a single decamer sequence as the minimal unit of binding in human cells. Our study also reveals many new p53 targets, some of which are frequently mutated in cancer.

Unexpected p53 binding patterns in untreated cells and cells treated with the solvent DMSO

Approximately 500 sites were bound by p53 in the untreated (NT) control cells (Supplementary Figure S1C), but only one of these corresponded to a site associated with a known p53 target gene. This was surprising in light of previous reports of p53 being prebound to known targets. However, using ChIP-qPCR, we did observe low levels of binding to several known p53 target genes, such as the well-studied *CDNK1A*, consistent with previous reports that also examined U2OS cells (31,32). Nevertheless, the newly discovered p53 ChIP-seq binding sites in untreated cells do not have a p53-binding motif (Figure 2B). Recent studies in mouse embryonic stem cells also reported p53 binding to numerous sites in NT control cells (33,34). The absence of a p53 motif in >90% of the p53 sites in NT cells (Figure 2B) suggests one of the following possibilities: p53 binding to DNA structures, indirect binding of p53 through other TFs such as those presented in Supplementary Table S3 or non-specific binding. Furthermore, many of the sites bound by p53 in

NT cells are not bound by p53 after the various drug treatments (Supplementary Figure S1C).

Surprisingly, the number of DMSO-induced p53 sites was comparable with that after DXR treatment, and the majority of these sites contained a p53 motif (Figure 2B). The validation of several of these sites by ChIP-PCR confirmed that DMSO treatment induced ~2-fold more p53 occupancy at most of the 36 sites tested in comparison with the NT condition. In support of our findings, Nikulenkov *et al*. (11) also recently reported p53 binding globally in mock-treated MCF7 cells, although the nature of the control solvent was not described. Little change in p53 protein level was observed after 24 h treatment with 0.1% DMSO, consistent with other reports (35,36). Thus, the effects of DMSO are not simply owing to a change in p53 levels (Figure 1A and Supplementary Figure S1A). The general view of a relationship between p53 levels, binding and transcriptional activation are challenged by our findings with DMSO. Despite the appearance of nearly 3000 p53-binding sites (Figure 1B), DMSO treatment affected the expression of only 20 genes, none of which had associated p53 binding (Figure 3B), suggesting DMSO modulates p53 binding specificity and/or chromatin structure. These data support the conclusions that (i) p53 binding does not predict transactivation, (ii) DMSO may not be an appropriate control to assess p53 binding/function and (iii) DMSO treatment could be used as an agent to poise p53 binding without any transactivation.

Dramatic discordance between DXR- and Nutlin-induced p53 binding and expression

Although >50% of the Nutlin-induced expression changes also were induced in response to DXR, this overlap was reduced to 33% when one considers only genes that bind p53 near a TSS (Figure 3A and F). The limited overlap among the differentially expressed genes after DXR or Nutlin treatments is consistent with previous reports of p53-mediated stress-specific transcriptional responses (37–40). We choose a 24 h time point because this is a common window used in the field for several of the activating p53 drugs, including those used in our study. In addition, given that therapeutic treatments with the agents we used are continuous and can be long-term, we consider our experimental design as a reasonable approach for addressing the consequences of drug treatment of tumors with chemotherapeutic agents. In a recent study that also used U2OS cells, only 36% of the changes in gene expression overlapped following treatment with Actinomycin D or Etoposide for 24 h (12).

The differential regulation of genes in response to the various stresses is thought to result in part from post-translational modifications of p53 (and probably other proteins) through treatment-specific activation of various signaling pathways and/or recruitment of various co-factors and binding proteins (1,2). Acetylation of the p53 DNA-binding domain was shown recently to be required for transactivation, but not for sequence-specific binding (41). The relatively small number of expression changes for Nutlin treatment might be due to few p53 post-translational modifications (42) as compared with the effects of DXR treatment (43).

We also found substantial differences in p53 binding after the various treatments. After DXR or DMSO treatments, nearly 50% of the sites bound by p53 overlapped. However, for Nutlin, <5% overlapped with DXR and DMSO sites. The large increase in sites after Nutlin compared with DXR treatment was unexpected and suggests that Nutlin treatment leads to changes in chromatin architecture or nucleosome occupancy that are not anticipated for a drug selected for its specific effect on the interaction between p53 and MDM2 (42).

