

Discovery of novel tumor suppressor p53 response elements using information theory

Ilya G. Lyakhov¹, Annangarachari Krishnamachari² and Thomas D. Schneider^{3,*}

¹Basic Research Program, SAIC-Frederick, Inc., NCI at Frederick, Frederick, MD, USA, ²Bioinformatics Centre, Jawaharlal Nehru University, New Delhi -110067, India and ³Center for Cancer Research Nanobiology Program, NCI at Frederick, Frederick, MD, USA

Received January 7, 2008; Revised March 31, 2008; Accepted April 1, 2008

ABSTRACT

An accurate method for locating genes under tumor suppressor p53 control that is based on a well-established mathematical theory and built using naturally occurring, experimentally proven p53 sites is essential in understanding the complete p53 network. We used a molecular information theory approach to create a flexible model for p53 binding. By searching around transcription start sites in human chromosomes 1 and 2, we predicted 16 novel p53 binding sites and experimentally demonstrated that 15 of the 16 (94%) sites were bound by p53. Some were also bound by the related proteins p63 and p73. Thirteen of the adjacent genes were controlled by at least one of the proteins. Eleven of the 16 sites (69%) had not been identified previously. This molecular information theory approach can be extended to any genetic system to predict new sites for DNA-binding proteins.

INTRODUCTION

p53 is a transcription factor that acts as a tumor suppressor and modulates expression of genes related to the cell cycle, DNA repair, apoptosis and angiogenesis (1). More than half of the tumors in some cancer types have mutations in p53 (2–4). Whole genome scanning for transcription factor sites in conjunction with experimental confirmation can fill the gaps in our understanding of gene regulation networks (5,6).

El-Deiry *et al.* (7) proposed the p53 consensus sequence 5'-PuPuPuC(A/T)(T/A)GPyPyPy-[0–13 bp]-PuPuPuC(A/T)(T/A)GPyPyPy-3', which consists of two decameric sequences separated by 0–13 bp. Microarray experiments have established that hundreds of genes are controlled by p53 (8–12). Attempts have been made to predict p53 binding sites using base frequency weight matrices (13–16) and hidden Markov models (17).

We present a p53 DNA binding model, based on Claude Shannon's information theory (18,19), which sharply distinguishes between specific and non-specific DNA-binding sites (20,21). This theory has been applied to genetic control systems including replication (22,23), transcription (5,24–26), splicing (27) and translation (28). It also has application beyond genetic control in characterizing molecular states and patterns in general (29,30). The information measure consistently accounts for sequence variability and conservation in universal units, bits of information (31). [A bit is a unit of information that allows one to distinguish between two states (19).] We analyzed binding sequences from two earlier studies (7,32) and from a collection of proven natural sites (i.e. naturally occurring experimentally confirmed sites).

To identify p53 response elements (REs), El-Deiry *et al.* selected more than 500 human genomic fragments and tested them for p53 protein binding *in vitro* (7). Twenty REs were identified. Since p53 binds a decameric site as a dimer, both the sequences and their complementary strands for the two decamers of the p53 RE were used to construct an information theory model (Figure 1a). Because tetrameric p53 binds to two dimeric sites, a total of 80 sequences are available from the 20 REs.

Another attempt to identify p53 REs used cyclic amplification and selection of targets (32). p53 was incubated with DNA containing degenerate bases, the p53–DNA complex was purified, the DNA was amplified by PCR and this cycle was repeated. A decameric model built from 17 sequenced DNAs is shown in Figure 1b.

METHODS

Methods are provided in the Supplementary Data.

RESULTS

To construct a natural model, we analyzed p53 decameric sites from 35 previously identified p53-controlled genes, containing the experimentally confirmed sites

*To whom correspondence should be addressed. Tel: +1 301 846 5581; Fax: +1 301 846 5598; Email: toms@ncifcrf.gov

(Supplementary Materials, Table 1). Individual information was calculated by the method given previously (33) and the sites with information less than zero were removed from the list because these sites can not bind a protein specifically, according to the second law of thermodynamics (30). We aligned the remaining 66 decameric sites and their complementary strands (34), created a sequence logo (27) and generated an individual information weight matrix with a range from -4 to $+5$ (33).

