

Parallel analysis of tetramerization domain mutants of the human p53 protein using PCR colonies

Joshua Merritt · Kim G. Roberts · James A. Butz ·
Jeremy S. Edwards

Received: 23 October 2006 / Accepted: 1 August 2007 / Published online: 5 September 2007
© Springer Science+Business Media B.V. 2007

Abstract A highly-parallel yeast functional assay, capable of screening approximately 100–1,000 mutants in parallel and designed to screen the activity of transcription activator proteins, was utilized to functionally characterize tetramerization domain mutants of the human p53 transcription factor and tumor suppressor protein. A library containing each of the 19 possible single amino acid substitutions (57 mutants) at three positions in the tetramerization domain of the human p53 protein, was functionally screened in *Saccharomyces cerevisiae*. Amino acids Leu330 and Ile332, whose side chains form a portion of a hydrophobic pocket that stabilizes the active p53 tetramer, were found to tolerate most hydrophobic amino acid substitutions while hydrophilic substitutions resulted in the inactivation of the protein. Amino acid Gln331 tolerated

essentially all mutations. Importantly, highly parallel mutagenesis and cloning techniques were utilized which, in conjunction with recently reported highly parallel DNA sequencing methods, would be capable of increasing throughput an additional 2–3 orders of magnitude.

Keywords Polonies · p53 · Genomics · High-throughput DNA sequencing

Introduction

The ability to use genotype data to understand complex phenotypes, for example the ability to predict the susceptibility to infectious disease or the likelihood of developing cancer based upon one's genome sequence, was one of the early dreams of the Human Genome Project. The publication of the human genome sequence (Lander et al. 2001; Venter et al. 2001) and ongoing research devoted to describing genetic variation within the human population (Masood 1999) are initial steps in this effort. One of the next major steps in the Human Genome Project is to understand the physiological effect of genetic diversity in the human population. Specifically, it is critical to identify and characterize the subset of SNPs and other mutations which impact biochemical function. In order to achieve this goal several additional questions must be addressed. *How do we identify functionally important mutations? How do we identify functionally important residues within a protein? How do we analyze the relation between single point mutations and the complex integrated functions in a cellular system?* Herein, we seek to develop a framework for addressing these questions in a high-throughput fashion. Specifically, we developed a highly parallel functional assay based upon the ability of the human p53 (p53) protein to initiate

Present Address:

J. Merritt (✉)
Vaccine Research Center, NIAID, NIH, Bethesda,
MD 20892, USA
e-mail: merrittj@niaid.nih.gov

J. Merritt · K. G. Roberts · J. A. Butz
Department of Chemical Engineering, University of Delaware,
Newark, DE 19716, USA

Present Address:

J. A. Butz
Schering-Plough Corporation, Kenilworth, NJ 07033, USA

J. S. Edwards
Molecular Genetics and Microbiology and the Cancer Research
and Treatment Center, University of New Mexico Health
Sciences Center, Albuquerque, NM 87131, USA
e-mail: jsedwards@salud.unm.edu

J. S. Edwards
Chemical and Nuclear Engineering Department, University of
New Mexico, Albuquerque, NM 87131, USA

transcription of target proteins under the control of a p53 response element mediated promoter in *Saccharomyces cerevisiae* (Scharer and Iggo 1992). We then used this functional assay to screen the activity of 57 single codon mutants (all possible single amino acid substitutions) at positions Leu330, Gln331 and Ile332 of the human p53 gene.

Our assay was inspired by the FASAY screen initially reported by Flaman et al. (1995) and its variations (e.g. (Jia et al. 1997)) and, in fact, utilized a reporter strain developed for this application (Tomso et al. 2005). Unlike these assays, which utilize separate colony growth on solid agar to identify and isolate functional and non-functional p53 expressing strains and standard DNA sequencing to identify a specific inactivating mutation, we utilized mixed mutant growth competitions, polymerase colony (polony) (Mitra and Church 1999) and primer extension sequencing technology (Mitra et al. 2003) similar to methods we have reported previously (Merritt et al. 2003; Merritt et al. 2005). The primary advantage of our methodology is that mutant enrichment (via mixed strain growth competition) and identification of the associated mutation(s) (polony based) are highly parallel. Our assay has the ability to screen the function of approximately 100–1,000 strains in parallel. Further, by applying recently reported ultrahigh throughput DNA sequencing (Margulies et al. 2005; Shendure et al. 2005) and making minor modifications, throughput could be increased several orders of magnitude.

As a target for mutation analysis, the p53 gene is of great interest (Hernandez-Boussard et al. 1999) due to the high prevalence of mutations in the gene in almost every type of human cancer. p53 is a tumor suppressor gene that binds DNA sequences (Kern et al. 1991) and activates the transcription of various genes including several that induce cell-cycle arrest and apoptosis (Chappuis et al. 1999). The p53 monomer contains three primary domains associated with this function—an N-terminal transactivation domain, a central DNA binding domain and a tetramerization domain located near the C-terminus (Ko and Prives 1996). The majority of identified mutations associated with cancer (87%) are localized to the DNA binding domain (Levine et al. 1995). However, mutations which inactivate the protein have also been identified in the transactivation and tetramerization domains (Chene and Bechter 1999). The amino acid positions screened in this work were localized in a portion of the tetramerization domain encoding a β -sheet substructure believed to stabilize the assembled functional p53 tetramer.

Briefly, a strain library was constructed in which p53 mutants were expressed in a p53 reporter strain of *Saccharomyces cerevisiae*. This strain was designed such that functional p53 expression activated expression of the ADE2 gene which was essential for growth of the strain in adenine deficient medium. Library growth enriched strains

expressing p53 mutants with higher activities in the culture and depleted strains expressing less active p53 mutants from the culture. Similar growth competition methodology has been used previously, primarily to determine the function of native genes via parallel analysis of whole-gene deletion libraries (Merritt and Edwards 2004). Using this method we identified both tolerated and non-tolerated mutations in the three amino acid positions tested. The identities of tolerated and non-tolerated amino acid substitutions were interpreted in conjunction with reported structural and epidemiological data.

Materials and methods

Strain and media

Yeast growth experiments were conducted in the AR-HGEF7 strain#2152 (Tomso et al. 2005). Briefly, the strain has been modified to express the ADE2 gene under the control of a p53 response element mediated promoter. As a result, the strain can be utilized for standard colony color screening (Flaman et al. 1995) or selective growth in YPD for active p53 expression in media low or deficient in adenine, respectively.