To address possible changes in chromatin after drug treatment, we examined H3K4me3 and H3K27me3 modifications. Previously, H3 and H4 histone acetylation changes were observed for 20% of target genes in human cells when p53 is overexpressed (44) or after 5-fluorouracil treatment (45), whereas different patterns of H3K4 di- and H3K27 tri-methylation were detected in repressed versus activated p53 genes in mouse embryonic stem cells (34). Of the 36 genes examined in depth, a strong relationship between p53 binding and changes in H3K4 trimethylation, a modification generally associated with gene activation, was observed. Here, we establish that increases in H3K4me3 is actually a marker specific to binding, as 69–86% of bound sites following DMSO, DXR or Nutlin treatment were accompanied by increases in H3K4me3, regardless of changes in gene expression. Unlike H3K4me3, the decrease in H3K27me3 was treatment specific and did not relate to binding or expression.

No direct association between p53 binding and induced gene expression changes was found after activating p53, as four times as many genes changed expression after DXR than after Nutlin treatment, even though Nutlin induced six times more p53 binding sites. Comparing our binding and gene expression results in U2OS cells with those from Smeenk *et al.* (12) using 24 h treatments with Actinomycin D or Etoposide, we found 270 common p53 binding sites that are near a TSS for all treatments (DMSO, DXR, Nutlin, Etoposide and Actinomycin D; Supplementary Data Set 6). However, only 11 genes associated with those p53-binding sites were upregulated in common after treatment with DXR, Nutlin, Actinomycin D or Etoposide (excluding DMSO). This finding suggests a common p53 binding profile that precedes a stress-specific response. Overall, as was established for DMSO, binding alone is not sufficient for inducing p53-mediated changes in expression.

The p53 minimal unit of binding *in vivo* is a half-site

The commonly accepted p53 consensus motif is two decamers with a spacer that can extend to greater than 10 bp based on *in vitro* results (5,6). Recent genome-wide p53 binding studies have shown that the most prevalent p53 binding motif, regardless of treatment, lacks any spacer. Similarly, we found that the predominant p53 motif bound after DMSO, DXR or Nutlin treatments is a 20-mer sequence comparable with that in recent reports. In addition, we established that *in vivo* p53 can bind just one decamer sequence (half site).

The *in vivo* low incidence of p53-bound sites with decamers separated by a spacer relative to tandem decamers without spacers was unexpected, given the widely accepted views of p53 binding being relatively independent of spacer length based on *in vitro* studies (46,47) As the likelihood of two decamer sequences in the genome is equal, regardless of whether they occur next to each other or are separated by a spacer, the number of two decamers separated by up to 15 bases is expected to be 15 times more prevalent than those separated by a zero spacer. But we observed that p53 sites containing decamers separated by up to 15 bases was >37 times less common than those containing a zero spacer (Figure 2B). These results are consistent with our previous findings, where we showed a higher p53 binding affinity for 0 or 1 base spacers compared with two or more base spacers (48).