We built a flexible p53 model containing two rigid decameric sites (Figure 1c) and a flexible spacer (Figure 1d) (25,28) (Supplementary Materials, Table 2). The natural model (Figure 1c) is more accurate than the consensus models of El-Deiry *et al.* (7) and Funk *et al.* (32) because ignoring the base frequencies and instead counting matches to a consensus inappropriately makes conserved bases equally important for binding (35). Also, they used artificially selected sets of sites (Figure 1a and b), but we

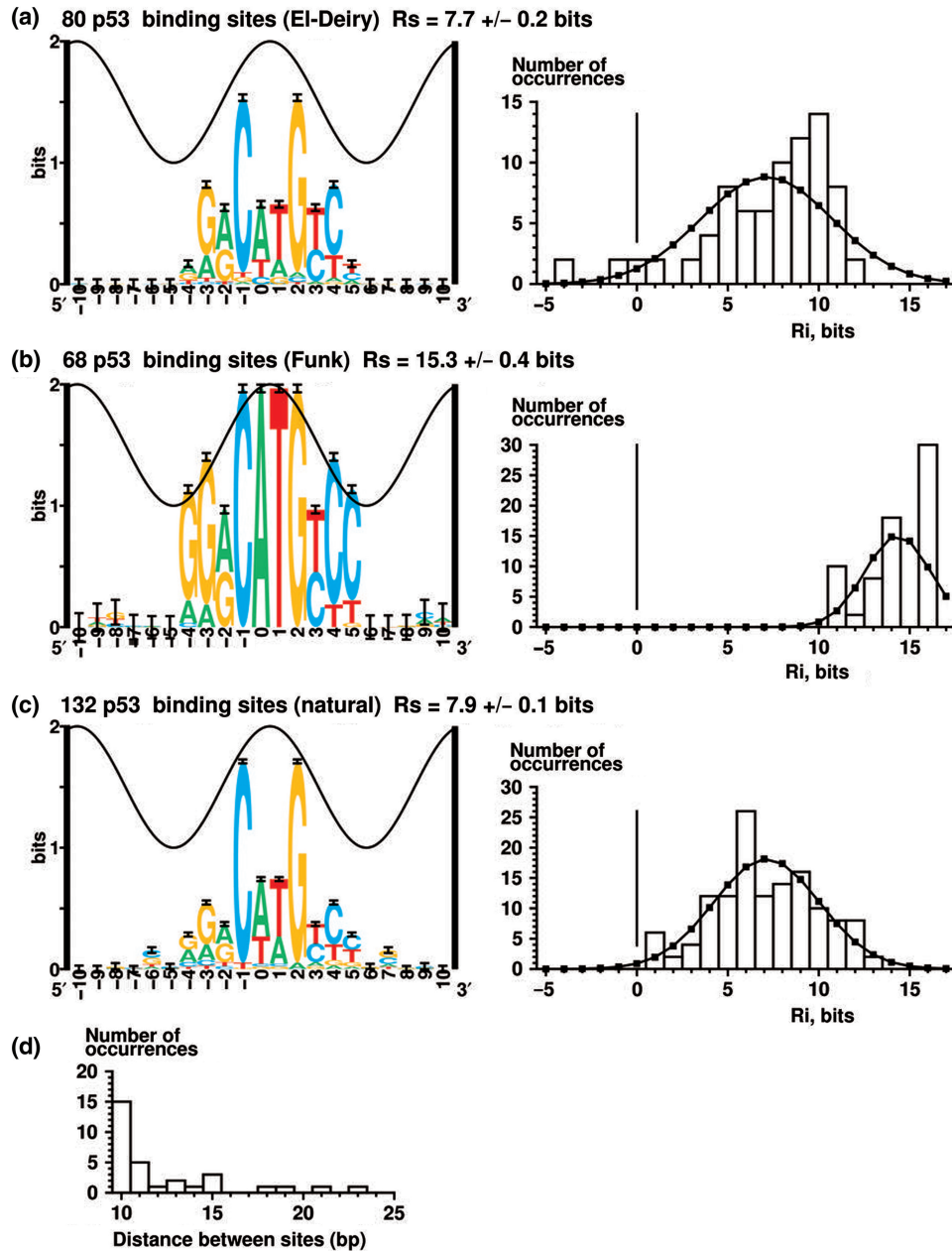


Figure 1. Decameric p53 models. The sequence logos (left) and individual information (R_i) distribution histograms (right) for individual binding sites come from (a) El-Deiry *et al.* (7), (b) Funk *et al.* (32) and (c) our collection of proven natural sites (Supplementary materials, Table 1). R_s is the total information (area of the sequence logo) and also the average of the R_i distribution. Error bars indicate the standard deviation of the information based on sample size. The peaks of sine waves (wavelength of 10.6 bp) above the logos represent the DNA major groove facing the protein, as determined by X-ray crystallography (48) and methylation interference (7). As with a number of other binding sites (22), the sequence logos of p53 follow the accessibility sine wave, especially the Funk data, which are remarkably close. On the distribution histograms a Gaussian curve with the same mean and standard deviation as the data is shown and a vertical line indicates 0 bits of information. (d) Histogram of the distance between natural decameric sites.

used only naturally occurring experimentally confirmed sites. SELEX and other artificial selection methods can give inconsistent results (36).

The El-Deiry model (Figure 1a) is similar to the natural one (Figure 1c). We suggest that the Funk model (Figure 1b) has excess information because strong sites were selected. According to molecular information theory, R_s is the total information in the binding site model, while R_f is the information required to locate the sites in a genome (31). These two independently determined numbers are expected to converge during evolution (37). In the natural p53 model, accounting for the variable distance between the decamers reduces the total information of two decamers to $R_s = 12.3 \pm 3.1$ bits (25,28). In comparison, El-Deiry *et al.* (7) found 18 REs in 530 DNA fragments

ranging from 139 to 470 bp long, which requires $RF = 13.1 \pm 0.8$ bits. Thus, the information R_s in our binding site model is close to the information R_f needed to find the sites in the genome, as occurs for other genetic systems (31,37), suggesting that the natural model is reasonable.

