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

Shriprakash Sinha$^{*,a,\dagger}$

Abstract
Sensitivity analysis plays a crucial role in observing the behaviour of output of a variable given variations in the input. In colorectal cancer, $\beta$-catenin based transcription complex ($TRCMPLX$) plays a major role in driving the Wnt signaling pathway. In this manuscript, the variation in the effect of 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 variation in predictive behaviour of $TRCMPLX$ conditional on gene evidences follows 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 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.

1 Introduction
From Sinha$^1$, the following two sections help build the background before delving into the problem of sensitivity analysis.

1.1 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$^3$). 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 (Gregorieff and Clevers$^4$). 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 ($Dvl$) causing phosphorylation. It is also thought that Wnts cause phosphorylation of the LRP via cainase 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 $\beta$-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 $\beta$-catenin in the cytoplasm. As $\beta$-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 $\beta$-catenin which is then recognized by $Fbox/WD$ repeat protein $\beta-TrCP$. $\beta-TrCP$ is a component of ubiquitin ligase complex that helps in ubiquitination of $\beta$-catenin thus marking it for degradation via the proteasome. Cartoons depicting the phenomena of Wnt being inactive and inactivated.
active are shown in figures 1(A) and 1(B), respectively.

1.2 Epigenetic factors

One of the widely studied epigenetic factors is methylation (Costello and Plass\textsuperscript{5}, Das and Singal\textsuperscript{6}, Issa\textsuperscript{7}). 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.\textsuperscript{8}). Similarly, methylation of genes in the Dickkopf (DKK)\textsuperscript{x} Niehrs\textsuperscript{9}, Sato et al.\textsuperscript{10}), Dapper antagonist of catenin (DACT\textsuperscript{x} Jiang et al.\textsuperscript{11}) and Wnt inhibitory factor-1 (WIF1 Taniguchi et al.\textsuperscript{12}) 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 Allis\textsuperscript{13}) can affect gene expression (Peterson et al.\textsuperscript{14}). 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.\textsuperscript{11}). 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.\textsuperscript{11}).

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\textsuperscript{1}.

1.3 Problem statement

In Sinha\textsuperscript{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\textsuperscript{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\textsuperscript{16} point to two findings namely, • the robust fold changes of β-catenin and • the transcriptional machinery of the Wnt pathway depends on the fold changes in β-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\textsuperscript{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\textsuperscript{16}. In the current manuscript, preliminary analysis of results in Sinha\textsuperscript{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 the work of Adler et al.\textsuperscript{17} on power and logarithmic law albeit at a coarse level. Also, the 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{18}).

1.4 The logarithmic psychophysical law

Masin et al.\textsuperscript{18} states the Weber’s law as follows -

Consider a sensation magnitude $\gamma$ determined by a stimulus magnitude $\beta$. Fechner\textsuperscript{19} (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{19} (vol 1, p. 65) attributed to the German physiologist Ernst Heinrich Weber the empirical finding Weber\textsuperscript{20} that $\Delta \gamma$ remains constant when the relative stimulus increment $\frac{\Delta \beta}{\beta}$ remains constant, and named this finding Weber’s law. Fechner\textsuperscript{19} (vol 2, p. 10) underlined that Weber’s law was empirical.

It has been found that Bernoulli’s principle (Bernoulli\textsuperscript{21}) is different from Weber’s law (Weber\textsuperscript{20}) 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{18} shows that Weber’s law is a special case of Bernoulli’s principle and can be derived as follows - Equation 1 depicts the increment in sensation represented by $\Delta \gamma$ to be proportional to change in stimulus represented by $\Delta \beta$.

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

where $b$ is a constant and $\alpha$ is a threshold. To evaluate the increment, the following equation 2 and the ensuing simplification
Fig. 2 Influence diagram of $\mathcal{M}_{PBK + EI}$ contains partial prior biological knowledge and epigenetic information in the form of methylation and histone modification. Diagram drawn using Cytoscape\textsuperscript{15}. In this model the state of Sample is distinguished from state of TRCMPLX that constitutes the Wnt pathway.

gives -

\[
\Delta \gamma = b \times \log \left( 1 + \frac{\Delta \beta}{\beta} \right) - b \times \log \frac{\beta}{\alpha} = b \times \log \left( 1 + \frac{\Delta \beta}{\beta} \right) - b \times \log \frac{\beta}{\alpha} = b \times \log \left( 1 + \frac{\Delta \beta}{\beta} \right)
\]

(2)

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

\[
\Delta \gamma \circ \frac{\Delta \beta}{\beta}
\]

(3)

were $\circ$ means "is constant when there is constancy of" from Masin \textit{et al.}\textsuperscript{18}. 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{19} 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.

