Human single-nucleotide polymorphisms alter p53 sequence-specific binding at gene regulatory elements

Omari J. Bandele, Xuting Wang, Michelle R. Campbell, Gary S. Pittman and Douglas A. Bell*

Environmental Genomics Group, Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, USA

Received May 18, 2010; Accepted August 10, 2010

ABSTRACT

p53 coordinates the expression of an intricate network of genes in response to stress signals. Sequence-specific DNA binding is essential for p53-mediated tumor suppression. We evaluated the impact of single-nucleotide polymorphisms (SNPs) in p53 response elements (p53RE) on DNA binding and gene expression in response to DNA damage. Using a bioinformatics approach based on incorporating p53 binding strength into a position weight matrix, we selected 32 SNPs in putative and validated p53REs. The microsphere assay for protein–DNA binding (MAPD) and allele-specific expression analysis was employed to assess the impact of SNPs on p53-DNA binding and gene expression, respectively. Comparing activated p53 binding in nuclear extracts from doxorubicin- or ionizing radiation (IR)-treated human cells, we observed little difference in binding profiles. Significant p53 binding was observed for most polymorphic REs and several displayed binding comparable to the p21 RE. SNP alleles predicted to lower p53 binding indeed reduced binding in 25 of the 32 sequences. Chromatin immunoprecipitation-sequencing in lymphoblastoid cells confirmed p53 binding to seven polymorphic p53 REs in response to doxorubicin. In addition, five polymorphisms were associated with altered gene expression following doxorubicin treatment. Our findings demonstrate an effective strategy to identify and evaluate SNPs that may alter p53-mediated stress responses.

INTRODUCTION

Although most of the 14.5 million validated human single-nucleotide polymorphisms (SNPs) listed in the National Center for Biotechnology Information (NCBI) dbSNP database (Build 131) are likely nonfunctional, some can alter cellular responses to a variety of perturbations such as DNA damage, and thereby increase susceptibility to diseases including cancer (1,2). Coding SNPs alter amino acid sequences and can modify structural and biological properties of encoded proteins. Alternatively, regulatory SNPs affect non-coding gene regions such as introns, enhancers, silencers and promoters. These non-coding sequence variants can modify transcription factor binding sites (3) or generate novel binding sites (4–6), which may modulate gene expression in an allele-specific manner (7–11). Nevertheless, the impact of non-coding regulatory SNPs on transcription factor binding and gene expression has been relatively unexplored.

The p53 tumor suppressor is a prominent transcription factor that regulates the expression of an intricate network of genes in response to DNA damage and other stress signals (12). This protein coordinates transcriptional programs primarily by binding to specific DNA sequences, known as p53 response elements (p53REs). The p53 consensus DNA-binding sequence is composed of two decameric half-sites, RRRWWGYYY (where W = A or T, R = purine and Y = pyrimidine), separated by a 0–13-bp spacer (13–15). Importantly, sequence-specific DNA binding is essential for p53 to exert its tumor suppressive activity (16–18).

Although p53 plays a significant role in protecting cells against DNA damage (e.g. carcinogenesis), the biological impact of SNPs in p53 binding sites is largely unknown. Experimental studies have identified over 150 functional
human p53REs (12) and genome-wide analyses have proposed many putative sites where the p53 protein may bind (19,20). It is likely that many SNPs occur within p53 binding sites, and we have previously identified several polymorphisms in putative p53REs that alter p53-DNA binding and transactivation capacity in an allele-specific manner (21,22). These findings suggest that polymorphisms in non-coding gene regulatory regions can modify transcriptional responses and may represent genetic risk factors for disease susceptibility. Using a previously described unique bioinformatics approach (22) that incorporates experimentally based p53 binding data, we conducted a comprehensive search for functional regulatory SNPs in experimentally validated (e.g. bona fide) and putative p53REs that might affect binding. This *in silico* strategy provides quantitative predictions of allele-specific binding and was used to identify and select candidate SNPs that were expected to alter p53 binding and gene expression in response to DNA-damaging agents.

The microsphere assay for protein–DNA binding (MAPD) (22,23) was used to evaluate allele-specific p53 binding to selected sequences. This assay utilizes fluorescent microspheres coated with oligonucleotides that contain specific binding sequences. Microsphere-oligonucleotide targets were incubated in nuclear extracts from cells that contain endogenous levels of ‘activated’ p53, and DNA-bound p53 was detected using a flow cytometer. An advantage of this assay is its ability to produce quantitative and biologically relevant protein–DNA binding measurements.