The average information content of the flexible p53 model is 12.3 ± 3.1 bits and 50% of the distances between the natural p53 REs and their promoters are < 300 bp, so we scanned genomic sequences around identified promoters (range -300 to $+100$) with the flexible p53 model, using a 12-bit cutoff. Each decameric site was at least 5 bits. We chose these parameters to identify the strongest sites. The sequences of human chromosomes 1 and 2 [ftp://ftp.ncbi.nih.gov/genomes/H_sapiens/ (38)] were scanned and 16 sites were found (Table 1; Supplementary Materials, Table 3).

There are two more proteins from the p53 family, p63 and p73, which are involved in cell cycle arrest and apoptosis (39). These transcription factors have DNA-binding domains, similar to p53, and, therefore, bind some of the p53 REs that cause transcription activation of p53-dependent genes. Because our initial data set may contain binding sites for these additional proteins, the sites predicted by the genomic scan were confirmed by electrophoretic mobility shift assays (EMSA) with p53, p63 and p73 proteins (Figure 2). p53 bound all predicted sites except S100A6. p63 does not bind the KCNA2-2, S100A6 and BZW1 sites, and weakly binds LEPRE1, KCNA2-1. All other sites form stable complexes with the p63 protein. p73 binds all sites except S100A6. Therefore, 15 out of the 16 sites show affinity to the p53, p63 or p73 proteins.

We used a luciferase reporter assay to confirm the predicted sites in human cells. Promoters containing each site were cloned upstream of the luciferase gene, the plasmids were co-transfected with expression vectors for p53, p63, p73 and a negative control into HEK293 cells, and luciferase activity was measured (Figure 3). Seven genes (CLCA2, FLJ43374, UGT1A6, FLJ38753, KCNA2, PROM2 and H6PD) were activated > 5 -fold by at least one

Table 1. Genes containing the predicted p53 response elements

Gene name	Information content, bits	Description
H6PD	12.5	Hexose-6-phosphate dehydrogenase
FLJ38753	12.0	Hypothetical protein
LEPRE1	12.2	Proteoglycan, potential growth suppressor
MGC955	12.2	Hypothetical protein
RPS8	13.0	Ribosomal protein S8
CLCA2	12.1	Calcium-activated ion channel protein
KCNA2	14.2, 12.2	Potassium channel protein
S100A6	12.1	S100 calcium binding protein A6 (calcyclin)
RDH14	13.5	Retinol dehydrogenase
DQX1	13.3	DEAQ box polypeptide 1 (RNA-dependent ATPase)
VPS24	14.6	Transmembrane protein sorting
PROM2	13.0	Prominin 2
U5-200KD	12.4	U5 snRNP-specific protein, RNA helicase
BZW1	13.3	Basic leucine zipper protein
UGT1A6	13.9	UDP glycosyltransferase
FLJ43374	12.6	Hypothetical protein

The KCNA2 gene contains two p53 REs.

