Nonmonotonic Pathway Gene Expression Analysis Reveals Oncogenic Role of p27/Kip1 at Intermediate Dose

Hien H Nguyen¹, Susan C Tilton²,³, Christopher J Kemp² and Mingzhou Song¹

¹Department of Computer Science, New Mexico State University, Las Cruces, NM, USA. ²Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. ³Department of Environmental and Molecular Toxicology, Oregon State University, Corvallis, OR, USA.

ABSTRACT: The mechanistic basis by which the level of p27Kip¹ expression influences tumor aggressiveness and patient mortality remains unclear. To elucidate the competing tumor-suppressing and oncogenic effects of p27Kip¹ on gene expression in tumors, we analyzed the transcriptomes of squamous cell papilloma derived from Cdkn1b nullizygous, heterozygous, and wild-type mice. We developed a novel functional pathway analysis method capable of testing directional and nonmonotonic dose response. This analysis can reveal potential causal relationships that might have been missed by other nondirectional pathway analysis methods. Applying this method to capture dose-response curvatures in papilloma gene expression data, we show that several known cancer pathways are dominated by low-high low gene expression responses to increasing p27 gene doses. The oncogene cyclin D1, whose expression is elevated at an intermediate p27 dose, is the most responsive gene shared by these cancer pathways. Therefore, intermediate levels of p27 may promote cellular processes favoring tumorigenesis—strikingly consistent with the dominance of heterozygous mutations in CDKN1B seen in human cancers. Our findings shed new light on regulatory mechanisms for both pro- and anti-tumorigenic roles of p27Kip¹. Functional pathway dose-response analysis provides a unique opportunity to uncover nonmonotonic patterns in biological systems.

KEYWORDS: p27, functional pathway analysis, tumorigenesis, nonmonotonic patterns

RECEIVED: June 24, 2017. ACCEPTED: September 16, 2017.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work is partially supported by US NIH NCi Partnership for the Advancement of Cancer Research, NMSU/Fred Hutch grants US4CA132838 (NISSU) and US4CA132838 (Fred Hutch), US NIH New Mexico IDEA Networks of Biomedical Research Excellence Grant 2P20GM103451-14, US NIH Mountain West Clinical Translational Research Grant 1U01GM104944-2, US NSF Advances in Biological Informatics Grant 1661331, and US NIH grant U01 CA176303.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHORS: Christopher J Kemp, Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA. Email: cj.kemp@fhcrc.org
Mingzhou Song, Department of Computer Science, New Mexico State University, Las Cruces, NM 88003, USA. Email: jsong@cs.nmsu.edu

Introduction

In many human cancers, the expression level of the cyclin-dependent kinase (CDK) inhibitor p27Kip¹ is strongly associated with prognosis and patient mortality.¹ Although p27 is a CDK inhibitor, it also has cytoplasmic functions and regulates cell migration.²,³ As death from cancer is primarily due to metastatic spread of the disease, this suggests that p27 could contribute to patient mortality by affecting invasion and metastatic dissemination. However, the mechanistic basis by which p27 influences tumor aggressiveness and patient mortality remains unclear. p27 can affect cell proliferation,⁴ migration,²,³ apoptosis,⁵–⁷ and differentiation,⁸,⁹ through both CDK-dependent and CDK-independent mechanisms.¹⁰,¹¹ However, the relative contributions of these processes to cancer outcome are unknown. In addition, p27 was the first tumor suppressor gene shown to be haploinsufficient, a genetic concept that is now widely accepted.¹²,¹³ Haploinsufficiency in this case means the loss of only 1 allele predisposes to tumor development. This property of p27 makes the study of its phenotypic effects even more complicated. Indeed, the low expression of p27 was widely observed to correlate with poor prognosis in various cancers, including prostate,¹⁴ breast,¹⁵ colorectal,¹⁶,¹⁷ and ovarian cancers.¹⁸ However, when its interaction with the cell cycle is disabled by deleting its CDK inhibitory domain, p27 has oncogenic properties, causing multiple tumor phenotypes by modulating prometastatic factors such as cell migration and invasion.¹⁹,²⁰ Increased motility and invasiveness are essential processes for cancer cell metastatic dissemination.²¹ However, findings of p27’s roles on cell migration and invasion have been inconsistent. High expression levels of p27 were reported to promote migration in cerebral cortex²² and more recently in bladder cancer cells.²³ Meanwhile, low expression of p27 is associated with invasion in neuroblastoma, bone marrow,²³ and early cervical carcinoma.²⁴ A non-monotonic function of p27 was first proposed for its role in cell proliferation²⁵ but not yet for cell motility. Furthermore, the positive relation of p27 to cell migration and its negative relation to invasion suggest an intermediate dose of p27 at which both cell migration and invasion could be active. This hypothesis motivated us to analyze gene expression in chemically induced mouse squamous cell papillomas derived from p27 null, heterozygous, and wild-type mice to capture potential nonmonotonic effects not feasible with a 2-dose...
experiment. We characterized the transcriptomic responses using a novel informatics method for functional pathway analysis.

Pathway analysis, studying gene activities collectively, provides statistical robustness over single-gene approaches. Many analysis tools are effective at ranking pathways by their responses to experimental conditions. These approaches can be grouped into 3 categories: overrepresentation analysis (ORA), functional class scoring (FCS), and topology based. Addressing several limitations of ORA—an attempt in early days, the FCS approaches, including gene set enrichment analysis and its variations, compute a pathway statistic cumulative over all gene-level statistics. However, recent topology-based pathway analysis methods integrate gene interaction information to understand pathway rewiring but require larger sample sizes than ORA or FCS. These pathway analysis approaches, however, were designed to determine the association between pathway responses and experimental conditions, not optimal for the complex functional dependency of pathway responses to dosage. Indeed, nonmonotonic dose-response curves are widely observed in biology such as in studies of hormones and endocrine-disrupting chemicals.