In this study, we measured the ability of p53 in nuclear extracts from human lymphoblastoid and U2OS osteosarcoma cells treated with doxorubicin (DOXO) or ionizing radiation (IR) to bind six *bona fide* and 26 putative p53REs containing SNPs. Most elements displayed significant binding above the negative control, and several putative elements were bound at levels similar to *bona fide* p53 binding sites. For most tested sequences, the alleles that were predicted to lower p53 binding indeed reduced binding, and the observed effects were independent of cell type or treatment. Chromatin immunoprecipitation-sequencing (ChIP-Seq) studies in cultured human cells confirmed that activated p53 bound to several sequences containing candidate SNPs. Furthermore, gene expression analysis demonstrated that several SNPs were associated with reduced gene transcription. These results illustrate that our bioinformatic approach combined with the MAPD binding assay and gene expression analysis is an effective strategy to identify functional SNPs that modify p53-mediated responses to stress signals and to discover novel p53REs.

**MATERIALS AND METHODS**

**Cell culture**

HapMap (http://hapmap.ncbi.nlm.nih.gov/) CEU human lymphoblastoid cells (Coriell Cell Repositories, Camden, NJ, USA) were grown in RPMI 1640 medium supplemented with 15% heat-inactivated fetal bovine serum (Invitrogen, Carlsbad, CA, USA) and 1X antibiotics/antimycotics (Gibco, Carlsbad, CA, USA). Human U2OS osteosarcoma cells (HTB-96, ATCC, Manassas, VA, USA) were grown in McCoy’s 5A medium supplemented with 10% heat-inactivated fetal bovine serum and 1X antibiotics/antimycotics. Cells were incubated at 37 °C with 5% CO2. During experiments, cells were grown in petri dishes and treated with 0.3 or 0.6 μg/ml doxorubicin (Sigma, St. Louis, MO, USA) for 18 h or exposed to 15 Gy ionizing radiation using a Shepherd 137Cesium irradiator at a dose rate of 0.85 krad/min followed by a 5-h incubation at 37°C.

**Western blot**

Nuclear proteins were isolated from untreated and treated cells using the Protein Extract Kit (Active Motif, Carlsbad, CA, USA). Proteins were quantified using the Bradford assay and a HTS 7000™ Bio Assay Reader (Perkin Elmer, Waltham, MA USA). Equal amounts (10 μg) of nuclear proteins were electrophoresed on a NuPAGE® 4–12% BisTris gel (Invitrogen) and transferred to a 0.45-μM nitrocellulose membrane (Invitrogen). Membranes were probed with primary anti-p53 mouse monoclonal antibodies (DO-7; BD Biosciences, San Jose, CA, USA) or anti-β-actin mouse polyclonal antibodies (C-11; Sigma). Proteins were detected using goat anti-mouse HRP-conjugated secondary antibodies (Bio-Rad, Hercules, CA, USA) and the enhanced chemiluminescence (ECL) detection system (Amersham, Piscataway, NJ, USA).

**Identification of putative functional SNPs in p53REs**

By implementing a modification of the approach of Veprintsev and Fersht (24), we created a binding-based position weight matrix (BPWM) model for quantitatively predicting p53-DNA binding (22) and incorporated this into our previous SNP identification strategy (25,26). Briefly, a delta-binding matrix was generated by subtracting experimentally measured p53 binding values for 60 variants of the ConA sequence (a perfect match to the p53 consensus binding motif) from binding to the unmodified sequence. Equations (2), (A.1) and (A.2) from Veprintsev and Fersht (24) were applied to the delta-binding matrix to generate the BPWM model.

In our current study, the BPWM model was used to identify and score SNPs within the human genome that alter p53-DNA binding. Briefly, utilizing the delta-binding matrix, each allele was scored independently, and the difference between the two scores was used to estimate the SNP impact on p53-DNA binding. We identified SNPs in putative p53REs by searching the NCBI dbSNP database for polymorphisms in sequences that matched the p53-DNA binding motif. We identified SNPs in validated p53REs by mapping NCBI dbSNP database entries into a list of 150 experimentally discovered p53REs.

Several SNPs in putative p53REs (rs2838769, rs3806624, rs1658728, rs3761624 and rs1465952) have previously been tested in our group using MAPD (22) and
luciferase expression assays (21). These SNPs were included in this study to evaluate their effects in different cell types and under different experimental conditions.