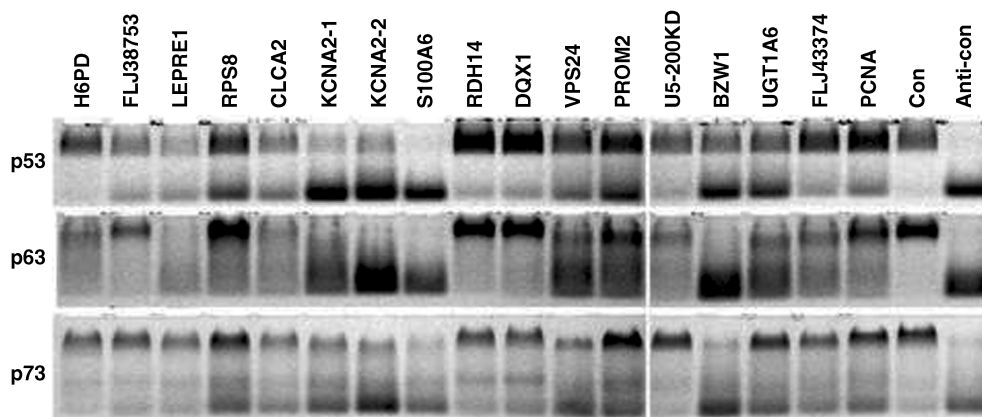


Figure 2. Electrophoretic mobility shift assays (EMSA) with hairpin oligonucleotides containing predicted p53 binding sites (Supplementary materials, Table 6) using the p53, p63 and p73 proteins. The bottom bands are unbound oligonucleotides, and the top bands are protein-oligonucleotide complexes. Names of the predicted genes are marked on the top. The 'PCNA' oligonucleotide containing the p53 RE from the promoter of the PCNA gene (49) is a positive control. The KCNA gene contains three close decameric p53 sites. Oligonucleotide 'KCNA-1' contains sites 1 and 2, oligonucleotide 'KCNA-2' contains sites 2 and 3. The 'Con' oligonucleotide containing the consensus p53 binding site is a positive control. 'Anti-con' has no p53 binding sites and is a negative control.

of the proteins. Five genes (RPS8, DQX1, VPS24, RDH14 and U5-200KD) showed 2- to 5-fold activation by at least one of the proteins. Two divergently expressed genes (MGC955 and LEPRE1), which contain a common p53 RE, showed ~2-fold repression by p53 but not by p63 or p73. Two genes (S100A6 and BZW1) showed no regulation by p53, p63 or p73. Therefore, by the luciferase reporter assay, 14 genes out of the 16 were activated or repressed by at least one of the proteins.

Total RNA from p53-, p63-, p73-transfected cells was isolated and analyzed by qPCR using primers for detection of the 16 predicted genes and the GAPDH and actin controls (Supplementary Materials, Table 4).

The qPCR data (Supplementary Materials, Tables 5A–D) are summarized in Figure 3. Four genes (CLCA2, FLJ43374, KCNA2 and PROM2) were activated >5-fold. Three genes (FLJ38753, DQX1 and VPS24) showed 2- to 5-fold activation by at least one of the proteins. Five genes (RPS8, RDH14, U5-200KD, S100A6 and LEPRE1) showed no activation by p53, p63 or p73. The UGT1A6, MGC955 and BZW1 gene expression was undetectable by qPCR. Although H6PD appears to be repressed (Supplementary Materials, Table 5D), the computed errors are larger than the averages, so we did not report it in Figure 3. Therefore, by qPCR 7 out of the 16 genes showed regulation by the proteins.