To overcome the limitations of current pathway analysis methods for detecting nonmonotonic dose–response relationships, we present a novel method called functional pathway dose-response analysis to test whether pathway response is a mathematical function of gene dosage, without assuming a specific underlying functional form. This method is based on the recent FunChisq method which detects nonconstant and nonparametric functional dependency among random variables for causal inference. FunChisq demonstrated outstanding empirical effectiveness in the recent HPN-DREAM Breast Cancer Network Inference Challenges. A gene is functionally dependent on stimulant dosage if the gene’s expression value is uniquely determined by the dose. We show that our method is mathematically optimal in detecting functional dependency of pathway response to stimulant doses. We also found that the gamma distribution approximation of the null pathway response statistic is more accurate than the χ² distribution. Next, we applied FunChisq and functional pathway dose-response analysis to characterize how genes and pathways respond to Cdkn1b gene dosage in mouse papillomas derived from 3 genotypes: wild type (p27+/+), heterozygous Cdkn1b gene dosage in mouse papillomas and 17 pathways critical in cancer. The pathways involve a total of 18 443 genes. Among them are 2 interesting overlapped subsets: 14 pathways known to be affected by p27 dosage influencing many aspects of tumorigenesis. Therefore, complementing posttranscription modification and cellular localization explanations of p27’s oncogenic roles, our findings provide new insights toward p27 function and molecular mechanisms for gene regulation in tumor development.

Materials and Methods

Dose–response gene expression data and KEGG pathways

We collected samples of DMBA/TPA-induced skin papillomas (>5 mm) from C57BL/6J NIH F2 mice of 3 p27 genotypes: (1) nullizygous mutation (p27−/−), corresponding to a low dose; (2) heterozygous mutation for a null allele (p27+/−), corresponding to an intermediate dose; and (3) wild type (p27+/+), corresponding to a high dose. Four papillomas from each p27 genotype were examined. Whole-genome microarray analysis was performed for the samples using Illumina MouseRef-8 v2.0 Expression BeadChips (Illumina, San Diego, CA, USA; 25 600 probe sets). Total RNA was isolated with Trizol reagent (Invitrogen Life Technologies, Carlsbad, CA, USA) and cleaned with RNeasy Kit (Qiagen, Valencia, CA, USA). RNA integrity and purity were assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Illumina BeadChips were hybridized with 1 µg of labeled complementary RNA, washed, stained, and scanned (Illumina BeadStation Scanner) according to Illumina’s protocols. Data values with detection scores were compiled using BeadStudio (Illumina) and raw intensity data were quantile normalized by Robust Multi-Array Analysis summarization for statistical analysis.

We extracted 281 Mus musculus functional pathways and disease processes from KEGG. Among them are 2 interesting overlapped subsets: 14 pathways known to be affected by p27 and 17 pathways critical in cancer. The pathways involve a total of 18 443 genes.

The functional chi-square test for single-gene dose response

The functional χ² (FunChisq) test determines whether discrete random variable Y is a nonconstant function f : X → Y of discrete random variable X. Let r and s be the numbers of discrete levels of X and Y, respectively. We construct an r × s contingency table (ni,j) from observed data of X and Y, where X and Y are row and column variables, respectively. Let n be the sample size. The functional χ² statistic is defined as follows:

\[ \chi^2_{f/\chi \\rightarrow Y} = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(n_{i,j} - n_{i/s})^2}{n_{i/s}} - \sum_{j=1}^{s} \frac{(n_{s,j} - n/s)^2}{n/s} \]  \hspace{1cm} (1)
and it follows a $\chi^2$ distribution with $(r-1)(s-1)$ degrees of freedom under the null hypothesis that $X$ and $Y$ are statistically independent with the assumption that $Y$ is uniformly distributed. Important properties of the test statistic, such as its null distribution, optimality for functional dependency, and asymmetry with respect to $X$ and $Y$, were established previously.23

**Functional pathway dose-response analysis**

Let $P$ be the set of genes on a given pathway and $P$ also serves as an identifier of the pathway. To assess the functional dependency of responses of all genes on pathway $P$ on dosage $X$, we design a novel pathway dose-response statistic $\chi^2_{F,X,p}$ by summing up the functional $\chi^2$ statistics of member genes in the pathway:

$$\chi^2_{F,X,p} = \sum_{g \in P,} \chi^2_{F,X,g}$$

where $F$ is the collection of dose-response functions for each gene in the pathway. We do not assume the same dose-response curve for different genes in the pathway. This is important for pathways that include genes which respond to the stimulus either positively or negatively. Based on the optimality of the functional $\chi^2$ statistic for functional dependency,23 we derive the following theorem to establish the optimality of the pathway dose-response statistic.

**Theorem 1. (Optimality of pathway dose-response statistic.)** The pathway dose-response statistic $\chi^2_{F,X,p}$ is maximized when the responses of each gene $g$ in a pathway $P$ are nonconstant empirical functions of dosage $X$ represented by function set $F$ and is minimized to 0 when all gene responses are empirically independent of dosage.

A proof of the theorem is provided in Supplementary Note.

If the expression levels of each gene in a pathway are statistically independent, then $\chi^2_{F,X,p}$ follows a $\chi^2$ distribution under the null hypothesis of gene expression in the pathway being independent of dosage. The degrees of freedom of $\chi^2_{F,X,p}$ are the sum of those of $\chi^2_{F,g}$ over all $g$'s in $P$.