Remarkably, the number of half sites bound by p53 is as much as half of the total number of p53 sites containing two decamer sites separated by 1–15 nt spacers. Before this, no studies had addressed genome-wide p53 binding to half sites, although we had previously demonstrated p53 binding to half sites and transactivating at some half-site targets to levels comparable with full sites (26,49). *In vitro* binding by dimer p53 molecules to half sites has been reported, and recent p53-DNA crystal structures were based on half sites (50,51). As the DNA-binding domain is the most highly conserved domain among p53 family members, it is likely that some half
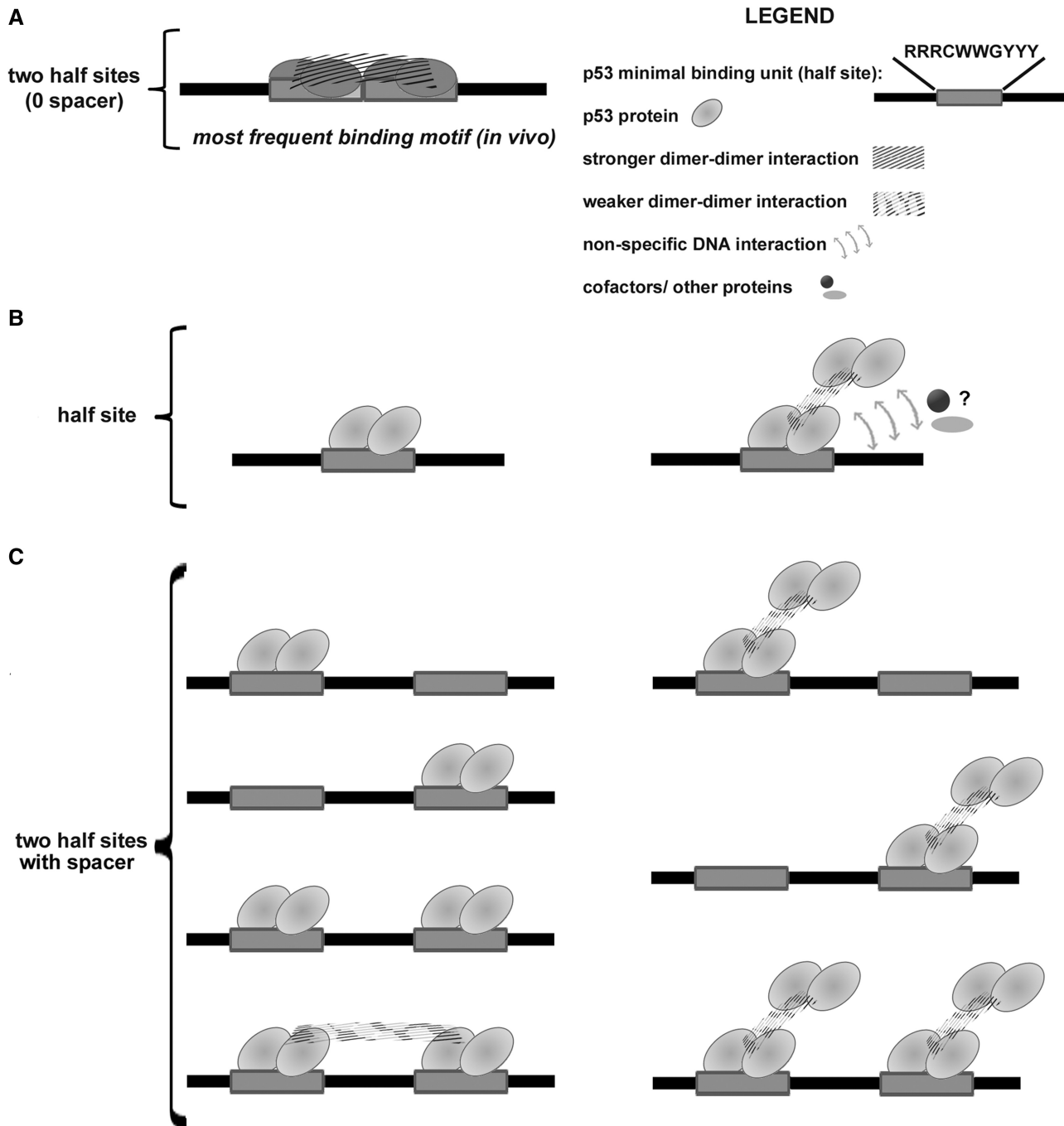
Figure 6. Proposed models for p53-DNA interaction at half-sites (1 decamer) and full sites (two decamers) with and without a spacer. The minimal p53-binding unit is a half site composed of the consensus decamer RRRCWWGYYY (except for the NT control) based on binding to half sites (decamers) and full sites (two decamers) with a spacer. The binding models assume that p53 exists as a dimer and tetramers are formed following p53 dimer binding to a target sequence. p53-bound half sites as well as full sites containing a spacer are much less frequent across the genome than full sites with zero spacer. (**A**) Efficient tetramer binding at full sites with zero spacer. Strong interactions between dimers and target sequences can occur when the dimers bind to a pair of decamers with no spacer. The dark grey ovals indicates possible changes in p53 (such as configuration) that enable the strong binding. (**B**) Binding to a half site, which is much less efficient than to a full site with no spacer, could be due to weaker interactions of the decamer sequence with just one p53 dimer or with a dimer of dimers, where one dimer makes contact with the decamer sequence, whereas the second dimer interacts non-specifically with nearby DNA. (**C**) Binding to a full site with a spacer is proposed to occur through p53-decamer interactions that are described for the half sites in (B) with the additional possibility of weak interactions between the two bound p53 dimers (lower left panel).

sites may serve as regulatory binding sites for other p53 family members, as demonstrated for p63 and p73 (49,52).