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

\[
e^{\Delta \gamma} = e^{b \times \log \left( 1 + \frac{\Delta \beta}{\beta} \right)}
\]

\[
k_p = e^{b \times \log \left( 1 + \frac{\Delta \beta}{\beta} \right)}; \text{ were } k_p = e^{\Delta \gamma}
\]

\[
k_p = \left( 1 + \frac{\Delta \beta}{\beta} \right)^b; \text{ since } e^{\log(x)} = x
\]

\[
\sqrt{k_p} = 1 + \frac{\Delta \beta}{\beta}
\]

\[
k_q \times \beta = \beta + \Delta \beta; \text{ were } \sqrt{k_p} = k_q
\]

\[
k_q \times \beta - \beta = \Delta \beta
\]

\[
(k_q - 1) \times \beta = \Delta \beta
\]

\[
k_q - 1 = \frac{\Delta \beta}{\beta}
\]

\[
k_r = \frac{\Delta \beta}{\beta}; \text{ the weber’s law s.t. } k_r = \sqrt[e]{e^{\Delta \gamma}} - 1
\]

(4)

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.

2 Materials and methods

The models purported by Sinha\(^1\) involving the biological knowledge as well as epigenetic information depicted by \(M_{PBK+E1}\) and biological knowledge excluding epigenetic information \(M_{PBK}\) were used to predict the state of TRCMPLX given the gene evidences. Figure 2 depicts the model \(M_{PBK+E1}\). The predictions were recorded over the varying effect of TRCMPLX on gene regulations via assignment of different values to conditional probability tables (cpt) of TRCMPLX while initializing the aforementioned BN models. This varying effect is represented by the term ETGN in Sinha\(^1\).

As a recapitulation, the design of the experiment is a simple 2-foldout experiment where one sample from the normal and one sample from the tumors 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.\(^11\), 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\(^1\). 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 the 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.

Two 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 and 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 preprocessing is required before the gene-gene interaction network can be inferred.

<table>
<thead>
<tr>
<th>Deviation study for model (M_{PBK+E1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta)</td>
</tr>
<tr>
<td>0.8</td>
</tr>
<tr>
<td>0.7</td>
</tr>
<tr>
<td>0.6</td>
</tr>
<tr>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 1 Deviation study for model \(M_{PBK+E1}\). \(Pr = \text{mean value of} Pr(\text{TRCMPLX} = \text{active}|\forall g_e\text{, evidences})\) over all runs.

<table>
<thead>
<tr>
<th>Deviation study for model (M_{PBK})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta)</td>
</tr>
<tr>
<td>0.8</td>
</tr>
<tr>
<td>0.7</td>
</tr>
<tr>
<td>0.6</td>
</tr>
<tr>
<td>0.5</td>
</tr>
</tbody>
</table>

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

3 Results and discussion

3.1 Logarithmic 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 \(M_{PBK+E1}\) and \(M_{PBK}\), respectively.

Figures 3, 4, 5 and 6 show the deviations represented in tables 1 and 2. Note that the number depicted in the tables are scaled in a nonuniform manner for observational purpose in the figures. Before reading the graphs, note that red indicates deviation of mean of \(Pr(\text{TRCMPLX} = \text{active}|\forall g_e\text{, evidences})\).
evidences) in normal test samples, blue indicates deviation of mean of \( \Pr(\text{TRCMPLX} = \text{active} | g_{\text{e}}, \text{evidences}) \) for normal cases. 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 \( M_{\text{PBK+EI}} \). Red - constant deviation in Normal, constant deviation in Tumor, Green - constant deviation in Weber’s law, Cyan - constant deviation in Bernoulli’s law.

