Sensitivity analysis of Wnt $\beta$-catenin based transcription complex might bolster power-logarithmic psychophysical law and reveal preserved gene gene interactions

Shriprakash Sinha$^{*, \dagger}$

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

Insight, Innovation and Integration
Let a sensation magnitude $\gamma$ be determined by a stimulus magnitude $\beta$. The Weber's laws states that $\Delta \gamma$ remains constant when the relative stimulus increment $\Delta \gamma / \beta$ remains constant. It has been found that this law is actually a derivation of Bernoulli's law were $\Delta \gamma \propto \log (\Delta \beta / \beta)$. Recently, such psychophysical laws have been observed in the behaviour of certain intra/extracellular factors working in the Wnt pathway and this manuscript tests the veracity of the prevalence of such laws, albeit at a coarse level, using sensitivity analysis on biologically inspired computational causal models. Sensitivity analysis observations reveal such natural behaviours in the signaling pathway at systems biology level.

Abstract
Sensitivity analysis plays a crucial role in observing the behaviour of output of a variable given variations in the input. Recently, it has been found that some factors of the Wnt signaling pathway show a natural behaviour that can often be characterized by psychophysical laws. In colorectal cancer, $\beta$-catenin based transcription complex (TRCMPLX) plays a major role in driving the Wnt signaling pathway. In this manuscript, the deviation in the predictive behaviour of TRCMPLX conditional on the evidences of regulated gene expressions in normal and tumor samples is observed by varying the initially assigned values of conditional probability tables (cpt) for TRCMPLX. Preliminary analysis shows that the deviation in predictive behaviour of TRCMPLX conditional on gene evidences follows power-logarithmic psychophysical law crudely, implying deviations in output are proportional to increasing function of deviations in input and showing constancy for higher values of input. As a second observation, this points towards stability in the behaviour of TRCMPLX and is reflected in the preserved gene gene interactions of the Wnt pathway inferred from conditional probabilities of individual gene activation given the status of another gene activation derived using biologically inspired Bayesian Network. Finally, as a third observation, based on the sensitivity analysis it was found that the psychophysical laws are prevalent among the gene-gene interaction network also.

1 Introduction

1.1 Problem statement

In Sinha$^1$, it has been hypothesized that the activation state of TRCMPLX in the Wnt signaling pathway is not always the same as the state of the test sample (normal/tumorous) under consideration. For this, Sinha$^1$ shows various results on the predicted state of TRCMPLX conditional on the given gene evidences, while varying the assigned probability values of conditional probability tables of TRCMPLX during initialization of the Bayesian Network (BN). It was found that the predicted values often increase with an increasing value in the effect of the TRCMPLX on the genes. In a recent development, Goentoro and Kirschner$^2$ point to two findings namely, • the robust fold changes of $\beta$-catenin and • the transcriptional machinery of the Wnt pathway depends on the fold changes in $\beta$-catenin instead of absolute levels of the same and some gene transcription networks must respond to fold changes in signals according to the Weber’s law in sensory physiology.

In accordance with the aforementioned phenomena noted in Sinha$^1$, it would be important to test the veracity of the observed logarithmic laws and their derivations (like the Weber’s law) employed in Goentoro and Kirschner$^2$. In the current manuscript, preliminary analysis of results in Sinha$^1$ shows that the variation in predictive behaviour of TRCMPLX conditional on gene evidences follows power and logarithmic psychophysical law crudely, implying deviations in output are proportional to increasing function of deviations in input and showing constancy for higher values of input. This relates to

$^\dagger$ This manuscript is an extension of Sinha$^1$ which was a part of the Netherlands Bioinformatics Centre (NBIC) BioRange-II project BR2.5 (code - IGD7/G).

$^a$ Independent Researcher, 104-Madhurisha Heights, Risali Sector, Bhilai-490006, India. E-mail: shriprakash.sinha@gmail.com

$^*\dagger$ This manuscript is an extension of Sinha$^1$ which was a part of the Netherlands Bioinformatics Centre (NBIC) BioRange-II project BR2.5 (code - IGD7/G).
the work of Adler et al.\textsuperscript{3} on power and logarithmic law albeit at a coarse level. Note that Goentoro and Kirschner\textsuperscript{2} shows results for the behaviour of fold change of $\beta$-catenin with respect to single parameter values. The current work, takes into account the behaviour of TRCMPLX conditional on affects of multiple parameters in the form of evidences of various intra/extracellular gene expression values working in the pathway. Bayesian networks help in integrating such multiple parameter affects via various causal arcs and sensitivity analysis aids in the study of the such natural behaviour.

As a second observation, the forgoing result points towards stability in the behaviour of TRCMPLX and is reflected in the preserved gene gene interactions of the Wnt pathway inferred from conditional probabilities of individual gene activation given the status of another gene activation derived using biologically inspired Bayesian Network. Note that Weber’s law has been found to be a special case of Bernoulli’s logarithmic law Masin et al.\textsuperscript{4}. Finally, as a third observation, it would be interesting to note if these behaviours characterized by psychophysical laws are prevalent among the dual gene-gene interactions or not. If the results are affirmative then the following important speculations might hold true • Not just one factor (as pointed out by Goentoro and Kirschner\textsuperscript{2}) but the entire network might be involved in such a behavior at some stage or the other. • The psychophysical law might not only be restricted to the intra/extracellular components but also to the interactions among the the intra/extracellular components in the pathway.

It is important to be aware of the fact that the presented results are derived from a static Bayesian network model. It is speculated that dynamic models might give much better and more realistic results.