ChIP-over-Input fold enrichment is a good indicator of binding affinity/stability (16). The lack of difference

in ChIP-over-Input enrichment between p53 sites containing half sites versus full sites with up to 15 nt spacer (Figure 2B) suggests that p53 binding to decamers separated by a spacer is comparable with p53 binding

one of the decamers. The pattern of p53 binding *in vivo* at half sites and spacer-containing sites lead us to propose a model (Figure 6) in which the decamer is the minimal unit of p53 binding across the human genome, and that two decamers with a zero spacer are highly preferred for stable/stronger binding.

New p53 targets are associated with cancer

Recent advances in whole-genome sequencing of cancers allowed us to explore possible relationships between the 219 differentially expressed and p53-bound putative p53 target genes and cancer-associated mutations. We found seven genes, *CDC27*, *EPHA5*, *NF2*, *ABCA12*, ASCC3, *CDH10* and *SCN2A* that were frequently mutated in human tumors. These genes, which are described in detail in Supplementary Data Set 5, are involved in different biological processes such as cell cycle regulation (*CDC27*), nervous system development and signaling (*NF2*, *EPHA5*, *SCN2A*), lipid metabolism (*ABCA12*), cellular adhesion (*CDH10*) and DNA repair (*ASCC3*). All these genes have been associated with the development of different diseases and malignancies. Mutations in the tumor suppressor gene *NF2* (also known as Merlin) synergize with p53 mutations in the development of meningiomas and other neural crest-derived tumors (53–56). The *NF2* gene product is involved in interactions between components of the cytoskeleton and the cell membrane.

In addition, from the set of validated p53 target genes, we found that the *FBXW7* gene (F-box WD repeat domain containing 7) was the only previously identified p53 target highly mutated in several types of tumors including hematopoietic and lymphoid tissue, large intestine, endometrium and biliary tract tissues (Supplementary Data Set 5). *FBXW7* is a tumor suppressor gene that encodes an E3 ubiquitin ligase, which is part of the SCF ubiquitin ligase complex that mediates ubiquitin-dependent proteolysis of several key oncoproteins, such as Cyclin E1, c-Myc and c-Jun (57,58). Mutations or deletions in *FBXW7* are commonly associated with endoreduplication, thereby altering cell cycle progression and chromosome stability. *FBXW7* and *TP53* mutations have been associated with intestinal and gastric tumors, contributing to a poor prognosis (59,60).

Except for *NF2*, recurrent mutations in the other seven genes were found in colon cancers. The frequency of p53 mutations in colon cancers is 54% (TCGA database). As 57–73% of the tumors containing mutations in one of the seven genes are wild-type for p53, these gene mutations may be important contributors to cancer development.

We have taken a comprehensive genome-wide approach to characterizing the p53 network that has revealed novel aspects including binding, expression and chromatin changes that together can dynamically influence the affinity of p53 for its target sites and participate in the fine tuning of the p53 transcriptional network. It will be interesting to compare responses in primary cells, *ex vivo*, or after patient treatments and assess individual variability in the network as has already been done for a limited set of p53 target genes (8). In addition, we have identified new p53 targets some of which are frequently mutated in human tumors. We have gone on to suggest that several of the new targets could play a role in tumor development. Given the importance of p53 in tumor biology, knowledge of new targets and cancer-associated mutants is expected to help in therapeutic approaches.