**Microsphere assay for protein–DNA binding**

Allele-specific p53 binding was evaluated using the MAPD assay, as previously described (22,23). Briefly, MicroPlex™-xTAG microspheres (Luminex, Austin, TX, USA) were hybridized to double-stranded oligonucleotides (Invitrogen) containing unique p53 binding sites and a biotin moiety on the 5′-end of the reverse strand. Hybridization reactions were carried out in a polymerase chain reaction (PCR) thermocycler. Microsphere-oligonucleotide targets were pooled and washed in a MultiScreen® HTS-BV filter plate (Millipore, Bedford, MA, USA) with Assay Buffer A (Cytokine Reagent Kit, Bio-Rad).

Sequence-specific p53 binding was examined 1 h after incubating targets in reaction mixtures that contained non-competing double-stranded oligonucleotides (TransAm p53 kit, Active Motif) and nuclear extracts from doxorubicin- or IR-exposed human cells. Negative and positive control sequences were examined in each reaction. Nuclear extracts from untreated cells served as an additional negative control. Microspheres were washed and incubated for 30 min with primary anti-p53 monoclonal antibodies (DO-7) followed by a 30-min incubation with phycoerythrin-conjugated secondary antibodies. p53 DNA binding was measured by flow cytometric analysis using a Bio-Plex® 200 System (Bio-Rad).

**p53 ChIP-Seq**

Chromatin immunoprecipitation methods and buffers were from the Agilent Mammalian ChIP-on Chip protocol (Agilent, Version 10.0, May 2008). Briefly, 10^6 DOXO or untreated lymphoblastoid cells (GM06993 and GM11992) were formaldehyde crosslinked and lysed to obtain nuclear cell pellets. Pellets were resuspended, chromatin was sonicated into 200–500-bp fragments using a Misonix 3000 sonicator (30 pulses—30 s on followed by 20 s off—at 100% power) (Farmingdale, NY, USA), and the supernatant was collected after centrifugation. Using either mouse monoclonal antibody DO-7 (BD Biosciences) or a non-specific rabbit IgG (Invitrogen) conjugated to secondary Dynal magnetic beads (Invitrogen), 5 × 10^7 isolated nuclei were immunoprecipitated overnight (4°C), beads were washed, and protein–DNA crosslinks reversed overnight (65°C). After digestion of RNA and cellular protein, DNA was isolated using phenol chloroform extraction and a Qiaquick PCR Purification Kit (Qiagen, Germantown, MD, USA) according to the manufacturer’s instructions. Purified ChIP DNA was dried, resuspended in double-distilled H2O, and quantified using PicoGreen (Invitrogen). p53 ChIP DNA from the two cell lines were combined for each condition (DOXO-treated, untreated, and IgG ChIP).

The National Center for Genome Resources (Santa Fe, NM, USA) sequenced the immunoprecipitated samples on the Illumina Genome Analyzer II. Briefly, sequencing adapters were ligated to immunoprecipitated DNA and size-selected by gel electrophoresis (250 ± 25 bp). Fragments were PCR-amplified (18 cycles) and sequenced. Short reads were mapped to the human reference genome (Build 36.3) using Burrows-Wheeler Alignment (BWA) Tool (27). By default, BWA finds an alignment (ungapped and gapped) with maximum edit-distance 2 to the 36 bp query sequence, except for disallowing gaps close to the end of the query.

Uniquely mapped short reads from p53-immunoprecipitated DNA were used to identify regions of the genome with significant enrichment in p53-associated DNA sequences, hereafter referred to as ‘peaks’ owing to their appearance in genome-wide density plots. The peak detection was performed by QuEST 2.4 software (28) using a ‘transcription factor binding site’ setting (bandwidth of 30 bp, region size of 300 bp) and default stringency (corresponding to 50-fold ChIP to input enrichment for the regions, and 3-fold ChIP enrichment for extending the regions).

**Gene expression analysis**

To evaluate the allele-specific effects of SNPs on gene expression, 23 HapMap CEU human lymphoblastoid cell lines were cultured in triplicate and treated with 0.6 μg/ml doxorubicin for 18 h. RNA was isolated with RNeasy columns (Qiagen) and cDNA was generated using the First-Strand Synthesis system (Invitrogen). For each biological replicate, target genes were amplified in triplicate using SYBR primers designed to span exon junctions for each target gene (see Supplementary Table S2) and Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA, USA). Fluorescence intensity was measured using a PRISM HTS 7900 System (Applied Biosystems) and the initial fluorescence (R0, value) of each amplified sample was calculated using the method described by Peirson and colleagues (29). Target values were normalized to GAPDH mRNA measurements.