1.2 The logarithmic psychophysical law

Masin et al.\textsuperscript{4} states the Weber’s law as follows -

Consider a sensation magnitude $\gamma$ determined by a stimulus magnitude $\beta$. Fechner\textsuperscript{5} (vol 2, p. 9) used the symbol $\Delta \gamma$ to denote a just noticeable sensation increment, from $\gamma$ to $\gamma + \Delta \gamma$, and the symbol $\Delta \beta$ to denote the corresponding stimulus increment, from $\beta$ to $\beta + \Delta \beta$. Fechner\textsuperscript{5} (vol 1, p. 65) attributed to the German physiologist Ernst Heinrich Weber the empirical finding Weber\textsuperscript{6} that $\Delta \gamma$ remains constant when the relative stimulus increment $\Delta \beta / \beta$ remains constant, and named this finding Weber’s law. Fechner\textsuperscript{5} (vol 2, p. 10) underlined that Weber’s law was empirical.

It has been found that Bernoulli’s principle (Bernoulli\textsuperscript{7}) is different from Weber’s law (Weber\textsuperscript{6}) in that it refers to $\Delta \gamma$ as any possible increment in $\gamma$, while the Weber’s law refers only to just noticeable increment in $\gamma$. Masin et al.\textsuperscript{4} shows that Weber’s law is a special case of Bernoulli’s principle and can be derived as follows - Equation 1 depicts the Bernoulli’s principle and increment in sensation represented by $\Delta \gamma$ is proportional to change in stimulus represented by $\Delta \beta$.

$$\gamma = b \times \log \frac{\beta}{\alpha}$$

were $b$ is a constant and $\alpha$ is a threshold. To evaluate the increment, the following equation 2 and the ensuing simplification gives -

$$\Delta \gamma = b \times \log \frac{\beta + \Delta \beta}{\alpha} - b \times \log \frac{\beta}{\alpha}$$

$$\Delta \gamma = b \times \log \left(1 + \frac{\Delta \beta}{\beta}\right)$$

$$\Delta \gamma = b \times \log \left(1 + \frac{\Delta \beta}{\beta}\right)$$

Since $b$ is a constant, equation 2 reduces to

$$\Delta \gamma \propto \frac{\Delta \beta}{\beta}$$

were $\propto$ means ”is constant when there is constancy of” from Masin et al.\textsuperscript{4}. The final equation 3 is a formulation of Weber’s laws in wordings and thus Bernoulli’s principles imply Weber’s law as a special case. Using Fechner\textsuperscript{5} derivation, it is possible to show the relation between Bernoulli’s principles and Weber’s law. Starting from the last line of equation 2, the following yields the relation.

$$\Delta \gamma = b \times \log \frac{\beta + \Delta \beta}{\beta}$$

$$e^{\Delta \gamma} = e^{b \times \log \frac{\beta + \Delta \beta}{\beta}}$$

$$q_p = \left(1 + \frac{\Delta \beta}{\beta}\right)^{p}; \text{ since } e^{\log(x)} = x$$

$$\sqrt[1/\beta]{q_p} = 1 + \frac{\Delta \beta}{\beta}$$

$$\left(\frac{\Delta \beta}{\beta}\right)^{q} = \left(\frac{\Delta \beta}{\beta}\right)^{p}; \text{ were } \sqrt[1/\beta]{q_p} = k_q$$

$$k_r = \frac{\Delta \beta}{\beta}; \text{ the weber’s law s.t. } k_r = \sqrt[1/\beta]{q_p} - 1$$

Equation 3 holds true given the last line of equation 4. In the current study, observation of deviation recorded in predicted values of state of TRCMPLX conditional on gene evidences show crude logarithmic behaviour which might bolster Weber’s law and Bernoulli’s principles. But it must be noted that
these observations are made on static causal models and observation of the same behaviour in dynamical setting would add more value.

Before delving into the details of the experimental setup the following two subsections from Sinha help build the background on Wnt pathway and the computational model used to infer the results in this paper.

1.3 Canonical Wnt signaling pathway

The canonical Wnt signaling pathway is a transduction mechanism that contributes to embryo development and controls homeostatic self renewal in several tissues (Clevers'9). Somatic mutations in the pathway are known to be associated with cancer in different parts of the human body. Prominent among them is the colorectal cancer case (Gregorierff and Clevers10). In a succinct overview, the Wnt signaling pathway works when the Wnt ligand gets attached to the Frizzled(fzd)/LRP coreceptor complex. Fzd may interact with the Dishevelled(Dev) causing phosphorylation. It is also thought that Wnts cause phosphorylation of the LRP via casein kinase 1 (CK1) and kinase GSK3. These developments further lead to attraction of Axin which causes inhibition of the formation of the degradation complex. The degradation complex constitutes of Axin, the β-catenin transportation complex APC, CK1 and GSK3. When the pathway is active the dissolution of the degradation complex leads to stabilization in the concentration of β-catenin in the cytoplasm. As β-catenin enters into the nucleus it displaces the Groucho and binds with transcription cell factor TCF thus instigating transcription of Wnt target genes. Groucho acts as lock on TCF and prevents the transcription of target genes which may induce cancer. In cases when the Wnt ligands are not captured by the coreceptor at the cell membrane, Axin helps in formation of the degradation complex. The degradation complex phosphorylates β-catenin which is then recognized by FboxWD repeat protein β-TrCP. β−TrCP is a component of ubiquitin ligase complex that helps in ubiquitination of β-catenin thus marking it for degradation via the proteasome. Cartoons depicting the phenomena of Wnt being inactive and active are shown in figures 1(A) and 1(B), respectively.