REFERENCES

1. Espinosa,J.M. (2008) Mechanisms of regulatory diversity within the p53 transcriptional network. *Oncogene*, **27**, 4013–4023.
2. Laptenko,O. and Prives,C. (2006) Transcriptional regulation by p53: one protein, many possibilities. *Cell Death Differ.*, **13**, 951–961.
3. Menendez,D., Inga,A. and Resnick,M.A. (2009) The expanding universe of p53 targets. *Nat. Rev. Cancer*, **9**, 724–737.
4. Riley,T., Sontag,E., Chen,P. and Levine,A. (2008) Transcriptional control of human p53-regulated genes. *Nat. Rev. Mol. Cell Biol.*, **9**, 402–412.
5. El-Deiry,W.S., Kern,S.E., Pietenpol,J.A., Kinzler,K.W. and Vogelstein,B. (1992) Definition of a consensus binding site for p53. *Nat. Genet.*, **1**, 45–49.
6. Funk,W.D., Pak,D.T., Karas,R.H., Wright,W.E. and Shay,J.W. (1992) A transcriptionally active DNA-binding site for human p53 protein complexes. *Mol. Cell. Biol.*, **12**, 2866–2871.
7. Menendez,D., Krysiak,O., Inga,A., Krysiak,B., Resnick,M.A. and Schönfelder,G. (2006) A SNP in the flt-1 promoter integrates the VEGF system into the p53 transcriptional network. *Proc. Natl Acad. Sci. USA*, **103**, 1406–1411.
8. Menendez,D., Shatz,M., Azzam,K., Garantziotis,S., Fessler,M.B. and Resnick,M.A. (2011) The toll-like receptor gene family is integrated into human DNA damage and p53 networks. *PLoS Genet.*, **7**, e1001360.
9. Yan,J., Menendez,D., Yang,X.P., Resnick,M.A. and Jetten,A.M. (2009) A regulatory loop composed of RAP80-HDM2p53 provides RAP80-enhanced p53 degradation by HDM2 in response to DNA damage. *J. Biol. Chem.*, **284**, 19280–19289.

10. Botcheva,K., McCorkle,S.R., McCombie,W.R., Dunn,J.J. and Anderson,C.W. (2011) Distinct p53 genomic binding patterns in normal and cancer-derived human cells. *Cell Cycle*, **10**, 4237–4249.

11. Nikulenkov,F., Spinnler,C., Li,H., Tonelli,C., Shi,Y., Turunen,M., Kivioja,T., Ignatiev,I., Kel,A., Taipale,J. *et al.* (2012) Insights into p53 transcriptional function via genome-wide chromatin occupancy and gene expression analysis. *Cell Death Differ.*, **19**, 1992–2002.

12. Smeenk,L., van Heeringen,S.J., Koeppel,M., Gilbert,B., Janssen-Megens,E., Stunnenberg,H.G. and Lohrum,M. (2011) Role of p53 serine 46 in p53 target gene regulation. *PLoS One*, **6**, e17574.

13. Smeenk,L., van Heeringen,S.J., Koeppel,M., van Driel,M.A., Bartels,S.J., Akkers,R.C., Denissov,S., Stunnenberg,H.G. and Lohrum,M. (2008) Characterization of genome-wide p53-binding sites upon stress response. *Nucleic Acids Res.*, **36**, 3639–3654.

14. Wei,C.L., Wu,Q., Vega,V.B., Chiu,K.P., Ng,P., Zhang,T., Shahab,A., Yong,H.C., Fu,Y., Weng,Z. *et al.* (2006) A global map of p53 transcription-factor binding sites in the human genome. *Cell*, **124**, 207–219.

15. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.

16. Jothi,R., Cuddapah,S., Barski,A., Cui,K. and Zhao,K. (2008) Genome-wide identification of *in vivo* protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res.*, **36**, 5221–5231.

17. Bailey,T.L. and Elkan,C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **2**, 28–36.

18. Bailey,T.L. and Gribskov,M. (1998) Combining evidence using p-values: application to sequence homology searches. *Bioinformatics*, **14**, 48–54.

19. Wingender,E., Chen,X., Hehl,R., Karas,H., Liebich,I., Matys,V., Meinhardt,T., Prüss,M., Reuter,I. and Schacherer,F. (2000) TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res*, **28**, 316–319.

20. Freudenberg,J.M., Joshi,V.K., Hu,Z. and Medvedovic,M. (2009) CLEAN: CLustering Enrichment ANalysis. *BMC Bioinformatics*, **10**, 234.

21. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.

22. Bolstad,B.M., Irizarry,R.A., Åstrand,M. and Speed,T.P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, **19**, 185–193.

23. Dai,M., Wang,P., Boyd,A.D., Kostov,G., Athey,B., Jones,E.G., Bunney,W.E., Myers,R.M., Speed,T.P., Akil,H. *et al.* (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.*, **33**, e175.

24. Smyth,G.K. (2005) Limma: linear models for microarray data. In: Gentleman,R., Carey,V., Dudoit,S., Irizarry,R. and Huber,W. (eds), *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. Springer, New York, USA, Chapter 23, pp. 397–420.