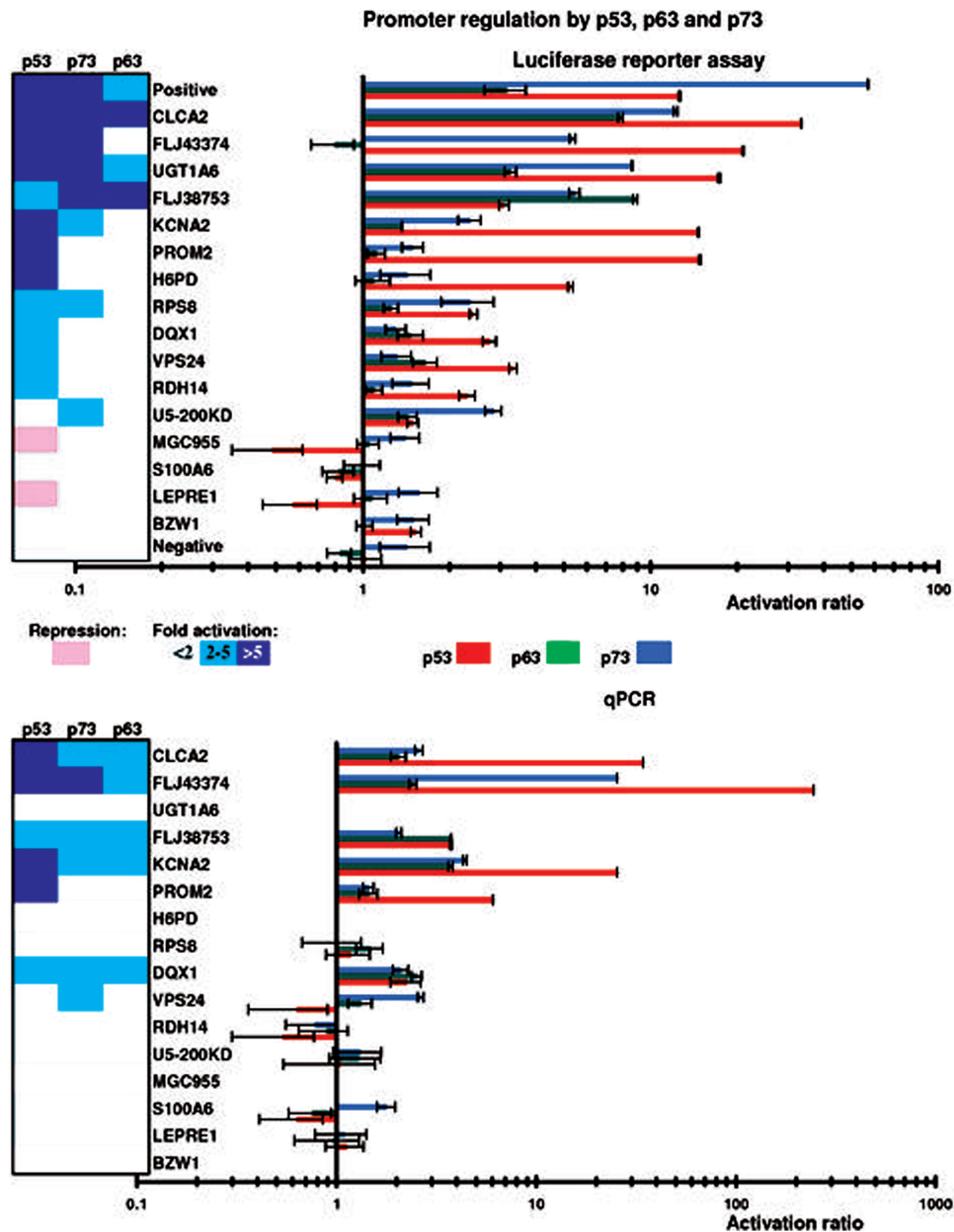


Figure 3. Transcriptional regulation of genes containing the predicted binding sites by p53, p63 and p73. The charts show the ratio between induced and non-induced luciferase signals (top chart) and qPCR signals (bottom chart). Rectangles at the left side represent promoter induction fold for p53, p63 and p73 proteins. The white area in the rectangles means that the signals were undetectable. Error bars indicate the standard deviation of the luciferase signal from two experiments.

Some of our findings are consistent with microarray data: CLCA2 is p53-, p63- and p73-inducible by 40-fold or more (40), VPS24 is p53-inducible by 2-fold (16), DQX1 and PROM2 in human mammary epithelial cells were upregulated after expression of p53 (41). The S100A6 gene was not induced by p53 (12), which is consistent with our results. The other genes we found have not been analyzed or did not show induction or repression in microarray experiments. Our method located 11 previously unidentified REs.

The only exceptional gene, S100A6, was not bound by any proteins *in vivo* or *in vitro*. We suggest that this site has an unusual DNA structure that blocks binding.