1.4 Epigenetic factors

One of the widely studied epigenetic factors is methylation (Costello and Plass11, Das and Singh12, Issa13). Its occurrence leads to decrease in the gene expression which affects the working of Wnt signaling pathways. Such characteristic trends of gene silencing like that of secreted frizzled-related proteins (SFRP) family in nearly all human colorectal tumor samples have been found at extracellular level (Suzuki et al.14). Similarly, methylation of genes in the Dickkopf (DKKx Niehrs15, Sato et al.16), Dapper antagonist of catenin (DACTx Jiang et al.17) and Wnt inhibitory factor-1 (WIF1 Taniguchi et al.18) family are known to have significant effect on the Wnt pathway. Also, histone modifications (a class of proteins that help in the formation of chromatin which packs the DNA in a special form Strahl and Allis19) can affect gene expression (Peterson et al.20). In the context of the Wnt signaling pathway it has been found that DACT gene family show a peculiar behavior in colorectal cancer (Jiang et al.17). DACT1 and DACT2 showed repression in tumor samples due to increased methylation while DACT3 did not show obvious changes to the interventions. It is indicated that DACT3 promoter is simultaneously modified by the both repressive and activating (bivalent) histone modifications (Jiang et al.17).

Information regarding prior biological knowledge in terms of known influence relations and epigenetic factors have been depicted in the figure represented by figure 2 from Sinha.

2 Materials and methods

The models purported by Sinha involving the biological knowledge as well as epigenetic information depicted by $M_{PBK}\pm E1$ and biological knowledge excluding epigenetic information $M_{PBK}$ were used to predict the state of TRCMLPX given the gene evidences. Figure 2 depicts the model $M_{PBK}\pm E1$. The predictions were recorded over the varying effect of TRCMLPX on gene regulations via assignment of different values to conditional probability tables (cpt) of TRCMLPX while initializing the aforementioned BN models. This varying effect is represented by the term ETGN in Sinha.
As a recapitulation, the design of the experiment is a simple 2-holdout experiment where one sample from the normal and one sample from the tumorous are paired to form a test dataset. Excluding the pair formed in an iteration of 2-hold out experiment the remaining samples are considered for training of a BN model. Thus in a data set of 24 normal and 24 tumorous cases obtained from Jiang et al.\textsuperscript{17}, a training set will contain 46 samples and a test set will contain 2 samples (one of normal and one of tumor). This procedure is repeated for every normal sample which is combined with each of the tumorous sample to form a series of test datasets. In total there will be 576 pairs of test data and 576 instances of training data. Note that for each test sample in a pair, the expression value for a gene is discretized using a threshold computed for that particular gene from the training set. Computation of the threshold has been elucidated in Sinha\textsuperscript{1}. This computation is repeated for all genes per test sample. Based on the available evidence from the state of expression of all genes, which constitute the test data, inference regarding the state of both TRCMPLX and the test sample is made. These inferences reveal information regarding the activation state of the TRCMPLX and the state of the test sample. Finally, for each gene $g_i$, the conditional probability $\Pr(g_i = \text{active}|g_k \text{ evidence}) \forall k$ genes. Note that these probabilities are recorded for both normal and tumor test samples.

Three observations are presented in this manuscript. The first observation is regarding the logarithmic deviations in prediction of activation status of TRCMPLX conditional on gene expression evidences. The second observation is preservation of gene gene interactions across deviations. To observe these preservations, first the gene gene interactions have to be constructed from the predicted conditional probabilities of one gene given the evidence of another gene (for all gene evidences taken separately). After the construction, further pre-processing is required before the gene-gene interaction network can be inferred. Finally, the third observation is to check whether these laws are prevalent among the gene-gene interactions in the network.

3 Results and discussion

3.1 Logarithmic-power deviations in predictions of $\beta$-catenin transcription complex

Let $\gamma$ be $\Pr(\text{TRCMPLX} = \text{active}—\text{all gene evidences}), \beta$ be the assigned cpt value of TRCMPLX during initialization of the Bayesian Network models and $\Delta \beta$ be the deviation in the assigned values of TRCMPLX during initialization. To compute $\Delta \gamma$, the 576 predictions of $\gamma$ observed at $\beta = 90\%$ is subtracted from the 576 predictions of $\gamma$ observed at $\beta = 80\%$ and a mean of the deviations recorded. This mean becomes $\Delta \gamma$. The procedure is computed again for different value of $\beta$. In this manuscript, the effect of constant and incremental deviations are observed. Tables 1 and 2 represent the deviations for models $\mathcal{M}_{PBK+EI}$ and $\mathcal{M}_{PBK}$, respectively.

Figures 3, 4, 5 and 6 show the deviations represented in
Non-uniform scaled deviations

Deviation study for model $M_{PBK+EI}$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\Delta \beta$</th>
<th>$\frac{\Delta \beta \beta}{\beta}$</th>
<th>$\log(1 + \frac{\Delta \beta \beta}{\beta})$</th>
<th>Pr in Normal</th>
<th>Pr in Tumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>0.1</td>
<td>0.125</td>
<td>0.117783</td>
<td>0.1400355</td>
<td>0.1097089</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1</td>
<td>0.1428571</td>
<td>0.1335314</td>
<td>0.13423754</td>
<td>0.10906427</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1</td>
<td>0.1666667</td>
<td>0.1541507</td>
<td>0.04384244</td>
<td>0.08052346</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1823216</td>
<td>0.0005872203</td>
<td>0.07294716</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1</td>
<td>0.125</td>
<td>0.117783</td>
<td>0.03055427</td>
<td>0.09151151</td>
</tr>
<tr>
<td>0.7</td>
<td>0.2</td>
<td>0.2857143</td>
<td>0.2513144</td>
<td>0.04479181</td>
<td>0.1823758</td>
</tr>
<tr>
<td>0.6</td>
<td>0.3</td>
<td>0.5</td>
<td>0.4054651</td>
<td>0.04917605</td>
<td>0.2628992</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.8</td>
<td>0.5877867</td>
<td>0.04976327</td>
<td>0.33584646</td>
</tr>
</tbody>
</table>

Table 1 Deviation study for model $M_{PBK+EI}$. Pr = mean value of $Pr(\text{TRCMPLX} = \text{active} | g_{e_i} \text{ evidences})$ over all runs.