25. Salmon,P. and Trono,D. (2006) Production and titration of lentiviral vectors. *Curr. Protoc. Neurosci.*, **37**, 4.21.1–4.21.24.

26. Jordan,J.J., Menendez,D., Inga,A., Noureddine,M., Bell,D.A. and Resnick,M.A. (2008) Noncanonical DNA motifs as transactivation targets by wild type and mutant p53. *PLoS Genet.*, **4**, e1000104.

27. Contente,A., Dittmer,A., Koch,M.C., Roth,J. and Dobbelstein,M. (2002) A polymorphic microsatellite that mediates induction of *PIG3* by p53. *Nat. Genet.*, **30**, 315–320.

28. Wang,B., Xiao,Z., Ko,H.L. and Ren,E.C. (2010) The p53 response element and transcriptional repression. *Cell Cycle*, **9**, 870–879.

29. Wang,B., Xiao,Z. and Ren,E.C. (2009) Redefining the p53 response element. *Proc. Natl Acad. Sci. USA*, **106**, 14373–14378.

30. Asthagiri,A.R., Parry,D.M., Butman,J.A., Kim,H.J., Tsilou,E.T., Zhuang,Z. and Lonser,R.R. (2009) Neurofibromatosis type 2. *Lancet*, **373**, 1974–1986.

31. Espinosa,J.M., Verdun,R.E. and Emerson,B.M. (2003) p53 functions through stress- and promoter-specific recruitment of transcription initiation components before and after DNA damage. *Mol. Cell*, **12**, 1015–1027.

32. Shaked,H., Shiff,I., Kott-Gutkowski,M., Siegfried,Z., Haupt,Y. and Simon,I. (2008) Chromatin immunoprecipitation-on-chip reveals stress-dependent p53 occupancy in primary normal cells but not in established cell lines. *Cancer Res.*, **68**, 9671–9677.

33. Lee,K.H., Li,M., Michalowski,A.M., Zhang,X., Liao,H., Chen,L., Xu,Y., Wu,X. and Huang,J. (2010) A genomewide study identifies the Wnt signaling pathway as a major target of p53 in murine embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **107**, 69–74.

34. Li,M., He,Y., Dubois,W., Wu,X., Shi,J. and Huang,J. (2012) Distinct regulatory mechanisms and functions for p53-activated and p53-repressed DNA damage response genes in embryonic stem cells. *Mol. Cell*, **46**, 30–42.

35. Hockley,S.L., Arlt,V.M., Brewer,D., Giddings,I. and Phillips,D.H. (2006) Time- and concentration-dependent changes in gene expression induced by benzo(a)pyrene in two human cell lines, MCF-7 and HepG2. *BMC Genomics*, **7**, 260.

36. Zhu,H., Smith,C., Ansah,C. and Gooderham,N.J. (2005) Responses of genes involved in cell cycle control to diverse DNA damaging chemicals in human lung adenocarcinoma A549 cells. *Cancer Cell Int.*, **5**, 28.

37. Krieg,A.J., Hammond,E.M. and Giaccia,A.J. (2006) Functional analysis of p53 binding under differential stresses. *Mol. Cell. Biol.*, **26**, 7030–7045.

38. Millau,J.F., Bastien,N., Bouchard,É.F. and Drouin,R. (2009) p53 Pre- and post-binding event theories revisited: stresses reveal specific and dynamic p53-binding patterns on the *p21* gene promoter. *Cancer Res.*, **69**, 8463–8471.

39. Staib,F., Robles,A.I., Varticovski,L., Wang,X.W., Zeeberg,B.R., Sirotin,M., Zhurkin,V.B., Hofseth,L.J., Hussain,S.P., Weinstein,J.N. *et al.* (2005) The p53 tumor suppressor network is a key responder to microenvironmental components of chronic inflammatory stress. *Cancer Res.*, **65**, 10255–10264.

40. Zhao,R., Gish,K., Murphy,M., Yin,Y., Notterman,D., Hoffman,W.H., Tom,E., Mack,D.H. and Levine,A.J. (2000) Analysis of p53-regulated gene expression patterns using oligonucleotide arrays. *Genes Dev.*, **14**, 981–993.