Because the member genes in a biological pathway are biochemically connected, their expression levels cannot be assumed independent in general. This implies that $\chi^2_{F,X,p}$ does not necessarily follow a $\chi^2$ distribution under the null hypothesis. To approximate the null distribution more accurately than the $\chi^2$ distribution, we developed a gamma approximation for the null distribution of $\chi^2_{F,X,p}$ using a bootstrap strategy. We permute samples instead of genes to maintain dependency among genes and randomly assign the permuted samples to stimulant doses, such that the experimental design is the same with the original except the swapped samples. For the 12 samples of transcriptome responding to 3 p27 doses, each permutation is a random assignment of 4 samples to each dose. The gamma distribution family $\{\Gamma(\nu, \sigma)\}$ of shape parameter $\nu$ and scale parameter $\sigma$ contains probability density functions in the form of

$$p(x | \nu, \sigma) = \frac{1}{\Gamma(\nu)\sigma^\nu} x^{\nu-1} e^{-x/\sigma} \quad (k \geq 0, \sigma \geq 0)$$

where the gamma function $\Gamma(k)$ is defined as follows:

$$\Gamma(k) = \int_0^\infty t^{k-1} e^{-t} dt$$

We first obtain the estimated mean $\hat{\mu}$ and variance $\hat{\sigma}^2$ of $\chi^2_{F,X,p}$ by bootstrapping. Then, we calculate the shape and scale parameters by $k = \hat{\mu}^2 / \hat{\sigma}^2, \hat{\theta} = \hat{\sigma}^2 / \hat{\mu}$. From the gamma distribution with the estimated parameters, we compute a $P$ value to indicate the statistical significance of $\chi^2_{F,X,p}$. In addition, even when $\chi^2_{F,X,p}$ is indeed $\chi^2$ distributed, gamma approximation is still applicable because $\chi^2$ distributions are a special case in the gamma distribution family.

**Pathway pattern enrichment analysis**

In each pathway, the member genes can differ greatly in responding to p27. We classify the responses into 4 patterns: increasing (INC), decreasing (DEC), high-low-high (HLH), and low-high-low (LHL). The enrichment analysis is to tell which pattern is overrepresented in a pathway, regarding the proportion of that pattern in the whole gene set. Let $N$ be the total number of genes in the gene pool; $K_p$ be the number of genes from the gene pool which respond in pattern $P$; $n$ be the number of genes in the analyzed pathway, and $k_p$ be the number of the genes in the pathway which respond in pattern $P$. Then, the probability of $k_p$ in a pathway is defined by

$$b_p(k_p; N, K_p, n) = \binom{K_p}{k_p} \binom{N-K_p}{n-k_p}$$

where $k_p$ follows a hypergeometric distribution if the pathway is equally enriched in pattern $P$ as in the entire gene pool.

**Results**

**Overview of the functional pathway dose-response analysis method**

Figure 1 illustrates functional pathway dose-response analysis, using the autophagy pathway as an example. The input includes a
list of genes on the pathway and gene expression data that were first discretized. The test statistic for a pathway dose-response is computed as the summation of gene-level functional \( \chi^2 \) statistics. Finally, the output of the analysis, the statistical significance of pathway response to stimulant dosage, is calculated based on an asymptotic gamma distribution whose parameters are estimated by bootstrapping through permutation.

**Advantage of functional pathway dose-response analysis by simulation study**

To validate the effectiveness of functional pathway dose-response analysis, we benchmarked it against an alternative method that replaces functional \( \chi^2 \) by the Pearson association \( \chi^2 \). The Pearson \( \chi^2 \) is a representative of many pathway statistics in the sense that it measures association but not mathematical functional dependency in a contingency table. We simulated pathway gene expression data modeled after 17 known cancer pathways from KEGG, at 3 levels of stimulant dosage in 4 biological replicates. The percentage of dose-responding genes in a pathway was controlled at 10%, 50%, and 90%. We also created null pathways whose genes do not functionally respond to stimulant dosage. Noise was generated using a house noise model at the noise levels of 20%, 30%, and 40%. Figure 2 shows areas under the receiver operating characteristic (ROC) and precision-recall curves of the 2 pathway \( \chi^2 \) methods under all 9 configurations of response percentages and noise levels. The actual ROC and precision-recall curves for each configuration are shown in Supplementary Figures S1 and S2. Both methods improved when more genes on the pathways are responsive to the stimulus. It is evident that the functional pathway dose-response analysis performs consistently better at noise levels of 20% and 30%. Although the performance of the 2 methods is almost comparable at high noise levels of 40%, functional pathway dose-response analysis consistently outperformed nonfunctional Pearson \( \chi^2 \) test under all settings.

**Genes with expression patterns most responsive to p27 dosage**

We first analyzed the papilloma gene expression data set for single-gene dose-response using FunChisq. The expression of each gene was quantized into 4 levels, using R package “Ckmeans.1d.dp” relying on dynamic programming to guarantee optimal clustering for sample points of each gene. Then, we generated 3 × 4 contingency tables with each gene’s discrete expression level as the column variable and p27 gene dose as the row variable. Then, the FunChisq test was applied to identify genes which are most functionally responsive to p27 dosage. To correct for the multiple testing effect, we control the false discovery rate (FDR) using R package “qvalue” while adjusting the FunChisq \( P \) values. When FDR is controlled at...
10%, 21 genes are significantly responsive to p27 dosage. Figure 3 illustrates monotonic and nonmonotonic expression patterns of 16 genes with respect to p27 dosage.