Media used in this work were the following. YPD (Sherman 1991): 10 g/l yeast extract (Fisher Scientific), 20 g/l peptone (Fisher Scientific), 20 g/l dextrose (Fisher Scientific). SD: 6.7 g/l Yeast Nitrogen Base Without Amino Acids (Difco), 20 g/l dextrose. SD was supplemented with 0.68 g/l Complete Supplement Mixture—Adenine-Leucine (Qbiogene) and, prior to introduction of the plasmid, 30 mg/l leucine (Sigma Aldrich) and either 30, 5 or 0 mg/l adenine (Sigma Aldrich) for general growth, color screening or selective growth, respectively. LB: 10 g/l Tryptone (Difco), 5 g/l yeast extract, 10 g/l NaCl. LB medium was supplemented with 50 mg/l ampicillin. Solid agar media were made using 15 g/l agar (Fisher Scientific) in the above media.

Nucleic acid manipulation and plasmid construction

Tagged human p53 expression vectors expressing all possible single amino acid substitutions at amino acid positions 330, 331, 332 (57 total) were constructed in a three step process outlined in Fig. 1 using the p415CYC1hG6PDTtag plasmid previously described (Merritt et al. 2005).

Final tagged mutant p53 expression vectors were assembled as follows. Both mutant constructs and the set of plasmids bearing the 200 bp tags were dual digested with HindIII and SpeI (Invitrogen) and gel purified. Purified

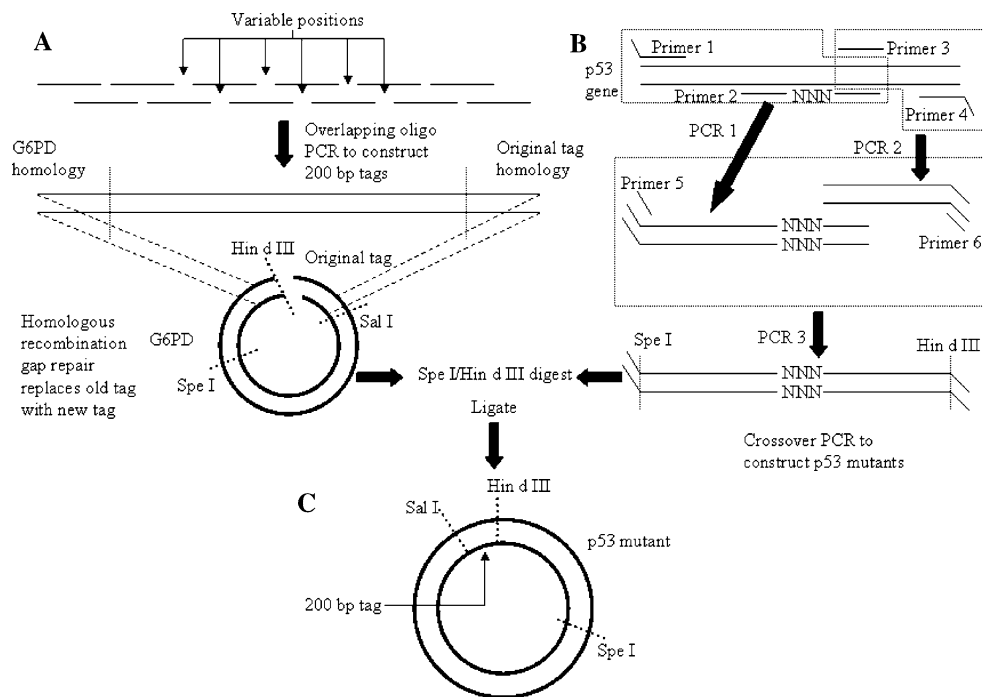


Fig. 1 Construction of tagged mutant p53 expression plasmids. Plasmids were constructed in a three step process. **(A)** 200 bp tags containing six variable positions were synthesized using overlapping oligonucleotide PCR; tags were cloned into the expression plasmid using homologous recombination mediated gap repair. **(B)** Mutant

p53 genes were constructed using mutagenic crossover PCR; both degenerate (-NNN-) and mutant specific primers were used. **(C)** Final plasmid assembly was done by ligating SpeI/HindIII digested tagged vector and mutant p53 genes from **(A)** and **(B)**

plasmid and individual mutant digests were mixed in equimolar quantities (specific tags were matched with each mutant p53) and ligated using the Quick Ligation Kit (New England Biolabs). Ligation reactions were transformed into competent cells, plasmid DNA from each ligation was isolated and the presence of the gene and tag were confirmed by triple digest with Spe I, Hind III and Sal I. The p53 coding region of each final construct was sequenced to confirm the identity of the mutation it carried.

Polony slides

Plasmid DNA from the growth competition experiments was used to quantify the relative concentrations of each mutant in culture as a function of time. Initially, plasmid DNA was isolated from culture samples using the Yeast RPM Plasmid Kit (Qbiogene). This plasmid DNA was then used as the template in the polony reactions described below as described in the literature (Merritt et al. 2003; Merritt et al. 2005). Single Base Extension (SBE) Sequencing of the sequence tags (STs) (Fig. 2A) was used to quantify polonies arising from each mutant strain in the mixed culture as described in the literature (Merritt et al. 2003; Merritt et al. 2005).

Gene expression analysis

Gene expression analysis of individual strains was conducted essentially to verify the expression level of the p53 gene as described (Merritt et al. 2003; Merritt et al. 2005; Mikkilineni et al. 2004). The ratio of ADE2:p53 was calculated and the % WT activity for each mutant tested was calculated using a calibration curve ($R^2 = 0.93$) prepared from the ADE2/p53 ratios of mutants with previously reported activity (Leu330Ala, Leu330His, Gln331Ala and Ile332Ala). Western blots were also performed to verify the protein expression level. The blot was initially contacted with anti-p53 (Bethyl Laboratories) at a dilution of 1/2000 and then with anti-Rabbit peroxidase linked secondary antibody (Amersham) at a dilution of 1/4000. Protein concentrations of each sample were normalized to the value of the Leu330Tyr mutant.