For the case of constant deviations (figure 3) in model \( M_{\text{PBK+EI}} \), it was observed that deviations in activation of \( \text{TRCMPLX} \) conditional on gene evidences for tumor test samples showed a logarithmic behaviour and were directly proportional to the negative of both the Weber’s and Bernoulli’s law. This can be seen by the blue curve almost following the green and cyan curves. For the case of deviations in activation of \( \text{TRCMPLX} \) conditional on gene evidences for the normal test samples showed an exponential behaviour and were proportional to the negative of both the Weber’s and Bernoulli’s law. Similar behaviour was observed for all the coloured curves in case of incremental deviations as shown in figure 4. The exponential behaviour for activation of \( \text{TRCMPLX} \) being active conditional on gene evidences correctly supports to the last line of equation 6, which is the derivation of Weber’s law from Bernoulli’s equation. It actually indicates deviations in activation of \( \text{TRCMPLX} \) 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 \( \text{TRCMPLX} \) 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 \( \text{TRCMPLX} \) 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 \( M_{\text{PBK}} \) does not capture and constrain the network as much as \( M_{\text{PBK+EI}} \) 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 Goentoro and Kirschner\(^{16}\), the presented results might be crude in terms of static observations, yet they show corresponding be-

---

**Fig. 3** Constant deviations in \( \beta \) i.e \( \text{ETGN} \) and corresponding deviations in \( \Pr(\text{TRCMPLX} = \text{active} | g_{\text{e}}, \text{evidences}) \) 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 \( M_{\text{PBK+EI}} \). Red - constant deviation in Normal, constant deviation in Tumor, Green - constant deviation in Weber’s law, Cyan - constant deviation in Bernoulli’s law.

**Fig. 4** Incremental deviations in \( \beta \) i.e \( \text{ETGN} \) and corresponding deviations in \( \Pr(\text{TRCMPLX} = \text{active} | g_{\text{e}}, \text{evidences}) \) 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 \( M_{\text{PBK+EI}} \). 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 $Pr(\text{TBCMPLX} = \text{active}|\text{ge})$ 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.

Fig. 6 Incremental deviations in $\beta$ i.e. ETGN and corresponding deviations in $Pr(\text{TBCMPLX} = \text{active}|\text{ge})$ 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 - incremental deviation in Normal, incremental deviation in Tumor, Green - incremental deviation in Weber’s law, Cyan - incremental deviation in Bernoulli’s law.

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 $Pr(g_i = \text{active}|g_k \text{ evidence}) \forall k$ 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 $g_{di}$ 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 $Pr(g_i = \text{active}|g_k \text{ evidence}) \forall k$ 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 $Pr(g_i = \text{active}|g_k \text{ evidence}) \forall k$ 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 $ge_k$ 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 $ge_k$ evidence $\forall k$ genes in the $j^{th}$ run for Tumor test sample. This creates a matrix at the end of the runs.

2. assign variables $ge, aa_N, ar_N, ra_N, rr_N, aa_T, ar_T, ra_T, rr_T$ to an empty vector $. Note - a (r) means activation (repression).

3. $\forall k$ genes except the $j^{th}$ one
   
   (a) if($k \neq i$)
      
      i. assign variables $tmpra_N, tmprT, tmpra_T, tmpr_T$ and $tmpr_{T}$ to 0.

haviour 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$^1$. The observations presented here might bolster the existence of behavioural phenomena in terms of logarithmic laws and its special cases.
**SFRP5 activation status apropos to gene evidences in Normal and Tumor samples**