Table 1 shows 21 genes with the most significant FDR-adjusted P values, response patterns, and the pathways to which each gene belongs. We categorize the expression responses of each gene to 3 increasing p27 doses into 4 patterns: increasing (INC), decreasing (DEC), low-high-low (LHL), and high-low-high (HLH). We call the former 2 monotonic and the latter 2 nonmonotonic response patterns. The criterion to assign the expression responses of each gene to a pattern is based on its median expression level under each dose. If the medians are strictly increasing or decreasing as p27 dosage increases, the pattern is INC or DEC, respectively. Otherwise, the pattern is either LHL or HLH, if the median at the p27 intermediate dose is either the highest or the lowest among the 3 medians of responses to each p27 dose. It is worth pointing out that while either the correlation or Pearson χ² test can detect INC and DEC patterns, they cannot capture LHL or HLH patterns.

Of these 21 genes, 2 are involved in guanosine triphosphatase (GTPase) signaling, and the GTPase HRAS is the major oncogenic driver of DMBA/TPA-induced papillomas41,42 and other tumor types. These genes also belong to diverse signaling pathways with oxytocin signaling and arachidonic acid metabolism being most prominent. This demonstrates a remarkable range of p27-driven effects on biological processes beyond cell cycle regulation.

**Figure 2.** Advantage of functional pathway dose-response analysis over alternative nonfunctional methods. Data were simulated using 17 known cancer pathways in KEGG in response to 3 doses of a hypothetical stimulus at 4 biological replicates. The percentage of dose-responding genes in a pathway was controlled at 10%, 50%, and 90%, and the noise levels were 20%, 30%, and 40%. Areas under (left) ROC and (right) precision-recall curves were calculated for pathway analysis using functional χ² versus nonfunctional Pearson χ² on the simulated data. ROC indicates receiver operating characteristic.

**Genes in cancer pathways nonmonotonically respond to p27 dosage**

Although single-gene dose-response analysis has highlighted individually responsive genes, we extended the analysis to biologically functional pathways to provide a comprehensive map for the influence of p27 on cancer gene networks with improved statistical power. Therefore, we performed functional dose-response analysis on 281 KEGG pathways with the p27 dose-response transcriptome data set and identified 198 statistically significantly dose-responsive pathways and disease processes at an FDR of 10% (meaning about 20 pathways—10% of 198—might be false positives). The complete ranked list of all 198 pathways and disease processes with pathway dose-response significance P values is provided in Supplementary Table S1. The P values were estimated by gamma distributions. Figure 4 shows that gamma approximation becomes more accurate than the χ² distribution after a few iterations of permutation and quickly converges to the true distribution afterward. Gamma approximation also converges faster to the true distribution than does the permutation test based on resampling.

For each pathway, we also computed the number of member genes which responded in INC, DEC, HLH, and LHL patterns, respectively. We were interested in whether a certain pattern is enriched in that pathway. With the known numbers of genes in each pattern and in total, we computed the probability that a pattern is overrepresented in a pathway using the hypergeometric distribution for the null hypothesis that a pattern is equally represented in the pathway versus in the entire collection of pathways.

Table 2 shows the results for 14 pathways and disease processes known to be affected by p27 from the literature. We can detect all 14 at an FDR of 10%. This provides additional evidence to simulation studies to support the effectiveness of the functional pathway analysis. The overrepresentation statistics show that the LHL pattern is significantly enriched in 2 pathways of cell cycle and microRNAs in cancer, the DEC pattern is enriched in PI3K-Akt signaling pathway, and both HLH
Cancer Informatics

and LHL are significantly enriched ($P < .05$) in the Epstein-Barr virus infection pathway.

Table 3 lists the response patterns and statistical significance of 17 KEGG pathways known to play important roles in cancer. Of the 17 cancer-related pathways, 16 are responsive to p27 gene dosage at 10% FDR. Figure 5 presents a comprehensive view of cancer gene networks with these pathways highlighted for their connections to p27. The networks include important pathways and genes that contribute to cancer hallmarks such as sustained angiogenesis, evading apoptosis, tissue invasion, and metastasis. We note that there are 2 ovals of p27 representing nuclear and cytoplasmic p27, respectively. Cytoplasmic p27, likely to have CDK-independent functions, affects several signaling pathways. It has direct interactions with the PI3K-Akt signaling and mTOR signaling pathways (Figure 5), which mediate sustained angiogenesis and evading apoptosis. Via these 2 pathways, p27 has indirect interactions with the Jak-STAT signaling pathway, the ErbB signaling pathway, and, in turn, the cytokine-cytokine receptor interaction pathway. Another indirect interaction is with the focal adhesion and adherens junction pathways, essential for cell migration and metastasis.

Table 3 shows 4 pathways enriched for LHL, 3 for DEC, 1 for HLH, and none for INC patterns. We also compared the enrichment statistics among the patterns in the same pathway to reveal the most enriched (smallest $P$ value) pattern. As a result, 10 cancer pathways have LHL as the most enriched, 4 pathways have HLH, 3 pathways have DEC, and none has INC as the most enriched. Evidently, the LHL response pattern is dominant in cancer pathways.

We took a further step to examine the member genes of the cancer pathways. Without prior expectation, Cyclin D1...
Table 1. Top individual genes that are most responsive to p27 dosage.