We searched all transcription starts in the human Reference Sequence [GenBank accessions NC_000001 to NC_000024, 2006, build 36 version 2, (42)] using the same parameters as before, and identified 198 potentially controlled genes (Supplementary Materials, Table 8). There were two missing RE compared to our previous search. The transcription start of H6PD is now annotated to be 10121 bp upstream, leaving the p53 site in the middle of the gene just after the start of the second exon, yet it activates luciferase expression *in vivo* (Figure 3). Genetic controls inside genes are known, an example for p53 is in the LIF gene (43). However, there are other possible explanations for this activation. Therefore, only 13 of the 16 genes are clearly activated by p53. Likewise, in the new sequence the annotated transcription start for S100A6 is shifted 214 bp upstream (old: NT_086596 referring to the mRNA NM_014624.2; new: NC_000001.9 referring to NM_014624.3), placing the predicted p53 site outside the searched region (−300 to +100), which explains why it was not located in the new search. This location does not explain the lack of p53, p63 or p73 binding to the sequence.

DISCUSSION

We used a proven mathematical method, molecular information theory, to measure DNA binding site conservation in universal units, bits. Not all positions in a site are equally important for protein binding (35). Information content measures the conservation of bases and allows for comparison of different positions (Figure 1). Summing the information content of the bases across the site gives the total information content of the site, which is an important physiological parameter (31,37). In contrast, summing any other function gives inconsistent results, and the total site information cannot be calculated (33). Because of its mathematical consistency, the molecular information theory-based flexible p53 model gave accurate gene predictions.

p53, p63 and p73 are known to have different affinities to different sequences (41,44) and they can be both activators and repressors (45–47). This is consistent with our data. Our model, built from experimentally proven, naturally occurring sites, is a combined model for p53, p63 and p73 that does not distinguish between the proteins. In order to build models specific for each particular protein, one would have to have 3 sets of sites, one for each of p53, p63 and p73. In contrast, the model built using El-Deiry's set is a

pure p53 model. The differences in conservation pattern between the models (Figure 1a versus 1c) may reflect the contribution of p63 and p73 proteins in our natural model. Even so, we used three different experimental methods with all three proteins to test the sites predicted using the natural model, and were able to show that all of the sites except one have *in vitro* or *in vivo* activity. Such a precise and reliable method for prediction of the p53 family response elements will allow further discovery of cancer-related genes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Fulbright Program and the United States Educational Foundation in India (USEFI) for supporting AK. We also thank Dr Christian Klein (Roche Diagnostics GmbH, Pharma Research, Penzberg, Germany) for p53 protein and p53 expressing plasmid p11435, Dr Hua Lu (Oregon Health Sciences University) for pcDNA3-HA-p73b plasmid, Dr Satrajit Sinha (State University of New York at Buffalo) for pColdIdNp63 plasmid, Dr Bimalendu Dasmahapatra (Schering-Plough Research Institute) for GST-His p73DBD plasmid, Dr Boopathy Ramakrishnan (SAIC-Frederick, Inc.) for helpful advice regarding p53 protein isolation, Peter Rogan, Danielle Needle and Peyman Khalichi for editorial comments on the manuscript. This publication has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. This Research was supported [in part] by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. Funding to pay the Open Access publication charges for this article was provided by the National Cancer Institute.

Conflict of interest statement. None declared.

REFERENCES

1. Levine,A.J., Hu,W., and Feng,Z. (2006) The P53 pathway: what questions remain to be explored? *Cell Death Differ.*, **13**, 1027–1036.
2. Petitjean,A., Mathe,E., Kato,S., Ishioka,C., Tavtigian,S.V., Hainaut,P. and Olivier,M. (2007) Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum. Mutat.*, **28**, 622–629.
3. Soussi,T. (2000) The p53 tumor suppressor gene: from molecular biology to clinical investigation. *Ann. N. Y. Acad. Sci.*, **910**, 121–137; discussion 137–139.
4. Soussi,T., Kato,S., Levy,P.P. and Ishioka,C. (2005) Reassessment of the TP53 mutation database in human disease by data mining with a library of TP53 missense mutations. *Hum. Mutat.*, **25**, 6–17.
5. Vyhldal,C.A., Rogan,P.K. and Leeder,J.S. (2004) Development and refinement of pregnane X receptor (PXR) DNA binding site model using information theory: insights into PXR-mediated gene regulation. *J. Biol. Chem.*, **279**, 46779–46786.