<table>
<thead>
<tr>
<th></th>
<th>ge</th>
<th>aa_N</th>
<th>ar_N</th>
<th>ra_N</th>
<th>rr_N</th>
<th>aot</th>
<th>arp</th>
<th>raT</th>
<th>rrT</th>
<th>gg1N</th>
<th>gg1T</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DKK1</td>
<td>0</td>
<td>360</td>
<td>216</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>360</td>
<td>216</td>
<td>DKK1</td>
<td>DKK1</td>
</tr>
<tr>
<td>2</td>
<td>DKK2</td>
<td>360</td>
<td>0</td>
<td>0</td>
<td>216</td>
<td>0</td>
<td>0</td>
<td>216</td>
<td>360</td>
<td>DKK2</td>
<td>DKK2</td>
</tr>
<tr>
<td>3</td>
<td>DKK3-1</td>
<td>360</td>
<td>0</td>
<td>0</td>
<td>216</td>
<td>0</td>
<td>0</td>
<td>216</td>
<td>360</td>
<td>DKK3-1</td>
<td>DKK3-1</td>
</tr>
<tr>
<td>4</td>
<td>DKK3-2</td>
<td>133</td>
<td>336</td>
<td>107</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>336</td>
<td>240</td>
<td>DKK3-2</td>
<td>DKK3-2</td>
</tr>
<tr>
<td>5</td>
<td>DKK4</td>
<td>0</td>
<td>480</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>460</td>
<td>116</td>
<td>DKK4</td>
<td>DKK4</td>
</tr>
<tr>
<td>6</td>
<td>DACT1</td>
<td>346</td>
<td>230</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>216</td>
<td>360</td>
<td>DACT1</td>
<td>DACT1</td>
</tr>
<tr>
<td>7</td>
<td>DACT2</td>
<td>312</td>
<td>218</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>0</td>
<td>264</td>
<td>312</td>
<td>DACT2</td>
<td>DACT2</td>
</tr>
<tr>
<td>8</td>
<td>DACT3</td>
<td>504</td>
<td>0</td>
<td>0</td>
<td>72</td>
<td>0</td>
<td>0</td>
<td>69</td>
<td>507</td>
<td>DACT3</td>
<td>DACT3</td>
</tr>
<tr>
<td>9</td>
<td>SFRP1</td>
<td>552</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>46</td>
<td>530</td>
<td>SFRP1</td>
<td>SFRP1</td>
</tr>
<tr>
<td>10</td>
<td>SFRP2</td>
<td>480</td>
<td>0</td>
<td>0</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>96</td>
<td>480</td>
<td>SFRP2</td>
<td>SFRP2</td>
</tr>
<tr>
<td>11</td>
<td>SFRP3</td>
<td>484</td>
<td>0</td>
<td>0</td>
<td>92</td>
<td>0</td>
<td>0</td>
<td>96</td>
<td>480</td>
<td>SFRP3</td>
<td>SFRP3</td>
</tr>
<tr>
<td>12</td>
<td>SFRP4</td>
<td>82</td>
<td>312</td>
<td>182</td>
<td>0</td>
<td>312</td>
<td>264</td>
<td>0</td>
<td>0</td>
<td>SFRP4</td>
<td>SFRP4</td>
</tr>
<tr>
<td>13</td>
<td>WIFI</td>
<td>0</td>
<td>408</td>
<td>168</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>398</td>
<td>178</td>
<td>WIFI</td>
<td>WIFI</td>
</tr>
<tr>
<td>14</td>
<td>LEFl</td>
<td>0</td>
<td>480</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>484</td>
<td>92</td>
<td>LEFl</td>
<td>LEFl</td>
</tr>
<tr>
<td>15</td>
<td>MYC</td>
<td>0</td>
<td>456</td>
<td>120</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>442</td>
<td>134</td>
<td>MYC</td>
<td>MYC</td>
</tr>
<tr>
<td>16</td>
<td>CCND1</td>
<td>0</td>
<td>480</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>480</td>
<td>96</td>
<td>CCND1</td>
<td>CCND1</td>
</tr>
<tr>
<td>17</td>
<td>CD44</td>
<td>0</td>
<td>376</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>384</td>
<td>192</td>
<td>CD44</td>
<td>CD44</td>
</tr>
</tbody>
</table>

**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, gg1N - gene-gene interaction with Normal, gg1T - gene-gene interaction with Tumor, <= - active and | - repressed.

ii. ∀ 576 runs iterated by a counter l
A. if (geN[l,k] == 1 and xN[l,k] < 0.5) increment tmprlN by 1
B. else if (geN[l,k] == 1 and xN[l,k] == 0.5) increment tmpraN by 1
C. else if (geN[l,k] == 2 and xN[l,k] < 0.5) increment tmpraN by 1
D. else if (geN[l,k] == 2 and xN[l,k] == 0.5) increment tmpraN by 1
E. if (geT[l,k] == 1 and xT[l,k] < 0.5) increment tmptrT by 1
F. else if (geT[l,k] == 1 and xT[l,k] == 0.5) increment tmpraT by 1
G. else if (geT[l,k] == 2 and xT[l,k] < 0.5) increment tmpraT by 1
H. else if (geT[l,k] == 2 and xT[l,k] == 0.5) increment tmpraT by 1