**RESULTS**

**In silico selection of candidate SNPs for functional evaluation**

Using a novel bioinformatics approach (Figure 1), we identified 6538 SNPs in the NCBI dbSNP (Build 129) database that were located in putative p53 binding sites and were predicted to alter p53–DNA interactions. We also identified 33 polymorphic sequences in experimentally validated p53 binding sites. After applying additional prioritization criteria, 32 SNPs were selected for functional evaluation. Six of these variants were located within 5 kb of the transcription start site of bona fide p53-responsive genes. SNP selection was based on the requirement that: (i) each site contained all four conserved core C and G nucleotides in at least one allele; (ii) each site contained at least three of the four consensus ‘W’ (A or T) nucleotides in at least one allele; (iii) each site contained a spacer ≤1-bp separating the two half-sites; (iv) the BPWM
score for the predicted strong allele was above 100 (an indicator of a moderate-to-high p53 binding sequence); and (v) the polymorphism was predicted to alter p53-DNA binding. The candidate SNPs and associated REs are described in Table 1.

MAPD evaluation of allele-specific p53 binding to polymorphic sequences

To overcome the limitations of traditional methods used to quantify protein–DNA binding, we have developed a microsphere-based DNA binding assay to measure p53-RE interactions (22). As seen in Figure 2, allele-specific binding was observed after incubating bead-oligonucleotide targets in nuclear extracts from human lymphoblastoid cells that were either untreated or treated with doxorubicin (DOXO). p53 binding levels in untreated samples were <2% of levels observed in treated samples. In addition, allele-specific differences in binding were not observed in untreated cells (data not shown).

In contrast, p53 in extracts from DOXO-treated cells bound to the predicted strong allele of 29 REs at levels that were above the negative control sequence (WRNC). Moreover, the tumor suppressor bound several validated REs at levels that were comparable to or greater than those observed for the positive control (a bona fide p53RE ~2 kb upstream of the p21 gene transcription start site) (see EDN2, RRM2B and TRIM22). Some putative REs bound p53 at levels similar to validated p53 binding sites (see ADARB1, EOMES, PSMA6, TNFSF14, TPP2 and UBA3). Compared to the predicted strong alleles, binding to the predicted weak alleles was significantly reduced for ~80% of the elements (P < 0.05, t-test). For 10 of these variants, the SNP diminished p53 binding by ≥50%.

Furthermore, the polymorphic sequences associated with FUCAI, PMAIP1, RRM1 and TLR8 nearly abolished the ability of p53 to recognize the binding sites. It should also be noted that our bioinformatic calculations were generally accurate for predicting p53 binding to REs with high scores (>100) and for estimating the SNP effect on binding.

ChIP-Seq confirms p53 binding to sequences containing SNPs

To date, only one p53 ChIP-seq study has been published, which was carried out in a colon cancer cell line treated with 5-fluorouracil (5-FU) (20). We applied ChIP-Seq to HapMap lymphoblastoid cells to determine whether DOXO induces p53 to bind to our selected sequences. Detailed information on the ChIP-Seq experiment and data analysis will be reported separately. Here, we only show limited data corresponding to the genomic regions selected for this study. After mapping sequenced p53-ChIP fragments to the human reference genome (Build 36.3), we observed that p53-bound sequences were enriched for seven of the 32 tested binding sites (Figure 3). It should be noted that four of these p53-bound sites correspond to sequences that had not been previously confirmed as p53 binding sites. Additional sequence analysis revealed that the candidate SNPs occurred within the p53-bound sites. These findings combined with our MAPD binding data suggest that a number of the bioinformatics-selected sequences are indeed bound by p53 in cultured human cells treated with DOXO and
that the candidate SNPs have the potential to modify p53 binding in vivo.

Doxorubicin induces allele-specific transcription in human cells

To determine if the selected SNPs alter gene transcription, we performed RT-qPCR on RNA isolated from untreated and DOXO-treated HapMap CEU cell lines that were homozygous for either the predicted strong or weak allele. Relative to untreated cells, nine of 11 analyzed genes were induced in response to DOXO treatment (Figure 4). Allele-specific differences in gene expression were observed for ADAR1, PSM4A6, TNFSF14, TRIM22 and UBA3 where the weak allele displayed lower levels of expression ($P < 0.05$, t-test). Expression levels for TPP2 and SFRS12 were not induced in response to DOXO, which suggests that they are not regulated by the tumor suppressor. The other tested genes had modest allelic differences in expression. These findings suggest that SNPs in p53 binding sites can alter gene expression in response to DNA-damaging agents.