Deviation study for model $M_{PBK}$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\Delta \beta$</th>
<th>$\frac{\Delta \beta \beta}{\beta}$</th>
<th>$\log(1 + \frac{\Delta \beta \beta}{\beta})$</th>
<th>Pr in Normal</th>
<th>Pr in Tumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>0.1</td>
<td>0.125</td>
<td>0.117783</td>
<td>0.1400355</td>
<td>0.1097089</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1</td>
<td>0.1428571</td>
<td>0.1335314</td>
<td>0.13423754</td>
<td>0.10906427</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1</td>
<td>0.1666667</td>
<td>0.1541507</td>
<td>0.04384244</td>
<td>0.08052346</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1823216</td>
<td>0.0005872203</td>
<td>0.07294716</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1</td>
<td>0.125</td>
<td>0.117783</td>
<td>0.03055427</td>
<td>0.09151151</td>
</tr>
<tr>
<td>0.7</td>
<td>0.2</td>
<td>0.2857143</td>
<td>0.2513144</td>
<td>0.04479181</td>
<td>0.1823758</td>
</tr>
<tr>
<td>0.6</td>
<td>0.3</td>
<td>0.5</td>
<td>0.4054651</td>
<td>0.04917605</td>
<td>0.2628992</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.8</td>
<td>0.5877867</td>
<td>0.04976327</td>
<td>0.33584646</td>
</tr>
</tbody>
</table>

Table 2 Deviation study for model $M_{PBK}$. Pr = mean value of $Pr(\text{TRCMPLX} = \text{active} | g_{e_i} \text{ evidences})$ over all runs.

Fig. 3 Constant deviations in $\beta$ i.e $ETGN$ and corresponding deviations in $Pr(\text{TRCMPLX} = \text{active} | g_{e_i} \text{ evidences})$ for both normal and tumor test samples. Corresponding Weber and Bernoulli deviations are also recorded. Note that the plots and the y-axis depict scaled deviations to visually analyse the observations. The model used is $M_{PBK+EI}$. Red - constant deviation in Normal, constant deviation in Tumor, Green - constant deviation in Weber’s law, Cyan - constant deviation in Bernoulli’s law.

For model $M_{PBK}$, the above observations do not yield consistent behaviour. In figure 5, for the case of constant deviations, only the deviations in activation of $TRCMPLX$ conditional on gene evidences for normal test samples exponential in nature and were found to be directly proportional to the negative of both the Weber’s and Bernoulli’s law. But the deviations in activation of $TRCMPLX$ conditional on gene evidences in tumor test samples show noisy behaviour. But this observation is not the case in incremental deviations for the same model. For the case of incremental deviations as represented in figure 6, the deviations in activation of $TRCMPLX$ conditional on gene evidences is directly proportional to both the Weber’s and Bernoulli’s law. The figure actually represents the plots with inverted values i.e negative values. A primary reason for this behaviour might be that $M_{PBK}$ does not capture and constrain the network as much as $M_{PBK+EI}$ which include epigenetic information. This inclusion of heterogeneous information adds more value to the biologically inspired network and reveals the hidden natural laws occurring in the
signaling pathway in both normal and tumor cases.

Finally, these observations present a crude yet important picture regarding the downstream transcriptional behaviour of signaling pathway in case of colorectal cancer. Psychophysical laws might not be constrained to a particular domain and as can be seen here, they might play an important role in shedding light on behaviour of the pathway. In context of Goenoro and Kirschner, the presented results might be crude in terms of static observations, yet they show corresponding behaviour of transcriptional activity in terms of psychophysical laws. Further investigations using dynamic models might reveal more information in comparison to the static models used in Sinha. The observations presented here might bolster the existence of behavioural phenomena in terms of logarithmic laws and its special cases.

3.2 Preservation of gene-gene interactions

The second part of this study was to find interactions between two genes by observing the conditional probability of activation status of one gene given the evidence of another gene. Let $g$ be a gene. To obtain the results, two steps need to be executed in a serial manner. The first step is to construct gene interactions based on the available conditional probabilities denoted by $\Pr(g_k = \text{active/repressed}|g_k \text{ evidence}) \forall k$ genes. The second step is to infer gene interaction network based purely on reversible interactions. Note that networks are inferred for gene evidences using normal and tumor test samples separately. The following sections elucidate the steps before explaining the implications.

3.2.1 Constructing gene-gene interactions

Before starting the construction of interactions from the conditional probabilities, assign a variable $g_{ij}$ as an empty list (say in R language). Then $\forall i$ genes, execute the following:

1. $\forall 576$ runs iterated by a counter $j$

(a) append $x_N$ with the vector whose elements are $\Pr(g_k = \text{active}|g_k \text{ evidence}) \forall k$ genes in the $j^{th}$ run for Normal test sample. This creates a matrix at the end of the runs.

(b) append $x_T$ with the vector whose elements are $\Pr(g_k = \text{active}|g_k \text{ evidence}) \forall k$ genes in the $j^{th}$ run for Tumor test sample. This creates a matrix at the end of the runs.