Red/White color screen for functional p53 expression

A red/white color screen for functional p53 expression similar to that previously described (Inga et al. 2002) was conducted on all p53 mutants. Three clones of each

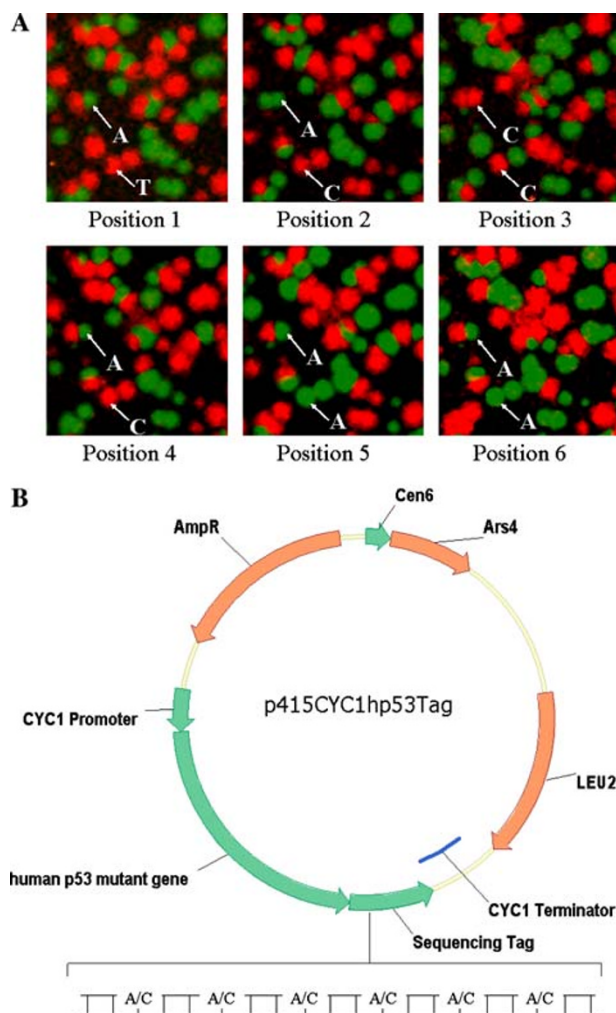


Fig. 2 (A) Polony method used to quantify mutants during p53 growth competition. Plasmid DNA was initially isolated from the culture and used as template in a polony PCR. Common primers amplified all STs, each associated with a different p53 mutant and resulting in an individual polony. Polonies were identified by conducting six sequential single-base extensions using fluorescently-labeled nucleotides. The position of each polony was manually logged after sybr green staining using Metamorph software (Universal Imaging). Data from each extension was assembled into the final code sequence using a simple routine developed in our laboratory. In the representative frames above, two polonies are traced through all six extensions. Upper: tag sequence “AACAAA” corresponds to Gln331Gly. Lower: tag sequence “TCCCAA” corresponds to Gln331Met. (B) Expression plasmids were constructed using the p415CYC1 vector which carries the constitutive CYC1 promoter

strain transformed with a different mutant p53 bearing plasmid were plated on media supplemented with a low concentration (5 mg/l) of adenine and grown at 30°C for four days. Strains were then scored on a scale of 1 (whitest) to 5 (reddest) by three independent evaluators. The average and standard deviation of the three scores was reported.

Growth competition for functional p53 expression

Mutant p53 bearing plasmids were pooled in equimolar quantities and this pool was used to transform the AR-HGEF7 strain. Approximately 1.5 µg of pooled plasmid was used in each of four transformations. To ensure library diversity, the four transformations were then pooled and this pool was used to inoculate medium for the growth competition. The competition pool was initially grown in leucine deficient medium supplemented with 30 mg/l adenine (selective for presence of p53 expression plasmid but non-selective for functional p53 expression) for approximately 10 culture generations. The library was then transferred to adenine deficient medium and grown to mid-log phase (OD_{600} 1.0–3.5). Cells from 10 ml of culture were harvested and stored at –80°C for analysis. This cycle was repeated two additional times so that the competition lasted 15–20 generations.

Results

Growth competition based p53 functional assay

A library of mutant human p53 expression plasmids was screened for functionality using a *S. cerevisiae* growth competition method. The concentrations of each mutant bearing strain in culture was measured at several time points using polonies and single base extensions to identify the unique tag associated with each mutant p53 gene. The specific growth rate of each mutant (greater than 0.5% of the population) was determined using a least-squares curve fitting routine based on the exponential growth equation of each mutant:

$$C_t = C_0 e^{\mu t}$$

Specifically, curve fits to the SBE data were performed using the exponential growth equations of the form:

$$X_{c(n,m)} = \frac{(X_{e(1,m)} \cdot e^{\mu_n \cdot t_m})}{\sum_n X_{e(1,m)} \cdot e^{\mu_n \cdot t_m}}$$

where X_e is an n (number of mutants) \times m (number of time points) matrix containing the experimentally measured percent concentrations of each mutant at each sampling, μ is a $n \times 1$ matrix containing the specific growth rate of each strain in the competition and t is a $1 \times m$ matrix containing the times at which samples were taken and mutant concentrations measured. All elements of μ were allowed to vary in order to minimize the sum of the square of the error ($\Sigma(\text{Error})^2$) between the calculated and measured matrix according to the equation:

$$\sum (\text{Error})^2 = \sum_m \sum_n (X_{c(n,m)} - X_{e(n,m)})^2$$

Results of the growth competition are summarized in Table 1. Although approximately equal amounts of each mutant plasmid were initially supplied to the growth competition, two distinct populations had arisen from this culture at the first sample point after selection was initiated. The first population consisted of mutants capable of growth

in the absence of adenine (i.e. expressing functional p53). The 30 mutants in this population had a narrow range of growth rates (average: 0.198 h^{-1} , standard deviation: 0.009 h^{-1}). The second population, presumably strains expressing non-functional p53, consisted of mutants not present in the competition culture at significant concentrations at the first sampling point after selection was initiated or at subsequent time points. It was therefore not possible to calculate growth rates for this population.