Allele-specific p53 binding is independent of cell type-specific effects

p53 is a versatile transcription factor with distinct roles that are influenced by cell type (30) and treatment (31,32). To evaluate cell-type specific contributions to p53-DNA binding, p53 binding to a subset of REs in nuclear extracts from DOXO-treated U2OS osteosarcoma cells was compared to results observed in lymphoblastoid cells. REs that displayed either weak p53 binding (<10% relative to the positive control) or no allele-specific binding changes in both cell lines were excluded from further analysis. As seen in Figure 5A, DOXO-activated p53 from U2OS cells bound each sequence at levels that were comparable to DOXO-activated p53 from lymphoblastoid cells (see Figure 2). A linear regression analysis of binding in lymphoblastoid versus U2OS extracts produced $r^2$ values of 0.92 and 0.94 for strong and weak alleles, respectively (data not shown). Linear regression analysis comparing SNP-induced binding changes in both cell lines demonstrated that the SNP effect on binding were also similar ($r^2 = 0.87$, Figure 5B).

**Table 1. Candidate p53RE SNPs**

<table>
<thead>
<tr>
<th>SNP rs#</th>
<th>Associated gene</th>
<th>Sequence alignment (bpwm)</th>
<th>Strong allele (BPWM)</th>
<th>Weak allele (BPWM)</th>
<th>ΔBPWM</th>
<th>Cellular process</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2837869</td>
<td>ADAR1</td>
<td>GGCAAGGTTGAGATTTCAC</td>
<td>G (131.3)</td>
<td>A (100.7)</td>
<td>12.6</td>
<td>Nucleic acid metabolism</td>
</tr>
<tr>
<td>rs2033654</td>
<td>ADCY3</td>
<td>GGAAGTTGCCTAACATGAGCA</td>
<td>G (108.3)</td>
<td>T (93.9)</td>
<td>14.4</td>
<td>Intracellular signaling</td>
</tr>
<tr>
<td>rs1668727</td>
<td>ARHGEF7</td>
<td>AACAGTATGCACTGGTCTT</td>
<td>G (124.2)</td>
<td>T (121.8)</td>
<td>2.4</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>rs3086624</td>
<td>EOMES</td>
<td>CACTTTGCCAGATTCTTTGCC</td>
<td>T (117.4)</td>
<td>C (105.9)</td>
<td>11.5</td>
<td>Development</td>
</tr>
<tr>
<td>rs12105811</td>
<td>CFLAR</td>
<td>GATCCGGTTTGCTGTTT</td>
<td>A (109.5)</td>
<td>G (105.2)</td>
<td>4.3</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>rs10917431</td>
<td>FUC1</td>
<td>GGGGCTCTCTACCTGCTT</td>
<td>A (123.8)</td>
<td>C (112.9)</td>
<td>10.9</td>
<td>Carbohydrate metabolism</td>
</tr>
<tr>
<td>rs9282701</td>
<td>HMOX1</td>
<td>GAGCGAAGCAAGACAGGCC</td>
<td>C (103.6)</td>
<td>T (103.0)</td>
<td>0.6</td>
<td>Oxidative stress response</td>
</tr>
<tr>
<td>rs11855354</td>
<td>IDH3A</td>
<td>TTTCCTTGTGCGGACAGCA</td>
<td>G (104.4)</td>
<td>A (90.0)</td>
<td>14.4</td>
<td>Carbohydrate metabolism</td>
</tr>
<tr>
<td>rs3755272</td>
<td>IL18R1</td>
<td>ATGCTATGCTGCCCTTGGT</td>
<td>A (124.0)</td>
<td>G (119.7)</td>
<td>4.3</td>
<td>Immune response</td>
</tr>
<tr>
<td>rs2410545</td>
<td>NAT1</td>
<td>CCTCCTGCACGAGCTGGAG</td>
<td>A (122.8)</td>
<td>G (111.3)</td>
<td>11.5</td>
<td>Xenobiotic metabolism</td>
</tr>
<tr>
<td>rs1946104</td>
<td>NOSM7</td>
<td>TACAGACACCACTGCTT</td>
<td>T (123.0)</td>
<td>C (118.7)</td>