| SYMBOL      | OFFICIAL GENE NAME                        | FDR-ADJUSTED P VALUE | RESPONSE | INVOLVED PATHWAYS                                                                 |
|-------------|-------------------------------------------|----------------------|----------|-----------------------------------------------------------------------------------|
| GPC3        | Glypican 3                                | .0172                | DEC      | Proteoglycans in cancer                                                            |
| EEF2        | Eukaryotic translation elongation factor 2| .0172                | HLH      | AMPK signaling pathway, oxytocin signaling pathway                                 |
| LMNA        | Lamin A/C                                 | .0172                | LHL      | Dilated cardiomyopathy                                                             |
| ADCY4       | Adenylate cyclase 4                       | .0172                | DEC      | Purine metabolism, dilated cardiomyopathy, oxytocin signaling pathway              |
| TNFRSF1A    | Tumor necrosis factor receptor superfamily, member 1A | .0172 | HLH | Apoptosis, tuberculosis, Epstein-Barr virus infection |
| RYR1        | Ryanodine receptor 1 (skeletal)           | .0217                | DEC      | Calcium signaling pathway, oxytocin signaling pathway                             |
| PSMB10      | Proteasome (prosome, macropain) subunit, beta type, 10 | .0478 | LHL | Proteasome                                                                        |
| DHRS9       | Dehydrogenase/reductase (SDR family) member 9 | .0973 | HLH | Retinol metabolism                                                                |
| CYB5R3      | Cytochrome b5 reductase 3                 | .0973                | LHL      | Amino sugar and nucleotide sugar metabolism                                        |
| AP1S1       | Adaptor-related protein complex 1, sigma 1 subunit | .0973 | LHL | Lysosome                                                                          |
| RGL2        | ral guanine nucleotide dissociation stimulator-like 2 | .0973 | INC | Ras signaling pathway                                                             |
| ELOVL1      | ELOVL fatty acid elongase 1               | .0973                | HLH      | Fatty acid elongation                                                              |
| SGPL1       | Sphingosine-1-phosphate lyase 1           | .0973                | INC      | Sphingolipid metabolism, sphingolipid signaling pathway                           |
| CHPT1       | Choline phosphotransferase 1              | .0973                | HLH      | Choline metabolism in cancer                                                       |
| MEF2C       | Myocyte enhancer factor 2C                | .0973                | HLH      | MAPK signaling pathway, oxytocin signaling pathway                                 |
| GIT2        | G protein-coupled receptor kinase-interacting ArfGAP 2 | .0973 | INC | Endocytosis                                                                       |
| HSD11B1     | Hydroxysteroid (11-ß) dehydrogenase 1     | .0973                | DEC      | Steroid hormone biosynthesis, chemical carcinogenesis                             |
| SERPINA1B   | Serine (or cysteine) peptidase inhibitor, clade A, member 1B | .0973 | DEC | Complement and coagulation cascades                                                |
| SAR1B       | Secretion associated, Ras-related GTPase 1B | .0973 | HLH | Protein processing in endoplasmic reticulum                                        |
| CYP4F18     | Cytochrome P450, family 4                 | .0973                | LHL      | Arachidonic acid metabolism                                                       |
| PTGS1       | Subfamily E, polypeptide 3 prostaglandin-endoperoxide synthase 1 | .0973 | HLH | Arachidonic acid metabolism, platelet activation                                   |

Abbreviations: DEC, decreasing; HLH, high-low-high; INC, increasing; LHL, low-high-low.

(Ccnd1)—a cell cycle regulator—emerges as the most responsive gene in several p27-responsive cancer pathways. Human CCND1 is an oncogene which is overexpressed in many tumors, especially in breast cancer. In terms of responding to p27 in this study, Ccnd1 ranks first in cell cycle, p53 signaling, Jak-STAT signaling, and Wnt signaling pathways. It also ranks third in focal adhesion pathway and fourth in the PI3K-Akt signaling pathway. More importantly, the response of Ccnd1 follows an LHL pattern, suggesting that Ccnd1 is active at the intermediate dosage of p27. Figure 6 shows the response of Ccnd1 to p27 dosage, which is consistent with the model proposed for prostate tumorigenesis by Gao et al.25

The dominance of nonmonotonic LHL response is evidence for the intermediate dose of p27 as cancer promoting; at
this dose, a large number of cancer pathways become highly expressed in favor of tumor development. When p27 dose deviates from the intermediate range, the cellular context may be induced to transition from oncogenic to tumor suppressing.

Figure 4. Convergence of gamma approximation to the null distribution in functional pathway dose-response analysis. Each box plot summarizes the deviation of the approximated \( P \) values from the true \( P \) values. The true \( P \) values were computed for all pathways by full permutation of the 12 samples from the papilloma gene expression data. The \( P \) values for the red box plot were computed for all pathways using a theoretical \( \chi^2 \) null distribution without permutation. Although both gamma approximation (green boxes) and random permutation test (yellow boxes) converge to the true \( P \) values after 5000 iterations, gamma approximation reduces errors faster than the random permutation test.

Table 2. Response of 14 pathways and disease processes known to involve p27.

| PATHWAY NAME                      | RESPONSE TO P27 | PATTERN ENRICHMENT | P VALUE | FDR-ADJUSTED \( P \) VALUE | DEC | INC | HLH | LHL |
|----------------------------------|----------------|--------------------|---------|---------------------------|-----|-----|-----|-----|
| With significantly enriched patterns |                |                    |         |                           |     |     |     |     |
| PI3K-Akt signaling pathway       | .0772          | .0119              | .7687   | .8553                     | .4929 |
| Cell cycle                       | .0800          | .9878              | .2272   | .9196                     | .0026 |
| Chronic myeloid leukemia         | .0800          | .7180              | .0504   | .5568                     | .5549 |
| Epstein-Barr virus infection     | .0817          | .9994              | .8413   | .0375                     | .0472 |
| MicroRNAs in cancer              | .0950          | .3719              | .5775   | .9782                     | .0147 |
| Significant in response to p27 at 10% FDR |      |                    |         |                           |     |     |     |     |
| Hepatitis B                      | .0635          | .6704              | .1319   | .7249                     | .3233 |
| Prostate cancer                  | .0635          | .4537              | .2326   | .3973                     | .6476 |
| HIF-1 signaling pathway          | .0646          | .6201              | .0881   | .6257                     | .5163 |
| ErbB signaling pathway           | .0646          | .9792              | .2778   | .4075                     | .0912 |
| Measles                          | .0772          | .8567              | .6647   | .2469                     | .1889 |
| FoxO signaling pathway           | .0800          | .5044              | .5403   | .1180                     | .7416 |
| Transcriptional misregulation in cancer | .0800      | .7843              | .1229   | .8571                     | .1357 |
| Viral carcinogenesis             | .0805          | .9158              | .1447   | .3806                     | .3924 |
| Small cell lung cancer           | .0865          | .8716              | .1505   | .5975                     | .2047 |