Table 1 Properties of the p53 tetramerization domain mutants

Mutant	WT AA	Mutant AA	Polony μ (h^{-1})	Red/White	%WT activity	p53 Western blot
Leu330Ala		Hydrophobic uncharged	–	3.7 ± 0.6	5.6 ± 4.2	–
Leu330Arg		Hydrophilic basic	–	4.3 ± 0.6	–	–
Leu330Asn		Hydrophilic uncharged	–	4.3 ± 0.6	7.2 ± 14.4	–
Leu330Asp		Hydrophilic acidic	–	5.0 ± 0	47.6 ± 8.0	$121.7\% \pm 8.3\%$
Leu330Cys		Uncharged hydrophobic	0.174	1.7 ± 0.6	–	–
Leu330Gln		Hydrophilic uncharged	–	3.0 ± 0	–	–
Leu330Glu		Hydrophilic acidic	–	5.0 ± 0	54.4 ± 22.3	–
Leu330Gly		Uncharged	–	5.0 ± 0	5.5 ± 5.4	–
Leu330His		Hydrophilic basic	–	4.3 ± 0.6	7.6 ± 10.0	–
Leu330Ile		Hydrophobic uncharged	0.196	1.3 ± 0.6	250.5 ± 7.4	–
Leu330Lys		Hydrophilic basic	–	5.0 ± 0	29.9 ± 16.2	–
Leu330Met		Hydrophobic uncharged	0.205	2.0 ± 1.0	72.8 ± 4.5	–
Leu330Phe		Hydrophobic uncharged	0.192	1.0 ± 0	–	$88.9\% \pm 11.3\%$
Leu330Pro		Uncharged	–	4.0 ± 0	–	–
Leu330Ser		Hydrophilic uncharged	0.186	4.7 ± 0.6	–	–
Leu330Thr		Hydrophilic uncharged	0.193	4.3 ± 0.6	–	–
Leu330Trp		Hydrophobic uncharged	–	4.0 ± 0	–	$98.4\% \pm 7.3\%$
Leu330Tyr		Hydrophobic ionizable	0.198	1.3 ± 0.6	150.6 ± 27.9	100.0%
Leu330Val	Hydrophobic uncharged	Hydrophobic uncharged	0.195	1.3 ± 0.6	164.9 ± 22.6	–
Gln331Ala		Hydrophobic uncharged	0.205	1.0 ± 0	77.6 ± 8.0	–
Gln331Arg		Hydrophilic basic	0.202	1.7 ± 0.6	–	–
Gln331Asn		Hydrophilic uncharged	0.203	2.0 ± 0	–	$80.5\% \pm 17.1\%$
Gln331Asp		Hydrophilic acidic	0.211	1.3 ± 0.6	–	–
Gln331Cys		Uncharged hydrophobic	–	3.0 ± 0	–	–
Gln331Glu		Hydrophilic uncharged	0.203	2.0 ± 0	–	–
Gln331Gly		Hydrophilic acidic	0.203	1.7 ± 0.6	–	–
Gln331His		Uncharged	0.217	1.0 ± 0	153.0	–
Gln331Ile		Hydrophilic basic	0.204	2.0 ± 0	–	–
Gln331Leu		Hydrophobic uncharged	0.209	2.0 ± 0	–	–
Gln331Lys		Hydrophilic basic	0.199	1.0 ± 0	–	–
Gln331Met		Hydrophobic uncharged	0.203	1.3 ± 0.6	–	–
Gln331Phe		Hydrophobic uncharged	0.186	2.0 ± 0	46.3	–
Gln331Pro		Uncharged	0.197	3.3 ± 0.6	37.9 ± 13.2	–
Gln331Ser		Hydrophilic uncharged	0.184	1.0 ± 0	–	–
Gln331Thr		Hydrophilic uncharged	0.194	2.0 ± 0	56.2 ± 6.8	–
Gln331Trp		Hydrophobic uncharged	0.192	4.0 ± 0	107.7 ± 14.7	–
Gln331Tyr		Hydrophobic ionizable	–	2.7 ± 0.6	–	–
Gln331Val	Hydrophilic uncharged	Hydrophobic uncharged	–	2.3 ± 0.6	–	–

Table 1 continued

Mutant	WT AA	Mutant AA	Polony μ (h^{-1})	Red/White	%WT activity	p53 Western blot
Ile332Ala		Hydrophobic uncharged	–	4.3 \pm 0.6	14.7	–
Ile332Arg		Hydrophilic basic	–	5.0 \pm 0	–	–
Ile332Asn		Hydrophilic uncharged	0.205	5.0 \pm 0	33.1	–
Ile332Asp		Hydrophilic acidic	–	5.0 \pm 0	–	–
Ile332Cys		Uncharged hydrophobic	0.202	1.0 \pm 0	–	86.1% \pm 6.9%
Ile332Gln		Hydrophilic uncharged	–	5.0 \pm 0	–	–
Ile332Glu		Hydrophilic acidic	–	5.0 \pm 0	26.9 \pm 6.9	122.2% \pm 25.5%
Ile332Gly		Uncharged	–	4.0 \pm 0	–	–
Ile332His		Hydrophilic basic	–	5.0 \pm 0	–	110.9% \pm 12.0%
Ile332Leu		Hydrophobic uncharged	0.198	1.0 \pm 0	84.2	–
Ile332Lys		Hydrophilic basic	–	5.0 \pm 0	–	–
Ile332Met		Hydrophobic uncharged	0.190	2.0 \pm 0	179.2 \pm 35.8	–
Ile332Phe		Hydrophobic uncharged	–	5.0 \pm 0	–	–
Ile332Pro		Uncharged	–	5.0 \pm 0	–	–
Ile332Ser		Hydrophilic uncharged	–	4.7 \pm 0.6	35.4	–
Ile332Thr		Hydrophilic uncharged	0.192	2.3 \pm 0.6	–	101.0% \pm 6.6%
Ile332Trp		Hydrophobic uncharged	–	4.7 \pm 0.6	–	–
Ile332Tyr		Hydrophobic ionizable	–	4.7 \pm 0.6	–	–
Ile332Val	Hydrophobic uncharged	Hydrophobic uncharged	0.198	2.7 \pm 0.6	56.0	–

The majority of tolerated mutations (16/30) were found at codon 331. Eight tolerated mutations were found at codon 330 and the remaining six tolerated mutations were at codon 332.

Red/white colony plate p53 functional assay

Red/white colony screening, among the most commonly applied p53 functional screens, was used to validate our growth competition methodology and results. Mutant p53 expression plasmids were individually used to transform the ARHGEF7 strain. Three colonies from each transformation were grown on low adenine agar plates and scored on a scale of 1 (whitest, active p53 mutants) to 5 (reddest, inactive p53 mutants) by three evaluators. The average and standard deviation of the three scores is reported in Table 1. Strains with a score less than 3.0 were assumed to express an active p53 mutant and those with a score of 3.0 or greater were assumed to express inactive p53.

Growth competition (GC) and Red/White assay (RW) results agreed in 88% of the cases tested (50/57). The GC correctly identified 93% (25/27) of the active p53 mutants identified by the RW. Gln331Tyr and Gln331Val were identified as active by the RW but missed by the GC. The resulting false negative rate was approximately 4% (2/57).