(c) append $ge_N$ with the vector whose elements are $g_{i_k}$ evidence $\forall k$ genes in the $j^{th}$ run for Normal test sample. This creates a matrix at the end of the runs.

(d) append $ge_T$ with the vector whose elements are $g_{i_k}$ evidence $\forall k$ genes in the $j^{th}$ run for Tumor test sample. This creates a matrix at the end of the runs.

- Fig. 4 Incremental deviations in $\beta$ i.e ETGN and corresponding deviations in $\Pr(TRCPMLX = \text{active}|g_{i_k} \text{ evidences})$ for both normal and tumor test samples. Corresponding Weber and Bernoulli deviations are also recorded. Note that the plots and the y-axis depict scaled deviations to visually analyse the observations. The model used is $M_{PBK+EII}$. Red - incremental deviation in Normal, incremental deviation in Tumor, Green - incremental deviation in Weber’s law, Cyan - incremental deviation in Bernoulli’s law.

- Fig. 5 Constant deviations in $\beta$ i.e ETGN and corresponding deviations in $\Pr(TRCPMLX = \text{active}|g_{i_k} \text{ evidences})$ for both normal and tumor test samples. Corresponding Weber and Bernoulli deviations are also recorded. Note that the plots and the y-axis depict scaled deviations to visually analyse the observations. The model used is $M_{PBK}$. Red - constant deviation in Normal, constant deviation in Tumor, Green - constant deviation in Weber’s law, Cyan - constant deviation in Bernoulli’s law.
Table 3 SFRP5 activation status in test samples conditional on status of individual gene activation (represented by evidence in test data) in Normal and Tumor samples. Measurements are taken over summation of all predicted values across the different runs of the 2-Hold out experiment. Here the notations denote the following: a - active, p - passive, N - Normal, T - Tumor, $ggIN$ - gene-gene interaction with Normal, $ggIT$ - gene-gene interaction with Tumor, $<>$ - active and | - repressed.


3. compute mean across columns of $xN$ and $xT$ to obtain averaged $Pr_N(g_i|g_k)$ and $Pr_T(g_i|g_k) \forall k$ gene evidences and $\forall i$ genes. Note $k, i \in 1,...,n$ if $n$ is the total number of genes.

4. assign a vector of $\hat{Pr}_N(g_i|g_k)$ $\forall k$ genes to $PggN$ and a vector of $\hat{Pr}_T(g_i|g_k)$ $\forall k$ genes to $PggT$

5. $\forall k$ genes except the $i^{th}$ one

(a) if($k \neq i$)

i. assign variables $tmpaa_N, tmpar_N, tmpaa_T, tmpar_T$ and $tmprr_T$ to 0.

ii. $\forall$ 576 runs iterated by a counter $l$

A. if($ge_N[l,k] == 1$ and $x_N[l,k] < 0.5$) increment $tmprr_N$ by 1

B. else if($ge_N[l,k] == 1$ and $x_N[l,k] >= 0.5$) increment $tmprr_N$ by 1

C. else if($ge_N[l,k] == 2$ and $x_N[l,k] < 0.5$) increment $tmpaa_N$ by 1

D. else if($ge_N[l,k] == 2$ and $x_N[l,k] >= 0.5$) increment $tmpaa_N$ by 1
Fig. 7 Gene gene interactions for normal case while using $M_{PBK+EI}$. Note that the effect of initialized cpt for $TRCMPLX$ is $90\%$ in tumorous case and $10\%$ in normal case. Diamond $<>$ means activation and straight bar $|$ means repression.

E. if($g_T[l,k] == 1$ and $x_T[l,k] < 0.5$) increment $tmprr_T$ by 1
F. else if($g_T[l,k] == 1$ and $x_T[l,k] >= 0.5$) increment $tmpar_T$ by 1
G. else if($g_T[l,k] == 2$ and $x_T[l,k] < 0.5$) increment $tmpra_T$ by 1
H. else if($g_T[l,k] == 2$ and $x_T[l,k] >= 0.5$) increment $tmpaa_T$ by 1

iii. Comment - store results
iv. append $ge$ with $g_k$, $rr_N$ with $tmprr_N$, $ar_N$ with $tmpar_N$, $ra_N$ with $tmpra_N$, $aa_N$ with $tmpaa_N$, $rr_T$ with $tmprr_T$, $ar_T$ with $tmpra_T$ and $aa_T$ with $tmpaa_T$

(b) store the variables in the previous step to a data frame (say in R language) to a variable $stats$.
(c) Comment - 1 means aa, 2 means ar, 3 means ra, 4 means rr

(d) assign variables $ggIN$ and $ggIT$ as empty vector $[]$
(e) $\forall j$ gene except the $i^{th}$ one under consideration
i. find the index $idx_N$ in stats that corresponds to 1 or 2 or 3 or 4
ii. if($idx_N == 1$) append $ggIN$ with interaction string $stats§g_j$ $<>$ $<> g_i$
iii. else if($idx_N == 2$) append $ggIN$ with interaction string $stats§g_j$ $|-$ $<> g_i$
iv. else if($idx_N == 3$) append $ggIN$ with interaction string $stats§g_j$ $|-$ $| g_i$
v. else if($idx_N == 4$) append $ggIN$ with interaction string $stats§g_j$ $|-$ $| g_i$
vi. find the index $idx_T$ in stats that corresponds to 1 or 2 or 3 or 4
vii. if($idx_T == 1$) append $ggIT$ with interaction string $stats§g_j$ $<>$ $<> g_i$
viii. else if($idx_T == 2$) append $ggIT$ with interaction string $stats§g_j$ $|-$ $<> g_i$