Abbreviations: DEC, decreasing; FDR, false discovery rate; HLH, high-low-high; INC, increasing; LHL, low-high-low. Pattern enrichment \( P \) values in bold indicate that the respective patterns are significantly enriched for the pathway or disease process.
Taken together, these results reveal a deeper understanding of the seemingly contradictory roles of p27 in cancer reported so far in the literature.

**Heterozygosity dominates somatic mutations in CDKN1B in human cancers**

To determine the mutational profile of p27 in human cancers, we examined somatic mutations in CDKN1B from data in TCGA (http://cancergenome.nih.gov/), where thousands of tumors and matched normal samples were collected for 30 major cancer types. We call a point mutation heterozygous when exactly 1 allele is different from the reference nucleotide, and a homozygous point mutation refers to both alleles differing from the reference nucleotide. The frequency of observed somatic mutations of CDKN1B in human cancers is shown in Figure 7A. A total of 114 tumor samples carry 15 homozygous and 99 heterozygous mutations in CDKN1B. Notably, all 5 observed CDKN1B mutations are heterozygous in 53 colon adenocarcinoma tumor samples, the highest ratio among all cancers (Figure 7) and in consonance with known correlation between p27 expression and poor prognosis of colorectal cancer.\(^\text{16,17}\) We further compared the occurrence of mutations in CDKN1B with 5 other known cancer-related genes TP53, BRCA1, BRCA2, RB1, and APC. Because these genes have large variations in the total number of mutations observed from the TCGA data sets, we show the percentages of heterozygous mutations over all mutations in tumor samples for each gene in Figure 7B. The bar plots indicate that CDKN1B heterozygosity is the highest among the 6 cancer genes. The increased percentage of CDKN1B heterozygous mutations in tumor samples relative to these genes are all statistically significant, as indicated by the heterogeneity $\chi^2$ test $P$ values in Figure 7C. This result reveals the dominance of CDKN1B heterozygosity in human cancers and is consistent with the intermediate dose of p27 being oncogenic.
We also compared copy number variation (CNV) of CDKN1B in cancer with 2 tumor suppressor genes, PTEN and TP53, as well as 2 oncogenes, MYC and EGFR. The CNV data were extracted from the Catalogue Of Somatic Mutations In Cancer (COSMIC) at cancer.sanger.ac.uk.45 Pair-wise differences between the copy numbers of the 5 genes are statistically significant (P < .001 after Bonferroni adjustment). Figure 8 shows that the copy number ranges of tumor suppressor genes PTEN and TP53 are lower, which is expected because the low expression of these genes correlates with the loss of cancer suppression. On the other side, the copy number ranges of MYC and EGFR are higher, which agrees with their oncogenic role. CDKN1B, once again, has a medium range of copy numbers, consistent with our observation that CDKN1B has oncogenic properties at the intermediate dosage.

**Discussion**

We have presented a novel pathway dose-response analysis method to uncover important cellular processes that might be regulated by p27. This method is unique for pathway analysis because it is functional and nonparametric. Functionality has long been used for causal inference, and nonparametric function dependency enables detection of previously unknown forms of nonmonotonic dose-responses. The pathway-level dose-response statistic, improved on statistical power of single-gene functional $\chi^2$, can identify most responsive pathways functionally and nonparametrically. Once a dose-response relationship has been validated, a parametric function$^{46}$ can be estimated for a more precise characterization of a specific dose-response curve.

Some human tumors show prominent cytoplasmic staining of p27.47 Mouse models show exceptions to strict dose
dependence of tumor suppression by p27, and tissue culture studies show that p27 has numerous cell cycle–independent functions. Together, these previous studies suggest that in addition to its well-known CDK inhibitory/tumor suppressor function, p27 may also have oncogenic activity in the cytoplasm, perhaps related to cell motility and metastasis. However, in the absence of in vivo data, the relevance of this remains unclear. Most critically, we do not know if p27 is prometastatic or antimetastatic, and more generally, we do not know the mechanistic basis for the association between p27 expression patterns in human tumors and patient mortality. This study has revealed, in an in vivo tumor model, that the largest proportion of genes on all known cancer pathways responded nonmonotonically to p27 dosage. Our pathway dose-response analysis not only detects direct links between p27 and a pathway but also suggests potential indirect interactions via cascading pathway responses. Among key cancer pathways (Table 3), our analysis established new links from p27 to 8 of them. Here, a
link is considered new if it has not been specified in KEGG database. These 8 pathways are adherens junction, apoptosis, cytokine–cytokine receptor interaction, focal adhesion, Jak-STAT signaling pathway, MAPK signaling pathway, and vascular endothelial growth factor signaling pathway. These connections support the CDK-independent functions of p27. Moreover, focal adhesion and adherens junction could be the link via which p27 influences cell migration and metastasis. Indeed, p27 is known to affect cell migration and metastasis via cyclin D1 and RhoA signaling pathway.49,50 These links to key cancer pathways suggest a more complex role of p27 in cancer. p27 could be much more active in its CDK-independent function than cell cycle inhibition.