17% of the mutants (5/30) that were identified as active by the GC were identified as inactive by the RW (false positives). However, in three of the five cases there is ambiguous or contradictory data. Gln331Trp had a RW score of 4.0 but expression profiling (see below) indicates that this mutant has approximately WT activity, suggesting that the GC rather than RW score is correct. Gln331Pro had a borderline RW score (3.3) and protein activity (37.9% WT) in the ambiguous range. Ile332Asn also had a measured protein activity (33.1%) in the ambiguous range. No additional data was available for the two remaining false positives, Leu330Ser or Leu330Thr. It is likely that the true number of false positives was between 2 and 5, resulting in a false positive rate of approximately 4–9%.

Gene expression assay for p53 activity

The ability of mutant p53 to activate transcription of the ADE2 gene was assayed for 24 individual strains by measuring transcript levels of p53 and ADE2 using polonies. An approximate p53 protein activity level, reported as %WT Activity in Table 1, was calculated for each mutant tested based upon the ratio of ADE2:p53 transcripts. In the 11 cases where p53 protein activity was greater than 55% of WT the GC identified all 11 and the

RW identified 10 (Gln331Trp was the exception) as active p53 mutants. Similarly, when protein activity was less than 30% of WT both the GC and the RW identified all mutants (7/7) as inactive. Six of the mutants tested had protein activity levels in the range 30–55% WT activity. In this range of activity, the GC and the RW generally agreed in evaluating mutants. In four of six cases, GC and RW concurred in scoring mutants either active or inactive; a fifth mutant (Gln331Pro) was scored active by the GC but had a borderline inactive RW score (3.3). However, between 30 and 55% WT activity, GC and RW scores were not well correlated with protein activities measured by the gene expression assay. For example, Leu330Glu had the highest protein activity score (54.4% WT) in the 30–55% range but was scored inactive by the GC and the RW (score of 5.0) whereas Gln331Phe had a lower measured protein activity (46.3% WT) but was scored active by the GC and the RW (score 2.0).

p53 protein expression analysis

Relative p53 protein expression levels of nine (arbitrarily selected) mutants from our mutant library were determined by Western blot (Table 1). p53 protein concentrations were normalized with the value of the Leu330Tyr mutant which was in the center of the range. Each mutant tested expressed p53 at a similar level (80.5–122.2% of the Leu330Tyr value) irrespective of protein activities measured by the GC, RW or the gene expression assay.

Discussion

The objective of this work was the development and application of an assay to identify functional and non-functional single amino acid changes in the human tumor suppressor and transcription activator, p53, and to do so in a highly parallel manner. There are currently several technologies available capable of assaying the function of p53. However, each is either inherently low throughput or is incompatible with recently reported ultrahigh-throughput DNA sequencing technologies (Margulies et al. 2005; Shendure et al. 2005), the application of which will likely revolutionize functional genomics. With these new technologies in mind, we designed a growth competition based p53 functional assay which utilizes highly-parallel methodology at every phase (i.e. mutant gene construction, expression vector construction, mutant strain library assemble and assay readout).

In this work, a library consisting of all possible single amino acid substitutions at three positions (57 total) within the tetramerization domain of the p53 protein was

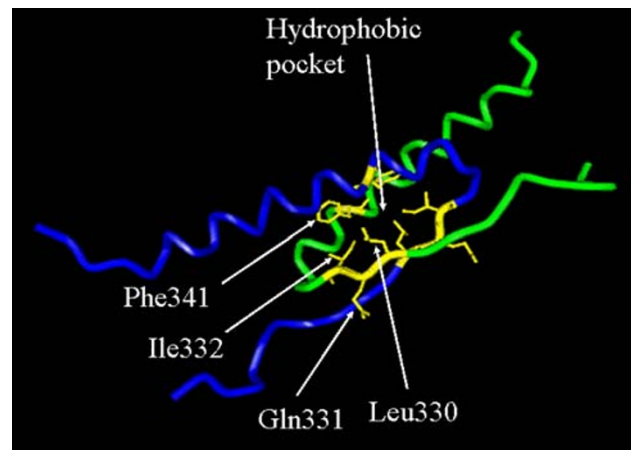


Fig. 3 Human p53 tetramerization domain. Interacting regions of two p53 monomers (peptide backbones visualized in green and blue) which form a dimer. A second identical dimer (not shown) mates with the first to form the final assembled protein. Side chains of the three residues tested in this work, Leu330, Gln331 and Ile332, and the side chain of Phe341 are highlighted. R-groups of Leu330, Ile332 and Phe341 are oriented toward each other and form a portion of the hydrophobic pocket, which stabilizes the p53 tetramer

screened for active and inactive mutants using our highly-parallel growth competition based functional assay. Although, most p53 mutations that are observed clinically are located in the DNA-binding domain, the importance of the tetramerization domain has been recognized and several mutation affecting residues 330–331 have been observed (reviewed in (Chene 2001)). Many reports have characterized such mutant proteins. The amino acid positions tested—Leu330, Gln331 and Ile332—are found within a β -strand spanning residues 326–333 (Fig. 3) which immediately precedes a 22 amino acid α -helix (residues 335–356). Functional tetramer formation arises from interactions between these subunit α -helices (Chene and Bechter 1999). Tetramer formation and stability are also dependant upon the interaction of Leu330 and Ile332 in the β -sheet with Phe341 in the α -helix. Side chains of these three amino acids are in close proximity and form a hydrophobic pocket necessary for protein activity (Rollenhagen and Chene 1998).

Results of our functional screen were consistent with this observation. Six of the eight tolerated mutations at amino acid 330 were hydrophobic or uncharged (Leu330-Cys) substitutions. The two hydrophilic substitutions that were tolerated (Leu330Ser and Leu330Thr) were likely false positives based upon their RW scores. Additionally, it appears that the size of the hydrophobic R-group substitution at position 330 had minimal impact on the function of the protein; only the smallest and largest R-group substitutions, Leu330Ala and Leu330Trp, resulted in inactive mutant proteins. Similar results were observed at Ile332.

Four of the six tolerated mutations were hydrophobic or uncharged (Ile332Cys). Based upon its RW score (5.0), it is likely that Ile332Asn, one of the tolerated hydrophilic substitutions, was a false positive. In Ile332, the size range of tolerated mutations was smaller than that of Leu330. Ile332Ala and each of the large ring structured amino acid mutants (i.e. Ile332Phe, Ile332Tyr and Ile332Trp) were inactive. A priori, it was not expected that mutations at Gln331 would have a significant impact on protein activity. Our functional screen confirmed this hypothesis; sixteen of the nineteen possible amino acid substitutions at this position were identified as functional. The three substitutions at Gln331 identified as nonfunctional (Gln331Cys, Gln331Tyr and Gln331Val) appear to be false negatives based upon their RW scores.