<td>4.3</td>
<td>Nucleic acid metabolism</td>
</tr>
<tr>
<td>rs12409754</td>
<td>PGD</td>
<td>GGGGATGTTGCGGATGCTT</td>
<td>T (118.2)</td>
<td>C (97.7)</td>
<td>1.4</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>rs7240884</td>
<td>PMAIP1</td>
<td>CACATATTGATGCTTGTGAT</td>
<td>G (112.1)</td>
<td>A (97.7)</td>
<td>14.4</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>rs1048990</td>
<td>PSMA6</td>
<td>TGCTTTTATACCCATGCTC</td>
<td>G (112.4)</td>
<td>C (97.8)</td>
<td>14.6</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>rs1003619</td>
<td>RANGAP1</td>
<td>CCACAGCCACAAAGGCGG</td>
<td>G (112.1)</td>
<td>C (97.7)</td>
<td>14.4</td>
<td>Cell signaling</td>
</tr>
<tr>
<td>rs6017447</td>
<td>RP2N2</td>
<td>CACCATGTCGGCCTGTTTT</td>
<td>C (119.6)</td>
<td>G (115.8)</td>
<td>3.8</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>rs1465952</td>
<td>RRM1</td>
<td>GGGATGCTACATGGTTTTT</td>
<td>C (120.6)</td>
<td>T (108.1)</td>
<td>12.5</td>
<td>Nucleic acid metabolism</td>
</tr>
<tr>
<td>rs27078</td>
<td>SFRS12</td>
<td>TGGCATGCTCTCAGTGGAT</td>
<td>C (125.6)</td>
<td>G (111.0)</td>
<td>14.6</td>
<td>Nucleic acid metabolism</td>
</tr>
<tr>
<td>rs50808</td>
<td>SLC22A12</td>
<td>CCACATGCTGACATGCTT</td>
<td>A (136.2)</td>
<td>G (132.0)</td>
<td>4.2</td>
<td>Ion transport</td>
</tr>
<tr>
<td>rs1678199</td>
<td>SP100</td>
<td>TGGGCACTTATGGGCGACC</td>
<td>A (125.4)</td>
<td>C (124.8)</td>
<td>0.6</td>
<td>Immune response</td>
</tr>
<tr>
<td>rs3761624</td>
<td>TLR8</td>
<td>AGGGAGATGAAACTTCCA</td>
<td>G (111.9)</td>
<td>A (101.1)</td>
<td>10.8</td>
<td>Immune response</td>
</tr>
<tr>
<td>rs1077667</td>
<td>TNFSF14</td>
<td>GGGCTGACAGACGCTGGCCA</td>
<td>A (123.5)</td>
<td>G (112.0)</td>
<td>11.5</td>
<td>Immune response</td>
</tr>
<tr>
<td>rs3783231</td>
<td>TPP2</td>
<td>CCAGAGCTAAAGCAGGTTT</td>
<td>T (126.4)</td>
<td>C (121.8)</td>
<td>4.6</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>rs10455025</td>
<td>TSLP</td>
<td>GAAAAGCTGAACATGAAA</td>
<td>A (123.8)</td>
<td>C (122.6)</td>
<td>1.2</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>rs3853154</td>
<td>UBA3</td>
<td>AGACATGATGACGCTC</td>
<td>C (118.1)</td>
<td>A (114.6)</td>
<td>3.5</td>
<td>Protein metabolism</td>
</tr>
<tr>
<td>rs11248877</td>
<td>UNKL</td>
<td>GGGGCTGTTGACATGCTCT</td>
<td>A (119.2)</td>
<td>G (115.0)</td>
<td>4.2</td>
<td>Nucleic acid metabolism</td>
</tr>
</tbody>
</table>
To further evaluate cell type-specific contributions to p53-RE binding, we compared binding in nuclear extracts from IR-exposed lymphoblastoid and U2OS cells. As seen in Supplementary Figure S1A (lymphoblastoid cells) and B (U2OS cells), IR-activated p53 binding levels were similar in both extracts ($r^2 = 0.95$ and 0.97 for strong and weak alleles, respectively; data not shown). The SNP effects on p53 binding also were similar in both cell types ($r^2 = 0.90$, Supplementary Figure S1C). These results demonstrate that the effects of the polymorphisms on p53-DNA binding are independent of cell-type specific differences between the two cell lines.