We have yet to propose a mathematical model for the biochemical mechanism that is consistent with the observed nonmonotonic dose–response. In control theory, the functional dependency of other genes on p27 with nonmonotonic patterns is decomposable to monotonic subsystems with negative feedback.51 This is a promising direction for future modeling work. A limitation of this study is that the gene expression data are not specific to protein modifications or subcellular localization. When additional location-specific protein data become available, the presented functional pathway dose-response analysis is readily applicable to deepen the biological findings. Thus, one direction is to explore the connection between the 2 cancer-induced/repressed promoter groups and the abundance and activity of cytoplasmic/nuclear p27 protein.

The presented functional pathway dose–response analysis is a novel method to capture functional dependency of the response of biological systems to stimuli. The main strength of this method is its capability to test for directional influence from dose to response, hence, suggesting potential causal relationships that might have been missed by undirected association analyses. Our study on p27Kip1 shows that it has nonmonotonic effects on expression of genes in almost all cancer pathways beyond its inhibitory role in cell cycle regulation. This complies with findings in the literature that p27 has both tumor suppressing and oncogenic activities, and other data support the generalization of this conclusion to human cancers.

Acknowledgements

The results shown here are in part based on data generated by the TCGA Research Network: http://cancergenome.nih.gov/. We thank Dr. Jun-Yuan Ji for constructive feedback on the manuscript.

Author Contributions

CJK oversaw the mouse tumor studies. SCT performed gene expression analysis. HHN and MS designed the functional pathway analysis method. HHN implemented the method, performed simulation studies, and applied the method to the mouse transcriptome data. All authors contributed to interpretation and manuscript preparation.

Availability of Supporting Data

The software (functional pathway response analysis) and the data set (normalized data set of transcriptome in papillomas of wild-type and p27-mutant mice) supporting the results of this article is available at http://www.cs.nmsu.edu/~joemsong/software/FuncPathResp.

REFERENCES

1. Chu IM, Hengst L, Slingerland JM. The Cdk inhibitor p27 in human cancer: prognostic potential and relevance to anticancer therapy. *Nat Rev Cancer*. 2007;7:23–30.
2. Baldassarre G, Belletti B, Nicoloso MS, et al. p27(Kip1)–statin interaction influences sarcoma cell migration and invasion. *Cancer Cell*. 2005;7:51–63.
3. McAllister SS, Becker-Hapak M, Pintucci G, Pagano M, Dowdy SF. Novel p27kip1 C-terminal scatter domain mediates Ras-dependent cell migration independent of cell cycle arrest. *Mol Cell Biol*. 2003;23:216–228.
4. Polya K, Kato JY, Solomon MJ, et al. p27kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. *Genes Dev*. 1994;8:9–22.
5. Eymin B, Hausq M, Droin N, Sorret O, Dimanche-Boitrel MT, Salery E. p27kip1 induces drug resistance by preventing apoptosis upstream of cytochrome c release and procaspase-3 activation in leukemic cells. *Oncogene*. 1999;18:3411–3418.
6. Eymin B, Sorret O, Droin N, et al. Caspase-induced proteolysis of the cyclin-dependent kinase inhibitor p27kip1 mediates its anti-apoptotic activity. *Onco gene*. 1999;18:4839–4847.
7. Katayose Y, Kim M, Rakkar AN, Li Z, Cowan KH, Seth P. Promoting apoptosis: a novel activity associated with the cyclin-dependent kinase inhibitor p27kip1. *Cancer Res*. 1997;57:5441–5445.
8. Harper BL, Wang A, Seth P, Jette AM. Up-regulation of p27kip1, p21WAF1/Cip1 and p53Mdm2 is associated with, but not sufficient for, induction of squamous differentiation. *J Cell Sci*. 1998;111:1185–1196.
9. Nguyen L, Besson A, Heng JH, et al. p27kip1 independently promotes neuronal differentiation and migration in the cerebral cortex. *Bull Mem Acad R Med Belg*. 2007;162:110–114.
10. Denicourt C, Saenz CC, Datnow B, Cui XS, Dowdy SF. Relocalized p27kip1 tumor suppressor functions as a cytoplasmic metastatic oncogene in melanoma. *Cancer Res*. 2007;67:9238–9243.
11. Besson A, Dowdy SF, Roberts JM. CDK inhibitors: cell cycle regulators and beyond. *Dev Cell*. 2008;14:159–169.
12. Fero ML, Randel E, Gurley KE, Roberts JM, Kemp CJ. The murine gene p27kip1 is haploinsufficient for tumour suppression. *Nature*. 1998;396:177–180.
13. Philipp-Staheli J, Payne SR, Kemp CJ. p27(Kip1): regulation and function of a novel inhibitor of cell cycle progression. *Biochim Biophys Acta*. 1997;1345:128–136.
14. Yang RM, Naitoh J, Murphy M, et al. Low p27 expression predicts poor disease-free survival of patients with prostate cancer. *J Urol*. 1998;159:941–945.
15. Porter PL, Malone KE, Heagerty PJ, et al. Expression of cell-cycle regulators p27kip1 and cyclin E, alone and in combination, correlate with survival in young breast cancer patients. *Nat Med*. 1997;3:222–225.
16. Frederdsen S, Burns J, Milne AM, et al. High level expression of p27kip1 and cyclin D1 in some human breast cancer cells: inverse correlation between the expression of p27(kip1) and degree of malignancy in human breast and colorectal cancers. *Proc Natl Acad Sci U S A*. 1997;94:6380–6385.
17. Palmqvist R, Stenling R, Öberg Landberg ÅG. Prognostic significance of p27kip1 expression in colorectal cancer: a clinicopathological characterization. *J Pathol*. 1999;188:18–23.
18. Sui L, Dong Y, Ohno M, et al. Jab1 expression is associated with invasive growth of epithelial ovarian tumors. *Clin Cancer Res*. 2001;7:4130–4135.
19. Lee J, Kim SS. The function of p27kip1 during tumor development. *Exp Mol Med*. 2009;41:765–771.
20. Reymond N, d’Agua BB, Ridley AJ. Crossing the endothelial barrier during metastasis. *Nat Rev Cancer*. 2013;13:858–870.
21. Park SL, Lee EJ, Kim JW, Moon SK. p27kip1 is involved in ERK1/2-mediated MMP-9 expression via the activation of NF-κB binding in the IL-8-induced migration and invasion of 5637 cells. *Int J Oncol*. 2014;44:1349–1356.
23. Koomoa DL, Geerts D, Lange I, et al. DFMO/eflornithine inhibits migration and invasion downstream of MYCN and involves p27Kip1 activity in neuroblastoma. *Int J Oncol*. 2013;42:1219–1228.