We validated our growth competition methodology and results in four separate ways. The first validation was the inclusion of mutants of known activity in the screen. Three of these mutants, Leu330His, Gln331Ala and Ile332Ala, had previously reported activities relative to wild type p53 ranging from approximately 0–30% (Chene and Bechter 1999; Rollenhagen and Chene 1998); each was identified by the GC as inactive. Gln331Ala, with a reported activity approximately 80% that of wild type p53 (Chene and Bechter 1999), was identified by the GC as active. The second validation was the concurrent screening of all mutant strains included in the GC using the standard Red/White colony screen for p53 activity. There was approximately 90% agreement between the GC and the RW; the majority of the conflicting results (5/7) were false positives, i.e. mutants identified by the GC as active but scored inactive by the RW. The rate of false negatives was approximately 4% while the rate of false positives was approximately 4–9%. Relative to the RW, the GC had a sensitivity of 93% and specificity of 83%. p53 protein expression measurements using Western blots were the third validation of the GC methodology. Each of the nine mutant p53 strains tested expressed the p53 monomer at a similar level although their activities spanned the entire range. mRNA transcription rates from the p415CYC1 expression vector have been previously demonstrated to vary little from clone to clone (Merritt et al. 2003; Merritt et al. 2005). Therefore, similar p53 monomer concentrations among the mutants tested indicate similar rates of protein translation and degradation. The fourth validation of the GC methodology was a direct measurement of transcriptional activation of the ADE2 reporter gene by p53 mutants using polonies. All mutants which had above 55% wild type p53 activity, as measured by the ratio of ADE2/p53 transcripts, were identified by the GC as active (the RW assay agreed in 10/11 cases). Similarly, all mutants which tested below 30% wild type activity were identified as inactive (7/7, both GC and RW). Mutants whose activity

was measured in the intermediate range, 30–55% wild type, had poor qualitative correlation between transcription based activity measurements and functional assays. It can therefore be concluded that our assay, as designed, is most effective at identifying p53 mutants with activities less than 30% or greater than 55% wild type activity.

The primary limitations of the growth competition p53 assay as designed and presented in this work were the inability to accurately measure mutant activities in the 30–55% wild type activity range and the inability to quantitatively measure the growth rates of strains expressing low activity p53 mutants. We hypothesize that both limitations arise from suboptimal assay design, i.e. a non-linear response in growth rate of the mutant bearing strain to changes in the specific activity of the mutant p53 being expressed by the strain. It may be possible to overcome these limitations by optimizing several parameters. For example, it may be possible to discriminate lower activity mutants under conditions of high p53 protein concentration or to discriminate high activity mutants under limiting protein concentrations. To this end, the assay could be redesigned to express p53 under

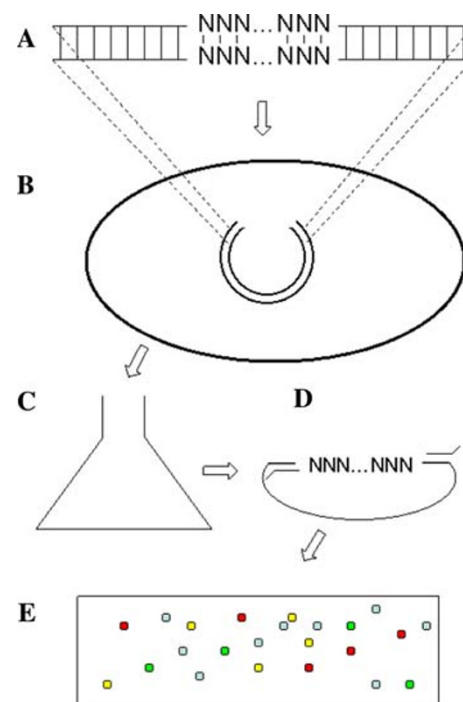


Fig. 4 Possible high throughput mutant construction and analysis. Only minor modifications to the procedures for mutant construction and analysis used in this work are required to greatly increase throughput and compatibility with ultrahigh throughput sequencing methods. **(A)** Random mutagenesis of portion of gene using mutagenic crossover PCR, error prone PCR, annealing degenerate single-stranded synthetic oligonucleotides, etc.; length of mutant portion of gene limited by sequencing method employed. **(B)** Construction of growth competition-ready mutant strain library using gap repair cloning. **(C)** Pooled mutant growth competition. **(D)** Mutant segment isolation and sequencing preparation using PCR. **(E)** Chip-based ultrahigh throughput sequencing

the control of a variable strength promoter such as CUP1 (Robinson et al. 1996) or GAL1 (Inga et al. 2002) or under the control of an optimized fixed strength promoter. A second parameter of the assay that might be optimized is the rate of reporter molecule production which allows growth in the presence of active p53. Several options are available to this end. The most attractive would involve switching the reporter from the ADE2 system to a HIS3 reporter. The His3p enzyme is stoichiometrically inhibited by 3-amino-1,2,4-triazole (AT) (Bitter et al. 2002). Therefore, varying amounts of AT could be used to modulate His3p levels to concentrations appropriate to discriminate different levels of activity in p53 mutants. Alternatively, different p53 response elements could be utilized in the reporter construct which result in different levels of reporter transcription (Tomso et al. 2005; Inga et al. 2002; Campomenosi et al. 2001; Resnick and Inga 2003).

As alluded to earlier, our growth competition methodology was designed to utilize recently reported ultrahigh throughput sequencing technologies. Our polony based method presents a significant advantage in throughput relative to standard screens such as the red/white colony plate screen utilized in this work. However, the physical size of polonies limits the number of data points per polony slide to approximately 1000. As a result, the number of different mutants that can be put into a single competition and the concentration range of individual mutants in the

competition that can be accurately measured are limited. Furthermore, the inability to directly sequence multinucleotide runs of polony DNA requires that mutants (a) be identified using the tagging method that we describe and (b) limits the throughput of mutant expression vector construction. The method reported by Margulies et al. (2005) would allow an increase in the number of data points per sequencing of approximately 10^2 – 10^3 fold and direct sequencing of approximately 100 bp. Therefore it would be possible to assess the impact of mutations at any position in a typical gene in approximately 10 sequencing runs. The method reported by Shendure et al. (2005) would allow an increase of approximately 10^4 – 10^5 fold points per run (with respect to polonies) but is somewhat more limited in sequence read length (approximately 26 bp for this application). With modest modifications to the system we describe here and use of high throughput sequencing technologies, whole gene mutation analysis could be readily accomplished thus helping to elucidate the functional consequences of SNPs (Fig. 4).