Allele-specific p53 binding is independent of treatment-specific effects

To determine if treatment-specific effects alter p53–DNA interactions, we compared binding in nuclear extracts from IR-exposed lymphoblastoid cells to results observed in extracts from DOXO-treated cells. In Figure 6A, the pattern of IR-induced p53 binding to the predicted strong alleles was similar to the pattern induced by DOXO (see Figure 2). As seen with DOXO, reduced binding to the weak alleles was also observed in extracts from IR-exposed cells ($P < 0.05$, t-test). Moreover, a strong correlation was shown for SNP-induced changes on p53 binding between DOXO and IR treatment ($r^2 = 0.86$, Figure 6B).

To further determine if treatment-specific effects modify p53-RE binding, we compared p53 binding in nuclear extracts from DOXO- and IR-exposed U2OS cells. As seen in Supplementary Figure S2, the SNP effects on p53 binding were also similar in response to both treatments ($r^2 = 0.73$). These findings indicate that the impact of sequence variation on p53-RE binding is mostly unaffected by treatment-specific effects induced by these DNA-damaging agents.
SNP sequence position influences the polymorphic effects on p53 binding

Although base changes involving the highly conserved C and G nucleotides at positions 4 and 7 in each p53RE half-site significantly reduce p53 binding, most functional REs deviate from the consensus binding sequence (22,33,34). Variations in p53RE sequences and p53 protein levels contribute greatly to gene transactivation mediated by the tumor suppressor (23,33,35). Understanding this relationship is important for determining how p53 influences cellular responses to stress. To further examine the effects of polymorphic sequences on p53-DNA binding, we evaluated ADARB1 and RRM2B RE allele pairs in the presence of increasing amounts (0–4 μg) of nuclear extracts from DOXO-treated lymphoblastoid cells. The RE associated with ADARB1 contains a G-to-A change at position 7 in the second half-site, while the RE associated with RRM2B has a C-to-G change adjacent to position 4 and 7. As seen in Figure 7A, modification of the preferred G in the ADARB1 RE dramatically reduced p53 binding. For example, at 4 μg of nuclear extracts, the predicted weak allele reduced p53 binding by 3.5-fold relative to the predicted strong allele.

The base change in the RRM2B RE had modest effects on p53 binding compared to the strong allele (see Figure 7B). At 4 μg of nuclear extracts, p53 binding was 0.44 and 0.39 for the predicted strong and weak allele,
respectively. These results suggest that polymorphisms that alter the highly favored bases at position 4 and 7 in p53RE half-sites can have significant effects on p53-mediated responses to DNA damage, whereas other changes may be better tolerated.

**DISCUSSION**

Non-coding regulatory SNPs in transcription factor binding sites can have significant biological consequences (1). We carried out a comprehensive genome-wide bioinformatics search using a unique prediction model (BPWM), to identify human SNPs in validated and putative p53-targeted sequences that were predicted to alter p53 binding following exposure to DNA-damaging agents. The selected SNPs were evaluated using the MAPD binding assay and gene expression analysis to determine the functional impact of each variant. We employed ChIP-Seq analysis to confirm p53 sequence-specific binding in cultured human cells in response to doxorubicin, a DNA-damaging agent. To our knowledge, this represents the first study that examines properties of SNPs that co-locate with ChIP-seq peaks. The allele-specific differences observed for p53 target genes following exposure to doxorubicin support the hypothesis that genetic variation in non-coding regulatory sequences may impact risk in environmentally induced diseases such as cancer (1). Our findings also demonstrate that this bioinformatics strategy can identify SNPs in gene regulatory regions that alter p53–DNA binding and gene expression. In addition, we believe regulatory SNPs that affect other transcription factors or DNA binding proteins can be identified and evaluated using a similar approach.

Many of the sequences we identified have not been experimentally validated as p53REs. Our results, using MAPD, gene expression and ChIP-seq methods, demonstrate that p53 binds putative polymorphic REs near ADARB1, HMOX1, PSMA6 and TNFSF14 in cultured human cells, which suggests that these genes are novel, previously unidentified components of the p53 transcriptional network. Several polymorphic sequences that were tested are associated with genes that have tumor suppressive activity, play roles in cell cycle regulation, or function in other cellular processes associated with p53-mediated signaling pathways. ADARB1 is a member of the adenosine deaminase family, which plays important roles during developmental processes. This RNA-editing enzyme is suggested to have negative effects on the development of...
**Figure 5.** Evaluation of cell type-specific effects on p53 binding to polymorphic alleles. (A) The MAPD assay was used to assess p53 binding to predicted strong (black bars) and weak (gray bars) alleles in nuclear extracts from DOXO-treated (0.6 µg/ml for 18 h) human U2OS cells. All other experimental conditions and data analysis was performed as described in Figure 2. Each allele pair displayed significant differences in binding ($P < 0.05$, t-test). Error bars represent the standard deviation of three independent experiments carried out in triplicate. (B) Regression analysis was used to compare the effect of each SNP on p53 binding in nuclear extracts from DOXO-treated human lymphoblastoid versus U2OS cells.