iii. Comment - store results
iv. append ge with geN, rrN with tmprN, arN with tmpraN, raN with tmpraN, aot with tmpraN, rrT with tmprT, arT with tmpraT, raT with tmpraT and aot with tmpraT
(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 gg1N and gg1T as empty vector []
(e) ∀ j gene except the ith one under consideration
i. find the index idxN in stats that corresponds to 1 or 2 or 3 or 4
ii. if (idxN == 1) append gg1N with interaction string statsgjN | | gg1N
iii. else if (idxN == 2) append gg1N with interaction string statsgjN | | gg1N
iv. else if (idxN == 3) append gg1N with interaction string statsgjN | | gg1N
v. else if (idxN == 4) append gg1N with interaction string statsgjN | | gg1N
vi. find the index idxN in stats that corresponds to 1 or 2 or 3 or 4
vii. if (idxT == 1) append gg1T with interaction string statsgjT | | gg1T
viii. else if (idxT == 2) append gg1T with interaction string statsgjT | | gg1T
ix. else if (idxT == 3) append gg1T with interaction string statsgjT | | gg1T
x. else if (idxT == 4) append gg1T with interaction string statsgjT | | gg1T
(f) assign statsg1N with gg1N
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.

Based on the above execution, for each gene a matrix is obtained that shows the statistics of how the status of gene is affected conditional on the individual evidences of the remaining genes. 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, $aa_T$ - $SFRP5$ is active (a) when a gene is active (a) in Tumor (T) sample, $ar_T$ - $SFRP5$ 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, $gg_{IN}$ - interaction of $SFRP5$ given the gene evidence based on majority voting among $aa_N$, $ar_N$, $ra_N$ and $rr_N$ and finally, $gg_{IT}$ - 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 $rr_N$ ($aa_T$, $ar_T$, $ra_T$ and $rr_T$) confirms the relation between genes using Normal (Tumor) samples. Active (repressed) for $SFRP5$ is based on discretization the predicted conditional probability $Pr(SFRP5 = \text{active}|g_j \text{ evidence})$ 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 $gg_{IN}$ and $gg_{IT}$, $<>$ 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 $DKK2$, $DKK3 - 1$,
Gene–gene interactions

| DACT2 | ⋯ | DKK1, SFRP4 | ⋯ | DKK1, DACT1 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | DKK2, SFRP1 | ⋯ | ⋯ | ⋯ | DKK2, DKK4 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ |
| DKK4− | ⋯ | DKK3−1, DACT3 | ⋯ | ⋯ | ⋯ | ⋯ | DKK3−1, SFRP2 | ⋯ | DKK3−1, SFRP3 | ⋯ | ⋯ | ⋯ | DKK3−1, CD44 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ |
| ⋯ | ⋯ | ⋯ | ⋯ | DKK3−2, DKK2 | ⋯ | ⋯ | DKK3−2, DKK3−1 | ⋯ | ⋯ | ⋯ | ⋯ | DKK3−2, DACT1 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ |
| ⋯ | ⋯ | DKK3−2, SFRP1 | ⋯ | ⋯ | DKK3−2, SFRP4 | ⋯ | ⋯ | DKK3−2, DKK2 | ⋯ | DKK3−2, DKK3−1 | ⋯ | ⋯ | DKK3−2, DACT1 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ |
| DKK4, SFRP3 | ⋯ | DKK4, SFRP5 | ⋯ | DKK4, WIFI1 | ⋯ | DKK4, LEF1 | ⋯ | DKK4, MYC | ⋯ | DKK4, CCND1 | ⋯ | DKK4, CD44 | ⋯ | DKK4, DACT1 | ⋯ | DKK4, DACT3 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ |
| DKK4, CD44 | ⋯ | DKK4, DKK4 | ⋯ | DKK4, DACT1 | ⋯ | DKK4, MYC | ⋯ | DKK4, CD44 | ⋯ | DKK4, DACT1 | ⋯ | DKK4, CCND1 | ⋯ | DKK4, DAKT1 | ⋯ | DKK4, DCT3 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ |
| DKK4 | ⋯ | DKK3−1 | ⋯ | DKK3−1, WIFI1 | ⋯ | DKK3−1, LEF1 | ⋯ | DKK3−1, MYC | ⋯ | DKK3−1, CCND1 | ⋯ | DKK3−1, CD44 | ⋯ | DKK3−1, DACT1 | ⋯ | DKK3−1, DACT3 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ |
| DKK3−1 | ⋯ | DKK3−1, DKK3−1 | ⋯ | ⋯ | ⋯ | ⋯ | DKK3−1, SFRP2 | ⋯ | DKK3−1, SFRP3 | ⋯ | DKK3−1, CD44 | ⋯ | DKK3−1, DACT1 | ⋯ | DKK3−1, DACT3 | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ | ⋯ |

**Table 4** Tabulated gene–gene interactions of figure 7 using $M_{DBK+EI}$ obtained in case of Normal samples. Here, the symbols represent the following - ⋯ activation and | ⋯ repression/suppression. Note that for Tumor cases, the interaction roles were found to be reversed, i.e. ⋯ | ⋯ in normal became ⋯ ⋯ in tumor, ⋯ ⋯ in normal became ⋯ ⋯ in tumor, ⋯ ⋯ in normal became | ⋯ in tumor and | ⋯ in normal became ⋯ ⋯ in tumor.