8 | Journal Name, 2010, [vol],1–14
Based on the above execution, for each gene a matrix is obtained that shows the statistics of how the status of a gene is affected conditional on the individual evidences of the remaining genes. Also, for each of the $i^{th}$ gene the averaged $\Pr_N(g_i|g_j)$ is also stored in vector $PggN$. Same is done for tumors cases. These two vectors are later used to test the veracity of existence of psychophysical laws in gene-gene interaction network. Table 3 represents one such tabulation for gene SFRP5. For all runs and all test samples, the following was tabulated in table 3: $aa_N$ - SFRP5 is active (a) when a gene is active (a) in Normal (N) sample, $ar_N$ - SFRP5 is active (a) when a gene is repressed (r) in Normal (N) sample, $ra_N$ - SFRP5 is repressed (r) when a gene is active (a) in Normal (N) sample, $rr_N$ - SFRP5 is repressed (r) when a gene is repressed (r) in Normal (N) sample, $ar_T$ - SFRP5 is active (a) when a gene is active (a) in Tumor (T) sample, $art_T$ - SFRP5 is active (a) when a gene is repressed (r) in Tumor (T) sample, $pa_T$ - SFRP5 is repressed (r) when a gene is active (a) in Tumor (T) sample, $g91N$ - interaction of SFRP5 given the gene evidence based on majority voting among $aa_N$, $ar_N$, $ra_N$ and $rr_N$ and finally, $g91T$ - interaction of SFRP5 given the gene evidence based on majority voting among $aa_T$, $ar_T$, $ra_T$ and $rr_T$. The highest score among $aa_N$, $ar_N$, $ra_N$ and $rr_N$ confirms the relation between genes using Normal (Tumor) samples. Active (repressed) for SFRP5 is based on discretization the predicted conditional probability $Pr(SFRP5 = active|g_j$ evidence) as $\geq 0.5$ ($< 0.5$). Active (repressed) for a particular gene evidence $g_j$ is done using discrete evidence. In table 3, under the columns $g91N$ and $g91T$, $<>$ implies the gene is active and $|$ implies the gene is repressed or passive.

It was found that $DKK1$, $DKK3 - 2$, $DKK4$ expressed similar repression behaviour as the standard genes $LEF1$, $MYC$, $CCND1$ and $CD44$ in Normal test samples (and vice versa for Tumor test samples). Also, $DKK2$ and $DKK3 - 1$ show similar activated behaviour as $DACT - 1/2/3$ and $SFRP - 1/2/3$ in Normal test samples (and vice versa for Tumor test samples). In comparison to $DKK2$, $DKK3 - 1$, $DACT - 1/2/3$ and $SFRP - 1/2/3$, which are activated along with $SFRP5$ in Normal test samples (repressed in Tumor test samples), genes $DKK1$, $DKK3 - 2$, $DKK4$, $LEF1$, $MYC$, $CCND1$ and $CD44$ were reversed while $SFRP3$ is activated in Normal test sample (roles reversed in Tumor cases). Genes which showed similar behaviour to $SFRP5$ might be affected by epigenetic factors, i.e these factors might play a role in suppressing the gene expression in Normal test samples. The reverse might be the case for genes that were suppressed in Tumor test samples.

It can also be seen that most of the interactions are reversible except for $SFRP4|<> SFRP5$ in Normal test sample and $SFRP4 <> SFRP5$ in Tumor test sample. This kind of interaction is deleted the existing set of interactions as they do not provide concrete information regarding the functional roles of the genes in normal and tumor cases. This attributes to one of the following facts (1) noise that might corrupt prediction values as can be seen in the columns of $aa_N$ ($aart_T$, $ar_N$ ($ar_TT$), $ra_N$ ($ra_TT$) and $rr_N$ ($rr_TT$) or (2) other multiple genes might be interacting along with $SFRP5$ in a combined manner and it is not possible to decipher the relation between $SFRP5$ and other genes. This calls for investigation of prediction of $SFRP5$ status conditional on joint evidences of two or more genes (a combinatorial problem with a search space order of $2^{17} - 17$, which excludes 17 cases of individual gene evidences which have been considered here). Incorporating multiple gene evidences might not be a problem using Bayesian network models as they are designed to compute conditional probabilities given joint evidences also (except at the cost of high computational time).

### 3.2.2 Inferring gene-gene interaction network

Next, after the construction of gene-gene interactions, it is necessary to infer the network. The inference of the estimated gene-gene interactions network is based on explicitly reversible roles in Normal and Tumor test samples. This means that only those interactions are selected which show the following property: $g_j <> g_i$ in Normal if and only if $g_j| <> g_i$ in Tumor, $g_j <> g_i$ in Normal if and only if $g_j| <> g_i$ in Tumor, $g_j| <> g_i$ in Normal if and only if $g_j| <> g_i$ in Tumor and finally, $g_j| <> g_i$ in Normal if and only if $g_j| <> g_i$ in Tumor. This restricts the network to only reversible gene-gene interactions in Normal and Tumor cases. Note that an interaction $g_j|2Rg_i (g_j|2Rg_i)$ is depicted by $Pr(g_i|g_j) (Pr(g_j|g_i)).$

Lastly, duplicate interactions are removed from the network for Normal samples. This is repeated for the network based on tumor samples also. This removal is done by removing one of the interactions from the following pairs $(g_j| <> g_i$ and $g_i| <> g_j$), $(g_j| <> g_i$ and $g_i| <> g_j$), $(g_j| <> g_i$ and $g_i| <> g_j$) and $(g_j| <> g_i$ and $g_i| <> g_j$). Figure 7 shows one such network after complete interaction construction, inference and removal of duplicate interactions in using Normal test samples with ETGN of 90% in $M_{PBK+EI}$. For the case of Tumor test samples with ETGN 90% in $M_{PBK+EI}$, only the reversal of interactions need to
be done. Table 4 represents these interactions in tabulated form.