6. Kohn, K.W. (1999) Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Mol. Biol. Cell*, **10**, 2703–2734.
7. 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.
8. Wang, L., Wu, Q., Qiu, P., Mirza, A., McGuirk, M., Kirschmeier, P., Greene, J.R., Wang, Y., Pickett, C.B. and Liu, S. (2001) Analyses of p53 target genes in the human genome by bioinformatic and microarray approaches. *J. Biol. Chem.*, **276**, 43604–43610.
9. 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.
10. Mirza, A., Wu, Q., Wang, L., McClanahan, T., Bishop, W.R., Gheys, F., Ding, W., Hutchins, B., Hockenberry, T., Kirschmeier, P. et al. (2003) Global transcriptional program of p53 target genes during the process of apoptosis and cell cycle progression. *Oncogene*, **22**, 3645–3654.
11. Ohki, R., Kawase, T., Ohta, T., Ichikawa, H. and Taya, Y. (2007) Dissecting functional roles of p53 N-terminal transactivation domains by microarray expression analysis. *Cancer Sci.*, **98**, 189–200.
12. 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.
13. Bourdon, J.C., Deguin-Chambon, V., Lelong, J.C., Dessen, P., May, P., Debuire, B. and May, E. (1997) Further characterisation of the p53 responsive element-identification of new candidate genes for trans-activation by p53. *Oncogene*, **14**, 85–94.
14. Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.*, **23**, 4878–4884.
15. Hoh, J., Jin, S., Parrado, T., Edington, J., Levine, A.J. and Ott, J. (2002) The p53MH algorithm and its application in detecting p53-responsive genes. *Proc. Natl Acad. Sci. USA*, **99**, 8467–8472.
16. Sbisà, E., Catalano, D., Grillo, G., Licciulli, F., Turi, A., Liuni, S., Pesole, G., De Grassi, A., Caratozzolo, M.F., D'Erchia, A.M. et al. (2007) p53FamTaG: a database resource of human p53, p63 and p73 direct target genes combining in silico prediction and microarray data. *BMC Bioinform.*, **8**, (Suppl. 1), S20.
17. Li, W., Meyer, C.A. and Liu, X.S. (2005) A hidden Markov model for analyzing ChIP-chip experiments on genome tiling arrays and its application to p53 binding sequences. *Bioinform.*, **21** (Suppl. 1), i274–i282.
18. Shannon, C.E. (1948) A Mathematical Theory of Communication. *Bell System Tech. J.*, **27**, 379–423, 623–656.
19. Pierce, J.R. (1980) *An Introduction to Information Theory: Symbols, Signals and Noise*. Dover Publications, Inc., New York.
20. Shultzaberger, R.K., Roberts, L.R., Lyakhov, I.G., Sidorov, I.A., Stephen, A.G., Fisher, R.J. and Schneider, T.D. (2007) Correlation between binding rate constants and individual information of *E. coli* Fis binding sites. *Nucleic Acids Res.*, **35**, 5275–5283.
21. Schneider, T.D. (2006) Claude Shannon: Biologist. *IEEE Engineering in Medicine and Biology Magazine*, **25**, 30–33.
22. Schneider, T.D. (2001) Strong minor groove base conservation in sequence logos implies DNA distortion or base flipping during replication and transcription initiation. *Nucleic Acids Res.*, **29**, 4881–4891.
23. Lyakhov, I.G., Hengen, P.N., Rubens, D. and Schneider, T.D. (2001) The P1 phage replication protein RepA contacts an otherwise inaccessible thymine N3 proton by DNA distortion or base flipping. *Nucleic Acids Res.*, **29**, 4892–4900.
24. Chen, Z. and Schneider, T.D. (2006) Comparative analysis of tandem T7-like promoter containing regions in enterobacterial genomes reveals a novel group of genetic islands. *Nucleic Acids Res.*, **34**, 1133–1147.
25. Shultzaberger, R.K., Chen, Z., Lewis, K.A. and Schneider, T.D. (2007) Anatomy of *Escherichia coli* σ^{70} promoters. *Nucleic Acids Res.*, **35**, 771–788.
26. Chen, Z., Lewis, K.A., Shultzaberger, R.K., Lyakhov, I.G., Zheng, M., Doan, B., Storz, G. and Schneider, T.D. (2007) Discovery of Fur binding site clusters in *Escherichia coli* by information theory models. *Nucleic Acids Res.*, **35**, 6762–6777.
27. Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.*, **18**, 6097–6100.
28. Shultzaberger, R.K., Bucheimer, R.E., Rudd, K.E. and Schneider, T.D. (2001) Anatomy of *Escherichia coli* ribosome binding sites. *J. Mol. Biol.*, **313**, 215–228.
29. Schneider, T.D. (1991) Theory of molecular machines. I. Channel capacity of molecular machines. *J. Theor. Biol.*, **148**, 83–123.
30. Schneider, T.D. (1991) Theory of molecular machines. II. Energy dissipation from molecular machines. *J. Theor. Biol.*, **148**, 125–137.
31. Schneider, T.D., Stormo, G.D., Gold, L. and Ehrenfeucht, A. (1986) Information content of binding sites on nucleotide sequences. *J. Mol. Biol.*, **188**, 415–431.
32. 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.
33. Schneider, T.D. (1997) Information content of individual genetic sequences. *J. Theor. Biol.*, **189**, 427–441.
34. Schneider, T.D. and Mastronarde, D. (1996) Fast multiple alignment of ungapped DNA sequences using information theory and a relaxation method. *Discrete Appl. Math.*, **71**, 259–268.
35. Schneider, T.D. (2002) Consensus sequence Zen. *Appl. Bioinform.*, **1**, 111–119.
36. Shultzaberger, R.K. and Schneider, T.D. (1999) Using sequence logos and information analysis of Lrp DNA binding sites to investigate discrepancies between natural selection and SELEX. *Nucleic Acids Res.*, **27**, 882–887.
37. Schneider, T.D. (2000) Evolution of biological information. *Nucleic Acids Res.*, **28**, 2794–2799.
38. Istrail, S., Sutton, G.G., Florea, L., Halpern, A.L., Moberly, C.M., Lippert, R., Walenz, B., Shatkay, H., Dew, I., Miller, J.R. et al. (2004) Whole-genome shotgun assembly and comparison of human genome assemblies. *Proc. Natl Acad. Sci. USA*, **101**, 1916–1921.
39. Murray-Zmijewski, F., Lane, D.P. and Bourdon, J.C. (2006) p53/p63/p73 isoforms: an orchestra of isoforms to harmonise cell differentiation and response to stress. *Cell Death Differ.*, **13**, 962–972.
40. Osada, M., Park, H.L., Nagakawa, Y., Begum, S., Yamashita, K., Wu, G., Kim, M.S., Trink, B. and Sidransky, D. (2006) A novel response element confers p63- and p73-specific activation of the WNT4 promoter. *Biochem. Biophys. Res. Commun.*, **339**, 1120–1128.
41. Perez, C.A., Ott, J., Mays, D.J. and Pietenpol, J.A. (2007) p63 consensus DNA-binding site: identification, analysis and application into a p63MH algorithm. *Oncogene*, **26**, 7363–7370.
42. International Human Genome Sequencing Consortium. (2004) Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931–945.
43. Hu, W., Feng, Z., Teresky, A.K. and Levine, A.J. (2007) p53 regulates maternal reproduction through LIF. *Nature*, **450**, 721–724.
44. Lokshin, M., Li, Y., Gaiddon, C. and Prives, C. (2007) p53 and p73 display common and distinct requirements for sequence specific binding to DNA. *Nucleic Acids Res.*, **35**, 340–352.
45. Ho, J. and Benchimol, S. (2003) Transcriptional repression mediated by the p53 tumour suppressor. *Cell Death Differ.*, **10**, 404–408.
46. Testoni, B. and Mantovani, R. (2006) Mechanisms of transcriptional repression of cell-cycle G₂/M promoters by p63. *Nucleic Acids Res.*, **34**, 928–938.
47. Racek, T., Mise, N., Li, Z., Stoll, A. and Pützer, B.M. (2005) C-terminal p73 isoforms repress transcriptional activity of the human telomerase reverse transcriptase (hTERT) promoter. *J. Biol. Chem.*, **280**, 40402–40405.
48. Cho, Y., Gorina, S., Jeffrey, P.D. and Pavletich, N.P. (1994) Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346–355.
49. Morris, G.F., Bischoff, J.R. and Mathews, M.B. (1996) Transcriptional activation of the human proliferating-cell nuclear antigen promoter by p53. *Proc. Natl Acad. Sci. USA*, **93**, 895–899.