24. Bouda J, Hes O, Koprivova M, Pesek M, Svoboda T, Boudova L. P27 as a prognostic factor of early cervical carcinoma. *Int J Gynecol Cancer*. 2013;23:164–169.

25. Gao H, Ouyang X, Banach-Petrosky W, et al. A critical role for p27kip1 gene dosage in a mouse model of prostate carcinogenesis. *Proc Natl Acad Sci U S A*. 2004;101:17204–17209.

26. Khatri P, Sirot A, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol*. 2012;8:e1002375.

27. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102:15545–15550.

28. Mooney MA, Nigg JT, McWeeney SK, Wilmot B. Functional and genomic context in pathway analysis of GWAS data. *Trends Genet*. 2014;30:210–220.

29. Zhang Y, Liu ZL, Song M. ChiNet uncovers gene rewired transcription subnetworks in tolerant yeast for advanced biofuels conversion. *Nucleic Acids Res*. 2015;43:4393–4407.

30. Corton TB, Nguyen HH, Said JI, Ouyang Z, Zhang J, Song M. Discerning mechanistically rewired biological pathways by cumulative interaction heterogeneity statistics. *Sci Rep*. 2015;5:9634.

31. Vandenberg LN, Colborn T, Hayes TB, et al. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. *Endocr Rev*. 2012;33:378–455.

32. Zhang Y, Song M. Deciphering interactions in causal networks without parametric assumptions. arXiv Molecular Networks. 2013;arXiv:1311.2707.

33. Hill SM, Heiser LM, Cokelaer T, et al. Inferring causal molecular networks: empirical assessment through a community based effort. *Nat Methods*. 2016;13:310–318.

34. Kohn M, Melnick R. Biochemical origins of the non-monotonic receptor-mediated dose-response. *J Mol Endocrinol*. 2002;29:113–123.

35. Kruck S, Merseburger AS, Hennenlotter J, et al. High cytoplasmic expression of p27(Kip1) is associated with a worse cancer-specific survival in clear cell renal cell carcinoma. *BJU Int*. 2012;109:1565–1570.

36. Hulit J, Lee RJ, Li Z, et al. p27kip1 repression of ErbB2-induced mammary tumor growth in transgenic mice involves Skp2 and Wnt/beta-catenin signaling. *Cancer Res*. 2006;66:8529–8541.

37. Bolstad BM, Irizarry RA, Åstrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformat*. 2003;19:185–193.

38. Song M, Zhang Y, Katarzoff AJ, Edgar BA, Buttritta L. Hunting complex differential gene interaction patterns across molecular contexts. *Nucleic Acids Res*. 2014;42:e57.

39. Wang H, Song M. Clkmeans.id.dp: Optimal k-means clustering in one dimension by dynamic programming. *R J*. 2011;3:29–33.

40. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A*. 2003;100:9440–9445.

41. Kemp CJ. Multistep skin cancer in mice as a model to study the evolution of cancer cells. *Semin Cancer Biol*. 2005;15:460–473.

42. Mascaux C, Iannino N, Martin B, et al. The role of RAS oncogene in survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br J Cancer*. 2005;92:131–139.

43. Alle KM, Henshall SM, Field AS, Sutherland RL. Cyclin D1 protein is overexpressed in hyperplasia and intraductal carcinoma of the breast. *Clin Cancer Res*. 1998;4:847–854.

44. Sutherland RL, Muñoz-Florez EA. Cyclin D1 and mammary carcinoma: new insights from transgenic mouse models. *Breast Cancer Res*. 2002;4:14–17.

45. Forbes SA, et al. COSMIC: exploring the world’s knowledge of somatic mutations in human cancer. *Nucleic Acids Res*. 2015;43:D805–811.

46. Kohn M, Melnick R. Biochemical origins of the non-monotonic receptor-mediated dose-response. *J Mol Endocrinol*. 2002;29:113–123.

47. Kruck S, Merseburger AS, Hennenlotter J, et al. High cytoplasmic expression of p27(Kip1) is associated with a worse cancer-specific survival in clear cell renal cell carcinoma. *BJU Int*. 2012;109:1565–1570.

48. Hulit J, Lee RJ, Li Z, et al. p27kip1 repression of ErbB2-induced mammary tumor growth in transgenic mice involves Skp2 and Wnt/beta-catenin signaling. *Cancer Res*. 2006;66:8529–8541.

49. Besson A, Goujon-West M, Schmidt A, Hall A, Roberts JM. p27kip1 modulates cell migration through the regulation of RhoA activation. *Genes Dev*. 2004;18:862–876.

50. Li Z, Jiao X, Wang C, et al. Cyclin D1 induction of cellular migration requires p27kip1 and p21cip1 during chemically induced skin carcinogenesis. *Oncogene*. 1999;18:4689–4698.