Finally, different networks were generated by varying the effect of TRCMPLX (ETGN) and compared for the normal test samples. Table 5 represents the different interactions that were preserved in network from ETGN 90% with respect to networks obtained from ETGN with values of 80%, 70%, 60% and 50%. It was found that most of the genetic interactions depicted in figure 7 were found to be preserved across the different variations in ETGN as shown in table 5. Out of the total n genes which construct a fully connected graph of n × (n−1) 2, it was observed that lesser number of interconnections were preserved. This preservation indicates towards the robustness of the genetic contributions in the Wnt signaling pathway in both normal and tumor test samples. Note that these observations are made from static models and dynamic models might reveal greater information.

3.3 Logarithmic-power deviations in prediction of gene-gene interactions

In the penultimate section on preservation of gene-gene interaction, it was found that some of the interactions remain preserved as there was change of the effect transcription complex. The transcription complex itself was found to follow a logarithmic-power psychophysical law. It would be interesting to observe if these laws are prevalent among the gene-gene interactions in the network.

In Sinha, the unknown behaviour of SFRP5 in the Wnt pathway has been revealed slightly using computational causal inference. In figure 7, SFRP5 shows preservation in the network and it’s interaction with other genetic factors involved in the model proposed in Sinha has been depicted. In one such paired interaction between SFRP5 and MYC, SFRP5
Missing gene-gene interactions for different values of ETGN

<table>
<thead>
<tr>
<th>90N-T1</th>
<th>80N-T1</th>
<th>70N-T1</th>
<th>60N-T1</th>
<th>50N-T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>90N-T1</td>
<td>(in 90N-T1) MYC</td>
<td>-</td>
<td>DACT1, CCND1</td>
<td>-</td>
</tr>
<tr>
<td>80N-T1</td>
<td>(in 90N-T1)</td>
<td>DACT3</td>
<td>&lt;</td>
<td>DACT1, MYC</td>
</tr>
<tr>
<td>70N-T1</td>
<td>(in 90N-T1) DACT3</td>
<td>&lt;</td>
<td>DACT1, MYC</td>
<td>-</td>
</tr>
<tr>
<td>60N-T1</td>
<td>(in 90N-T1) DACT3</td>
<td>&lt;</td>
<td>DACT1, MYC</td>
<td>-</td>
</tr>
<tr>
<td>50N-T1</td>
<td>(in 90N-T1) CD44</td>
<td>&lt;</td>
<td>&lt;</td>
<td>DKK3-1, SFRP1</td>
</tr>
</tbody>
</table>

Table 5 Tabulated missing gene-gene interactions of figure 7 using $M_{PBK+FI}$ obtained in case of Normal samples. Interactions found in Normal samples with 80%, 70%, 60% and 50% effect that are not found with 90% and vice versa have been recorded. Here, the symbols represent the following: - < - activation and | repression/suppression. Note that for Tumor cases, the interaction roles were found to be reversed, ie. < - in normal became | - in tumor, | - in normal became < - in tumor, < - < - in normal became | - in tumor and | - in normal became < - < - in tumor.

Deviation study for SFRP5 and MYC for tumor case

| $\Delta$ | $\frac{\Delta^2}{\sigma^2}$ | log$(1 + \frac{\Delta^2}{\sigma^2})$ | Pr$(SFRP5|M)$ | Pr$(MYC|SFRP5)$ |
|---------|----------------|------------------------|--------------|----------------|
| 0.8     | 0.1            | 0.125                  | 0.117785     | 0.001049836    | 0.000000e+00  |
| 0.7     | 0.1            | 0.1428571              | 0.1335314    | 0.00129522     | 0.000000e+00  |
| 0.6     | 0.1            | 0.1666667              | 0.1541507    | 0.001216178    | -5.551115e-17 |
| 0.5     | 0.1            | 0.2                    | 0.1823216    | 0.001310739    | 5.551115e-17  |
| 0.8     | 0.1            | 0.125                  | 0.117785     | 0.004706275    | 0.000000e+00  |
| 0.7     | 0.2            | 0.2857143              | 0.2513144    | 0.003656439    | 0.000000e+00  |
| 0.6     | 0.3            | 0.5                    | 0.4054651    | 0.002526918    | 0.000000e+00  |
| 0.5     | 0.4            | 0.8                    | 0.5877867    | 0.001310739    | 5.551115e-17  |

Table 6 Deviation study for Pr$(SFRP5|M)$ and Pr$(MYC|SFRP5)$ for tumor case

showed activation (repression) and MYC showed repression (activation) in normal (tumor) samples. As the change in the effect of transcription complex was induced via sensitizing the initially assigned cpt values, the deviations in the prediction of the gene-gene interaction network was observed to follow the logarithmic-power law crudely.

Table 6 and 7 show these deviations in the prediction of the interactions for both the normal and the tumor cases. The tables show how deviations are affected when the changes in the effect of the transcription complex are done at constant and incremental level. To summarize the results in these tables, graphs were plotted in figures 8 for Pr$(SFRP5|M)$ (constant deviations), 9 for Pr$(MYC|SFRP5)$ (constant deviations), 10 for Pr$(SFRP5|M)$ (incremental deviations) and 9 for Pr$(MYC|SFRP5)$ (incremental deviations).