$DACT - 1/2/3$ 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 $SFRP4$ and $SFRP5$. They are both SFRP5 is activated in Normal test sample and $SFRP4$ is activated 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) that might corrupt prediction values as can be seen in the columns of $aa_N (aa_T)$, $ar_N (ar_T)$, $ra_N (ra_T)$ and $rr_N (rr_T)$ 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 already 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 and finally, $g_j > - g_i$ in Normal if and only if $g_j < - g_i$. This restricts the network to only reversible gene–gene interactions in Normal and Tumor cases.

Lastly, duplicate interactions are removed from the networks for both Normal and Tumor cases. 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 $(g_j < - g_i)$ and $(g_i < - g_j)$ and $(g_i < - g_j)$. Figure 7 shows one such network after com-
Deviations in gene-gene interactions for different values of ETGN

<table>
<thead>
<tr>
<th>ETGN</th>
<th>Changes in Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>90N-T1</td>
<td>MYC</td>
</tr>
<tr>
<td>80N-T1</td>
<td>SFRP5</td>
</tr>
<tr>
<td>70N-T1</td>
<td>DACT3</td>
</tr>
<tr>
<td>60N-T1</td>
<td>SFRP5</td>
</tr>
<tr>
<td>50N-T1</td>
<td>DACT1, MYC</td>
</tr>
</tbody>
</table>

Table 5 Tabulated missing gene-gene interactions of figure 7 using $M_{PBK+E1}$ 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, i.e. - in normal became | in tumor. | in normal became - | in tumor, | | in normal became - | in tumor and - in normal became | - | in tumor.

complete interaction construction, inference and removal of duplicate interactions in using Normal test samples with ETGN of 90% in $M_{PBK+E1}$. For the case of Tumor test samples with ETGN 90% in $M_{PBK+E1}$, 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 TRC MPLX (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 $\frac{n \times (n-1)}{2}$, 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.

## 4 Future Work

In context of the above observations, dynamic models might reveal greater information regarding the psychophysical laws. Work by Goentoro and Kirschner\cite{16} 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 8 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 $\text{Pr}(\Delta N(t+1)|N(t+1), N(t+2), \Delta Wnt, \Delta APC)$.

Regarding sensitivity analysis, in nonlinear problems, it might be useful to use Sobol\cite{22} 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\cite{23}). Finally, with respect to the robustness of the gene-gene interaction network, the cur-
Fig. 8 A schematic diagram of a dynamic bayesian network model that might help study the fold change and the logarithmic psychophysical laws behind the changes.

recent 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.

Conflict of interest

None

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. (3) Netherlands Bioinformatics Centre (NBIC) for funding the project. (4) Dr. ir. R. H. J. Fastenau (Dean of Faculty of Electrical Engineering Mathematics and Computer Science), Dr. Prof. Ir. K. Ch. A. M. Luyben (Rector Magnificus) and Drs. D. J. van den Berg (President) at Delft University of Technology, for providing support for conducting this work and giving permission to submit the manuscript. (5) Dr. Wim Verhaegh (faculty at Netherlands Bioinformatics Centre and a principal scientist at Molecular Diagnostic Lab in Philips Research) for providing technical details of Naive Bayes model for replicating experiments in Verhaegh et al.\(^2\). (6) Dr. Marcel J. T. Reinders (scientific director of Bioinformatics Research at Netherlands Bioinformatics Centre and a professor at Delft University of Technology) for refining
the proposed model and suggesting inclusion of methylation data. (7) Dr. Robert P.W. Duin (retired associate professor at Delft University of Technology) for informal discussions on 2-holdout experiments and (8) Dr. Jeroen de Ridder (assistant professor at Delft University of Technology) and (9) Phd candidate Marc Hulsman (Delft University of Technology), for informal discussions on estimation of probability values on methylation data. (10) Lastly, 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.

References

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