Appendix

Oligonucleotide primers used in this report. All primers were supplied salt-free (Operon). [5Acrd] denotes a 5' acrydite modification. [phosp] denotes added phosphate group.

Name	Mutation		Sequence
279	Leu330	Ala	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGGGCGGTGAAATATTCTCCATCCA
273	Leu330	Arg	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGNNNGGTGAAATATTCTCCATCCA
273	Leu330	Asn	See Above
280	Leu330	Asp	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGATCGGTGAAATATTCTCCATCCA
281	Leu330	Cys	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGGCAGGTGAAATATTCTCCATCCA
282	Leu330	Gln	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGCTGGGTGAAATATTCTCCATCCA
283	Leu330	Glu	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGCTCGGTGAAATATTCTCCATCCA
284	Leu330	Gly	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGGCCGGTGAATATTCTCCATCCA
273	Leu330	His	See Above
273	Leu330	Ile	See Above
–	Leu330	Leu	–
285	Leu330	Lys	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGCTTGGTGAATATTCTCCATCCA
286	Leu330	Met	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGCATGGTGAATATTCTCCATCCA
287	Leu330	Phe	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGGAAGGTGAAATATTCTCCATCCA
273	Leu330	Pro	See Above
273	Leu330	Ser	See Above
273	Leu330	Thr	See Above
288	Leu330	Trp	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGCCAGGTGAAATATTCTCCATCCA
289	Leu330	Tyr	TCGGAACATCTCGAAGCGCTCACGCCCACGGATCTGATAGGTGAAATATTCTCCATCCA

Appendix continued

Name	Mutation		Sequence
273	Leu330	Val	See Above
290	Gln331	Ala	TCGGAACATCTCGAAGCGCTCACGCCACGGAT GG CAAGGGTGAAATATTCTCCATCCA
274	Gln331	Arg	TCGGAACATCTCGAAGCGCTCACGCCACGGAT NN NAAGGGTGAAATATTCTCCATCCA
291	Gln331	Asn	TCGGAACATCTCGAAGCGCTCACGCCACGGAT GT TAAGGGTGAAATATTCTCCATCCA
292	Gln331	Asp	TCGGAACATCTCGAAGCGCTCACGCCACGGAT AT CAAGGGTGAAATATTCTCCATCCA
293	Gln331	Cys	TCGGAACATCTCGAAGCGCTCACGCCACGGAT GC AAAGGGTGAAATATTCTCCATCCA
–	Gln331	Gln	–
274	Gln331	Glu	See above
294	Gln331	Gly	TCGGAACATCTCGAAGCGCTCACGCCACGGAT GC CAAGGGTGAAATATTCTCCATCCA
295	Gln331	His	TCGGAACATCTCGAAGCGCTCACGCCACGGAT GT GAAGGGTGAAATATTCTCCATCCA
274	Gln331	Ile	See above
274	Gln331	Leu	See above
296	Gln331	Lys	TCGGAACATCTCGAAGCGCTCACGCCACGGAT CT TAAGGGTGAAATATTCTCCATCCA
297	Gln331	Met	TCGGAACATCTCGAAGCGCTCACGCCACGGAT CA TAAGGGTGAAATATTCTCCATCCA
298	Gln331	Phe	TCGGAACATCTCGAAGCGCTCACGCCACGGAT GA AAAGGGTGAAATATTCTCCATCCA
274	Gln331	Pro	See above
274	Gln331	Ser	See above
274	Gln331	Thr	See above
274	Gln331	Trp	See above
299	Gln331	Tyr	TCGGAACATCTCGAAGCGCTCACGCCACGGAT AT AAAGGGTGAAATATTCTCCATCCA
274	Gln331	Val	See above

Appendix

Name	Sequence	Purpose
272	AAACACAAATACACACACTAATCTAATGGAGGAGCCGCAGTCAGATCCTAGCGTC	Common primer for p53, 5' end
271	TAAATTAATACTTCTATAGACACGCAAAACACAAATACACACACTAATCTAATG	Mutant p53 amplification and
278	CAGAGCTTGTGGGGGTTACCCACTTGTAGGTGCCCTCATACTGGTCAGTC	addition of restriction sites
276	CGTGGGCGTGAGCGCTTCGAGATGTTCCGA	Primers for p53, 3' end
277	CACTTGTAGGTGCCCTCATACTGGTCAGTCTGAGTCAGGCCCTTCTGTCTTGAAC	
221	CCAGTATGAGGGCACCTACAAGTGGGTGAACCC	DNA sequence tag synthesis
222	CAAGCTTTCAGAGCTTGTGGGGGTTACCCACTTGTAGGTGCCCTCATACTGG	
223	CACAAGCTCTGAAAGCTTGAGTAACGGGTCTTGTTCGC	
224	GCACTGCACTGGTGACCGGCGAACAAGACCCGTTAC	
225	GGTCACCAGTGCAGTGCTGCGTCTCACGGACTTC	
226	GAGTACGAGGTGATCTCCGGAAGTCCGTGAAGACGC	
227	GGAGTACCTCGTACTCTGATTGCTGTGCAGCTCAC	
228	AGAGACCTAACAGTAGGGAAACTGTGAGCTGCACAGCAATC	
229	GTTTCCCTACTGTTAGGTCTCTCCATGCTACACTCGTCGACAGCTGA	
230	TCAGCTGTGACGAGTGTAGCATGG	
238	CACAAGCTCTGAAAGCTTGTGTAACGGGTCTTGTTCGC	
239	GCACTGCACTGGTGACCTGCGAACAAGACCCGTTAC	
240	GGTCACCAGTGCAGTGCCGCGTCTCACGGACTTC	
241	GAGTACGAGGTGATCTCCTGAAGTCCGTGAAGACGC	
242	GGAGTACCTCGTACTCAGATTGCTGTGCAGCTCAC	
243	AGAGACCTAACAGTAGGGAAACAGTGAGCTGCACAGCAATC	