Considering figure 8, when deviations are constant in both Weber and Bernoulli formulation, the deviations in the prediction of Pr$(SFRP5|M)$ is observed to be logarithmic in the normal samples (apropos to the Weber and Bernoulli deviations represented by green and cyan curves). Deviation in predictions are depicted by the red (blue) curves for normal (tumor) samples. Such a behaviour is not observed for Pr$(MYC|SFRP5)$ as is depicted in figure 9. Note that the interaction for SFRP5 given MYC was observed to be reversible in normal and tumor cases. But this is not so with the interaction for MYC given SFRP5. It might be expected that the non conformance of logarithmic-power law for Pr$(MYC|SFRP5)$ may be due to the non preservation of the interaction of MYC given SFRP5. This is so because Pr$(SFRP5|M)$ depicts a reversible $SFRP5 < -|MYC$ ($MYC < -|SFRP5$) in the network on normal (tumor) samples, while Pr$(MYC|SFRP5)$ does not depict a reversible $MYC < - |SFRP5$ ($MYC < -|SFRP5$) in the network on normal (tumor) samples.

Similar behaviour was observed in the case of incremental deviations as depicted in figures 10 and 11. Analysis of behaviour of other gene-gene interactions can be observed in a similar way.
Fig. 8 Constant deviations in $\beta$ i.e ETGN and corresponding deviations in $\text{Pr}(SFRP5|MYC)$ for both normal and tumor test samples. Corresponding Weber and Bernoulli deviations were also recorded. Note that the plots and the y-axis depict scaled deviations to visually analyse the observations. The model used is $M_{PBK+EI}$. Red - deviation in $\text{Pr}(SFRP5|MYC)$ in Normal case using Weber’s law, Blue - deviation in $\text{Pr}(SFRP5|MYC)$ in Tumor using Weber’s law, Green - constant deviation in Webers law, Cyan - constant deviation in Bernoullis law.

Fig. 9 Same as figure 9 but for $\text{Pr}(MYC|SFRP5)$. Instead of constant deviations, incremental deviations are represented.

Fig. 10 Same as figure 9 but for $\text{Pr}(SFRP5|MYC)$. Instead of constant deviations, incremental deviations are represented.

Fig. 11 Same as figure 9 but for $\text{Pr}(MYC|SFRP5)$. Instead of constant deviations, incremental deviations are represented.

4 Future directions

In context of the above observations, dynamic models might reveal greater information regarding the psychophysical laws. Work by Goentoro and Kirschner \cite{2} employs sensitivity analysis methods to reveal such laws by tuning single parameters.
There might be a few ways to measure fold change in single an multi parameter settings. Future work might involve deeper study of the phenomena based on multi-parameter setting in a dynamic bayesian network model. If one incorporates nodes in between two time snapshots of \( \beta \)-catenin concentration in a dynamic bayesian network, one might be able to measure the changes at different phases of the signaling pathway. For example, in figure 12 a set of nodes measuring the different concentrations of \( \beta \)-catenin (say \( N \)) are depicted. In a dynamic bayesian network, the previous concentration at \( t \) is connected to the next concentration at \( t + 1 \). Also, to measure the effect of difference (\( \Delta N \)), a change in concentration can be measured. Computations regarding fold change (\( \Delta N \)) could then be estimated as posterior probabilities given the two concentrations, which the Bayesian networks can easily handle. In case more parameters need to be involved (say the effect of Wnt and APC together), nodes might be added as shown below. Then the fold change is conditional on \( N(t+1), N(t+2), \Delta Wnt \) and \( \Delta APC \) and is estimated as 

\[
Pr(\Delta N(t+1) | N(t+1), N(t+2), \Delta Wnt, \Delta APC).
\]

Regarding sensitivity analysis, in nonlinear problems, it might be useful to use Sobol\(^2\) indices to estimate the sensitivity of the parameters. These indices are a way to estimate the changes in a multiparameter setting thus helping one to conduct global sensitivity analysis instead of local sensitivity analysis (Glen and Isaacs\(^2\)). Finally, with respect to the robustness of the gene-gene interaction network, the current work employs a very simple algorithm to construct the network and infer preserved interactions across the range of values set for a particular parameter. This helps in eliminating interactions that do not contribute enough biological information in the pathway or are non existant and require further analysis by integration of more data. Work in these lines would require incorporation of bigger datasets.

5 Conclusion

In this preliminary work via sensitivity analysis, the variation in predictive behaviour of \( \beta \)-catenin based transcription complex conditional on gene evidences follows logarithmic psychophysical law crudely, implying deviations in output are proportional to increasing function of deviations in input and show constancy for higher values of input. This points towards stability in the behaviour of transcriptional activity downstream of the Wnt pathway. As a further development, this stability might reflect the preserved gene gene interactions of the Wnt pathway inferred from conditional probabilities of individual gene activation given the status of another gene activation derived using biologically inspired Bayesian Network. Finally, based on the sensitivity analysis it was observed that the psychophysical laws are prevalent among the gene-gene interactions network also.

\[
\text{Conflict of interest}
\]

None

\[
\text{Acknowledgement}
\]

Thanks to - (1) the Royal Society of Chemistry (RSC) for giving permission to reproduce parts of material in Shriprakash Sinha, Integr. Biol., 2014, DOI: 10.1039/C4IB00124A. (2) all anonymous reviewers who have helped in refining this manuscript. Finally, the author is indebted to Mr. Prabhat Sinha and Mrs. Rita Sinha for financially supporting this project while the author was on educational and work leave.

\[
\text{References}
\]

5. G. T. Fechner, Elemente der Psychophysik (2 vols), Breitkopf and Hartel, 1860.