41. Li,T., Kon,N., Jiang,L., Tan,M., Ludwig,T., Zhao,Y., Baer,R. and Gu,W. (2012) Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. *Cell*, **149**, 1269–1283.

42. Vassilev,L.T. (2007) MDM2 inhibitors for cancer therapy. *Trends Mol. Med.*, **13**, 23–31.

43. Saito,S., Yamaguchi,H., Higashimoto,Y., Chao,C., Xu,Y., Fornace,A.J. Jr, Appella,E. and Anderson,C.W. (2003) Phosphorylation site interdependence of human p53 post-translational modifications in response to stress. *J. Biol. Chem.*, **278**, 37536–37544.

44. Vrba,L., Junk,D.J., Novak,P. and Futscher,B.W. (2008) p53 induces distinct epigenetic states at its direct target promoters. *BMC Genomics*, **9**, 486.

45. Kaneshiro,K., Tsutsumi,S., Tsuji,S., Shirahige,K. and Aburatani,H. (2007) An integrated map of p53-binding sites and histone modification in the human ENCODE regions. *Genomics*, **89**, 178–188.

46. Veprintsev,D.B. and Fersht,A.R. (2008) Algorithm for prediction of tumour suppressor p53 affinity for binding sites in DNA. *Nucleic Acids Res.*, **36**, 1589–1598.

47. Weinberg,R.L., Veprintsev,D.B., Bycroft,M. and Fersht,A.R. (2005) Comparative binding of p53 to its promoter and DNA recognition elements. *J. Mol. Biol.*, **348**, 589–596.

48. Noureddine,M.A., Menendez,D., Campbell,M.R., Bandele,O.J., Horvath,M.M., Wang,X., Pittman,G.S., Chorley,B.N., Resnick,M.A. and Bell,D.A. (2009) Probing the functional impact of sequence variation on p53-DNA interactions using a novel microsphere assay for protein-DNA binding with human cell extracts. *PLoS Genet.*, **5**, e1000462.

49. Menendez,D., Inga,A. and Resnick,M.A. (2010) Estrogen receptor acting *in cis* enhances WT and mutant p53 transactivation at

canonical and noncanonical p53 target sequences. *Proc. Natl Acad. Sci. USA*, **107**, 1500–1505.

50. Kitayner,M., Rozenberg,H., Rohs,R., Suad,O., Rabinovich,D., Honig,B. and Shakked,Z. (2010) Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nat. Struct. Mol. Biol.*, **17**, 423–429.

51. Kitayner,M., Rozenberg,H., Kessler,N., Rabinovich,D., Shaulov,L., Haran,T.E. and Shakked,Z. (2006) Structural basis of DNA recognition by p53 tetramers. *Mol. Cell*, **22**, 741–753.

52. Cai,B.H., Chao,C.F., Lu,M.H., Lin,H.C. and Chen,J.Y. (2012) A half-site of the p53-binding site on the keratin 14 promoter is specifically activated by p63. *J. Biochem.*, **152**, 99–110.

53. Chang,Z., Guo,C.L., Ahronowitz,I., Stemmer-Rachamimov,A.O., MacCollin,M. and Nunes,F.P. (2009) A role for the p53 pathway in the pathology of meningiomas with *NF2* loss. *J. Neurooncol*, **91**, 265–270.

54. Joachim,T., Ram,Z., Rappaport,Z.H., Simon,M., Schramm,J., Wiestler,O.D. and von Deimling,A. (2001) Comparative analysis of the *NF2*, *TP53*, *PTEN*, *KRAS*, *NRAS* and *HRAS* genes in sporadic and radiation-induced human meningiomas. *Int. J. Cancer*, **94**, 218–221.

55. Menon,A.G., Anderson,K.M., Riccardi,V.M., Chung,R.Y., Whaley,J.M., Yandell,D.W., Farmer,G.E., Freiman,R.N., Lee,J.K., Li,F.P. *et al.* (1990) Chromosome 17p deletions and p53 gene mutations associated with the formation of malignant neurofibrosarcomas in von Recklinghausen neurofibromatosis. *Proc. Natl Acad. Sci. USA*, **87**, 5435–5439.

56. Robanus-Maandag,E., Giovannini,M., van der Valk,M., Niwa-Kawakita,M., Abramowski,V., Antonescu,C., Thomas,G. and Berns,A. (2004) Synergy of *Nf2* and *p53* mutations in development of malignant tumours of neural crest origin. *Oncogene*, **23**, 6541–6547.