Appendix continued

Name	Sequence	Purpose
175	CAACCGCCTCTTCTACCTGG	Colony PCR with 230
249B	GGGGACTAGTATGGAGGAGCCGCAGTCAGATC	p53 isolation
249C	GGGGAAGCTTTCAGTCTGAGTCAGGCCCTT	
327	CTACATGTGTAACAGTTTCTGCATGGGC	p53 mRNA expression analysis
328	[5Acrid]TTCTTTGGCTGGGGAGAGGAGC	
329	AGGGAGCACTAAGCGAGCAC	
333	AAACAATTATCGCTGGAGCTGGTGGG	ADE2 mRNA expression analysis
334	[5Acrid]AAGCGCCAAGCAGTCTGACAGC	
335	CACCACTTCTGTATCGGTGTG	
325	CTCACTCCAGCCACCTGAAGTCCAA	Sequence tag polony amplification
326	[5Acrid]GAGTGATGATGGAGAGACCTAACAGTAG	
324	GGCCTGACTCAGACTGAAAGCTTG	Sequence tag single base extension
252	GTAACGGGTCTTGTTCGC	
253	GGTCACCAGTGCAGTGC	
254	GCGTCTTACGGACTTC	
255	GGAGATCACCTCGTACTC	
256	GATTGCTGTGCAGCTCAC	

References

- Bitter GA, Schaeffer TN, Ellison AR (2002) Reporter gene regulation in *Saccharomyces cerevisiae* by the human p53 tumor suppressor protein. *J Mol Microbiol Biotechnol* 4(6):539–550
- Campomenosi P, Monti P, Aprile A, Abbondandolo A, Frebourg T, Gold B, Crook T, Inga A, Resnick MA, Iggo R et al (2001) p53 mutants can often transactivate promoters containing a p21 but not Bax or PIG3 responsive elements. *Oncogene* 20(27):3573–3579
- Chappuis PO, Estreicher A, Dieterich B, Bonnefoi H, Otter M, Sappino AP, Iggo R (1999) Prognostic significance of p53 mutation in breast cancer: frequent detection of non-missense mutations by yeast functional assay. *Int J Cancer* 84(6):587–593
- Chene P (2001) The role of tetramerization in p53 function. *Oncogene* 20(21):2611–2617
- Chene P, Bechter E (1999) Cellular characterisation of p53 mutants with a single missense mutation in the beta-strand 326–333 and correlation of their cellular activities with in vitro properties. *J Mol Biol* 288(5):891–897
- Flaman JM, Frebourg T, Moreau V, Charbonnier F, Martin C, Chappuis P, Sappino AP, Limacher JM, Bron L, Benhattar J et al (1995) A simple P53 functional assay for screening cell-lines, blood, and tumors. *Proc Natl Acad Sci USA* 92(9):3963–3967
- Hernandez-Boussard T, Rodriguez-Tome P, Montesano R, Hainaut P (1999) IARC p53 mutation database: a relational database to compile and analyze p53 mutations in human tumors and cell lines. *Hum Mutat* 14(1):1–8
- Inga A, Storici F, Darden TA, Resnick MA (2002) Differential transactivation by the p53 transcription factor is highly dependent on p53 level and promoter target sequence. *Mol Cell Biol* 22(24):8612–8625
- Jia LQ, Osada M, Ishioka C, Gamo M, Ikawa S, Suzuki T, Shimodaira H, Niitani T, Kudo T, Akiyama M et al (1997) Screening the p53 status of human cell lines using a yeast functional assay. *Mol Carcinog* 19(4):243–253
- Kern SE, Kinzler KW, Bruskin A, Jarosz D, Friedman P, Prives C, Vogelstein B (1991) Identification of P53 as a sequence-specific Dna-binding protein. *Science* 252(5013):1708–1711
- Ko LJ, Prives C (1996) p53: Puzzle and paradigm. *Genes Dev* 10(9):1054–1072
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921
- Levine AJ, Wu MC, Chang A, Silver A, Attiyeh EF, Lin J, Epstein CB (1995) The spectrum of mutations at the p53 locus. Evidence for tissue-specific mutagenesis, selection of mutant alleles, and a “gain of function” phenotype. *Ann NY Acad Sci* 768:111–128
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057):376–380
- Masood E (1999) US firm’s bid to sequence rice genome causes stir in Japan ... as consortium plans free SNP map of human genome. *Nature* 398(6728):545–546
- Merritt J, Edwards JS (2004) Assaying gene function by growth competition experiment. *Metab Eng* 6(3):212–219
- Merritt J, DiTonno JR, Mitra RD, Church GM, Edwards JS (2003) Parallel competition analysis of *Saccharomyces cerevisiae* strains differing by a single base using polymerase colonies. *Nucleic Acids Res* 31(15):e84
- Merritt J, Butz JA, Ogunnaike BA, Edwards JS (2005) Parallel analysis of mutant human glucose 6-phosphate dehydrogenase in yeast using PCR colonies. *Biotechnol Bioeng* 92(5):519–531
- Mikkilineni V, Mitra RD, Merritt J, DiTonno JR, Church GM, Ogunnaike B, Edwards JS (2004) Digital quantitative measurements of gene expression. *Biotechnol Bioeng* 86(2):117–124
- Mitra R, Church G (1999) In situ localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Res* 27(24):e34

- Mitra RD, Shendure J, Olejnik J, Edyta Krzymanska O, Church GM (2003) Fluorescent in situ sequencing on polymerase colonies. *Anal Biochem* 320(1):55–65
- Resnick MA, Inga A (2003) Functional mutants of the sequence-specific transcription factor p53 and implications for master genes of diversity. *Proc Natl Acad Sci USA* 100(17):9934–9939
- Robinson AS, Bockhaus JA, Voegler AC, Wittrup KD (1996) Reduction of BiP levels decreases heterologous protein secretion in *Saccharomyces cerevisiae*. *J Biol Chem* 271(17):10017–10022
- Rollenhagen C, Chene P (1998) Characterization of p53 mutants identified in human tumors with a missense mutation in the tetramerization domain. *Int J Cancer* 78(3):372–376
- Scharer E, Iggo R (1992) Mammalian P53 can function as a transcription factor in yeast. *Nucleic Acids Res* 20(7):1539–1545
- Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309(5741):1728–1732
- Sherman F (1991) Getting started with yeast. *Method Enzymol* 194:3–21
- Tomso DJ, Inga A, Menendez D, Pittman GS, Campbell MR, Storici F, Bell DA, Resnick MA (2005) Functionally distinct polymorphic sequences in the human genome that are targets for p53 transactivation. *Proc Natl Acad Sci USA* 102(18):6431–6436
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA et al (2001) The sequence of the human genome. *Science* 291(5507):1304–1351