**Figure 6.** Evaluation of treatment-specific effects on p53 binding to polymorphic alleles. (A) The MAPD assay was used to assess p53 binding to predicted strong (black bars) and weak (gray bars) RE alleles in nuclear extracts from IR-treated (15 Gy followed by a 5-h recovery) human lymphoblastoid cells. All other experimental conditions and data analysis was performed as described in Figure 2. Each allele pair displayed significant differences in binding ($P < 0.05$, t-test). Error bars represent the standard deviation of three independent experiments carried out in triplicate. (B) Regression analysis was used to compare the effect of each SNP on p53 binding in nuclear extracts from human lymphoblastoid cells treated with DOXO versus IR.
In this work, we have used de novo approaches to identify and test polymorphic REs that alter p53–DNA binding. These SNPs may also alter cellular stress responses and increase disease susceptibility in vivo. Our results further illustrate the significance of the conserved C and G in the p53–DNA binding motif. We demonstrate that this observation is apparent in both moderate (ADARB1 and TLR8) and low (PMAIP1 and RRMI) affinity p53 binding sites, which suggests that sequence variations involving these nucleotides will have the most significant biological consequences. In contrast, SNPs that alter nucleotides outside of the conserved C and G may have little impact and produce no obvious phenotypic effects.

In this study, we identified SNPs that alter p53-RE binding and impact gene expression; however, the functional, evolutionary and risk implications of these effects are not always obvious. For example, we observe occasionally that the less common, minor allele is the predicted strong allele, displaying both stronger binding (e.g. rs1077667, TNFSF14; rs1048990, PSMAs6) and stronger transactivation while the common (i.e. ancestral allele, shared with other primates) allele displays weaker binding. This suggests that the recently created, stronger allele may have some selective advantage in human lineages that have it. However, selective advantage (positive selection) or disadvantage (purifying selection) is related to reproductive success, while cancer risk is related to somatic processes that impact a post-reproductive disease outcome. Although it is unclear whether the creation of these specific strong-binding alleles in the human population affects physiology, survival or risk, transcriptional pathways may evolve in this way (45–47). Menendez et al. (6) have suggested that an SNP may create novel binding sites and connect transcriptional pathways. The impact of RE polymorphisms is further complicated by the possibility that specific p53RE ‘WW’ dinucleotide combinations can influence transcriptional pathways. The impact of RE polymorphisms is further complicated by the possibility that specific p53RE ‘WW’ dinucleotide combinations can influence transcriptional pathways. Thus, an SNP that occurs in a RE from which p53 represses transcription could decrease p53 binding and thereby increase gene expression. This scenario may explain the increased expression of PMAIP1 in cell lines homozygous for the weak allele. Evaluating SNPs in p53REs that are predicted to repress transactivation would be of great interest.

The p53 protein plays key roles during the cellular stress response (12), which can be influenced by cell type (30). A number of reports have demonstrated that different stress-inducing agents trigger a variety of p53 post-translational modifications (49,50) and differential transactivation of target genes (51,52). Two recent publications disagree with regard to whether different agents produce different p53 occupancy profiles at native binding sites in cells. Shaked et al. (53), using ChiP-on-chip, observed that global patterns of p53 occupancy differed very little among gamma irradiated, DOXO, 5-FU or UV treated HCT116 and U2OS cell lines. Millau et al. (54) using a more sensitive technique, observed unique p53 occupancy patterns over time at different p21 response elements following treatment with 5-FU, Nutlin-3, UV and gamma...
radiation. The MAPD assay evaluates the p53 binding step and provides high resolution and discrimination of small binding differences. The present findings using DOXO and IR in lymphoblastoid and U2OS osteosarcoma cells clearly demonstrate that very similar p53 binding patterns are produced from p53 activated under all experimental conditions tested. Thus, it is likely that other features in the transcription process, such as the chromatin environment or stress-specific cofactors, may contribute to cell type- and treatment-specific p53 binding and subsequent transactivation in cells (51,52). In addition, we demonstrate for the first time that individual SNPs produced similar effects on p53–DNA binding regardless of the cell type or stimulus used for activation.

The present study suggests that sequence variation in p53 binding sites can impact binding and potentially modify p53-mediated responses to DNA-damaging agents. The MAPD binding assay is a useful way to estimate the functional impact of nucleotide changes in target binding sites and could be applied to other transcription factors. Developing quantitative binding data for many regulatory elements may allow us to elucidate the impact of polymorphic variation on transcriptional networks.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We appreciate the support of the Intramural Research Program of the National Institute of Environmental Health Sciences, National Institutes of Health. We thank Drs. Maher Noureddine, Brian Chorley, Michael Resnick, Scott Auerbach and Daniel Menendez for helpful comments on this project and review of the article. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the article.

FUNDING
The Intramural Research Program of the National Institute of Environmental Health Sciences, National Institutes of Health (Z01-ES-100475-M-0001 and Z01-ES065079-15). Funding for open access charge: The Intramural Research Program, National Institute of Environmental Health Sciences.

Conflict of interest statement. None declared.

REFERENCES