57. Akhoondi,S., Sun,D., von der Lehr,N., Apostolidou,S., Klotz,K., Maljukova,A., Cepeda,D., Fiegl,H., Dafou,D., Marth,C. *et al.* (2007) *FBXW7/hCDC4* is a general tumor suppressor in human cancer. *Cancer Res.*, **67**, 9006–9012.

58. Mao,J.H., Perez-Losada,J., Wu,D., Delrosario,R., Tsunematsu,R., Nakayama,K.I., Brown,K., Bryson,S. and Balmain,A. (2004) *Fbxw7/Cdc4* is a p53-dependent, haploinsufficient tumour suppressor gene. *Nature*, **432**, 775–779.

59. Grim,J.E., Knoblaugh,S.E., Guthrie,K.A., Hagar,A., Swanger,J., Hespelt,J., Delrow,J.J., Small,T., Grady,W.M., Nakayama,K.I. *et al.* (2012) Fbw7 and p53 cooperatively suppress advanced and chromosomally unstable intestinal cancer. *Mol. Cell. Biol.*, **32**, 2160–2167.

60. Yokobori,T., Mimori,K., Iwatsuki,M., Ishii,H., Onoyama,I., Fukagawa,T., Kuwano,H., Nakayama,K.I. and Mori,M. (2009) p53-Altered *FBXW7* expression determines poor prognosis in gastric cancer cases. *Cancer Res.*, **69**, 3788–3794.

61. Kelsell,D.P., Norgett,E.E., Unsworth,H., Teh,M.T., Cullup,T., Mein,C.A., Dopping-Hepenstal,P.J., Dale,B.A., Tadini,G., Fleckman,P. *et al.* (2005) Mutations in *ABCA12* underlie the severe congenital skin disease harlequin ichthyosis. *Am. J. Hum. Genet.*, **76**, 794–803.

62. Lefèvre,C., Bouadjar,B., Karaduman,A., Jobard,F., Saker,S., Özguc,M., Lathrop,M., Prud'homme,J.F. and Fischer,J. (2004) Mutations in *ichthyin* a new gene on chromosome 5q33 in a new form of autosomal recessive congenital ichthyosis. *Hum. Mol. Genet.*, **13**, 2473–2482.

63. Forbes,S.A., Bindal,N., Bamford,S., Cole,C., Kok,C.Y., Beare,D., Jia,M., Shepherd,R., Leung,K., Menzies,A. *et al.* (2011) COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.*, **39**, D945–D950.

64. Stevens,K.N., Wang,X., Fredericksen,Z., Pankratz,V.S., Cerhan,J., Vachon,C.M., Olson,J.E. and Couch,F.J. (2011) Evaluation of associations between common variation in mitotic regulatory pathways and risk of overall and high grade breast cancer. *Breast Cancer Res. Treat.*, **129**, 617–622.

65. Walker,M.M., Ellis,S.M., Auza,M.J., Patel,A. and Clark,P. (2008) The intercellular adhesion molecule, cadherin-10, is a marker for human prostate luminal epithelial cells that is not expressed in prostate cancer. *Mod. Pathol.*, **21**, 85–95.

66. Kober,P., Bujko,M., Olędzki,J., Tysarowski,A. and Siedlecki,J.A. (2011) Methyl-CpG binding column-based identification of nine genes hypermethylated in colorectal cancer. *Mol. Carcinog.*, **50**, 846–856.

67. Fu,D.Y., Wang,Z.M., Wang,B.L., Chen,L., Yang,W.T., Shen,Z.Z., Huang,W. and Shao,Z.M. (2010) Frequent epigenetic inactivation of the receptor tyrosine kinase *EphA5* by promoter methylation in human breast cancer. *Hum. Pathol.*, **41**, 48–58.

68. Sanders,S.J., Murtha,M.T., Gupta,A.R., Murdoch,J.D., Raubeson,M.J., Willsey,A.J., Ercan-Sencicek,A.G., DiLullo,N.M., Parikshak,N.N., Stein,J.L. *et al.* (2012) De novo mutations revealed by wholeexome sequencing are strongly associated with autism. *Nature*, **485**, 237–241.