Drug Targets and Biomarker Identification from Computational Study of Human Notch Signaling Pathway

Saikat Chowdhury\textsuperscript{1,2} and Ram Rup Sarkar\textsuperscript{1,2*}

\textsuperscript{1}Chemical Engineering and Process Development Division, CSIR-National Chemical Laboratory, Pune, Maharashtra, 411008, India
\textsuperscript{2}Academy of Scientific & Innovative Research (AcSIR), New Delhi, 110 001, India

Abstract

Notch signaling pathway is widely implicated in controlling various cellular functions, cell fate determination, and stem cell renewal in human but aberrant activity in cancer stem cells may cause different types of cancers. Understanding the complexity of this pathway to identify important targets for cancer therapy and to suppress the pathway activity without affecting the normal functions is of utmost importance to clinical and experimental pharmacologists. For developing therapeutic strategy, non availability of detailed molecular interactions, complex regulations and cross talks with other pathways pose a serious challenge to get a coherent understanding of this pathway. This motivated us to reconstruct the largest human cell specific Notch pathway with more number of molecules and interactions available from literatures and databases. To identify probable drug targets and biomarkers for cancer prognosis, we also performed computational study of the pathway using structural and logical analysis and identified important hub proteins, cross talks and feedback mechanisms. The model simulation is validated using reported mRNA expression profile in Glioblastoma cell line and the predictions not only show significant accuracy but also able to identify the underestimated expressions. From our simulation, to identify novel combinations of drug targetable proteins and better substitute for GAMMA SECRETASE inhibition, we proposed two alternative scenarios: partial suppression of Notch target proteins by NICD1 & HIF1A; and complete suppression by NICD1 & MAML, in Glioblastoma cell line. This reconstructed Notch signaling pathway and the computational analysis for identifying new biomarkers and combinatorial drug targets will be useful for future in-vitro and in-vivo analysis to control different cancers.

Keywords: Notch signaling pathway; Reconstruction; Structural and logical analysis; Glioblastoma; Drug targets and biomarkers identification

Introduction

Notch signaling pathway is a conserved developmental pathway which is involved in cell fate determination at the stage of embryonic development and also plays important roles in tissue and organ development, hematopoiesis, vascular growth etc., [1-3]. Mostly activated by the membrane associated ligands of juxtaposed cells, this pathway then triggers most of the cellular functions including transcription of various genes, cell cycle progression, anti-apoptosis and simultaneously regulates other signaling pathways through the cross talks [4,5]. Activation of this pathway requires the involvement of two nearly adjacent cells, one of which acts as a transducer of the signal and the other receives and processes that signal [6] along with several intermediate reactions and further production of Notch intracellular domain (NICD) [6-9]. Translocation of NICD into the nucleus induces various cellular functions including cell division, cell cycle progressions, cell growth etc., through transcription of various Notch target genes and activation of other signaling and metabolic pathways in cell [10-24]. Hence, activation of this pathway is utmost important in stem cell development and cellular growth, and should be highly tuned with the activity of other pathways [25].

Sometimes this concerted process is found to work in a wrong direction [25-27], and failures in its normal functional activities can cause development of various types of cancers, such as, glioblastoma, breast cancer, osteosarcoma, prostate cancer and melanoma [28,29]. Most of the cancer cell lines have shown significant level of up-regulation of its activator proteins (Onco-proteins) and down regulations of its tumor suppressor proteins [30]. The “gain or loss” of functions of these Notch pathway associated proteins have clearly proved its correlation with cancer development, and hence can be used as a biomarker for cancer diagnosis. Various molecular biology experiments have also shown that inhibition of the activators of this pathway can drastically reduce the cancer progression in different stages [31], hence identification of drug targetable proteins and their small molecule inhibitors in the pathway to reduce various types of cancer development has always been an important field of research to the pharmacists and clinical biologists [31]. Recent identification of GAMMA SECRETASE as a probable drug target in the pathway is found to reduce the Notch pathway activity by not allowing it to cleave the Notch receptor in the membrane [32]. But, unfortunately the compound Semagacestat (LY450139), which inhibits the GAMMA SECRETASE and was in the Phase III trial of Alzheimer’s disease, failed to meet the desired goal as it was compromising with several risk factors including Skin cancer [33]. Also, the drugs developed against other target molecules, such as NOTCH1, NOTCH4, DLL4, NRARP etc., have not shown desirable result due to their toxicity and side effects [31]. Hence, inhibition of Notch pathway is now under the inspection of several cancer pathologists and the merits of targeting the Notch pathway have raised numerous questions as certain imbalance of this pathway can impose long term side effects such as, gastrointestinal toxicity and diarrhea [34].

On the other hand, identification of suitable and alternative drug targets for inhibition of this pathway in various cancers including Glioblastoma is undoubtedly useful and effective for cancer therapy...
[35], but requires the understanding of the exact mechanisms that are governing the normal functions of Notch signaling pathway in functional cells. Numerous experiments on different regulations, feedback loops, and cross talks of this pathway have been reported time to time by several research groups [16,21,22,25]. But unfortunately, the integrations of these experimental findings have not been performed properly and none of the signaling pathway database provides this extensive and up to date information and hence it has become impossible to predict the consequence of the inhibition of this pathway in a diseased situation. Moreover, study of the effects of several drug targets from a population of large number of proteins is also difficult through in-vitro and in-vivo analysis. Recent advancements in computational approaches, bioinformatics tools, and mathematical methods have contributed immensely in the understanding and analysis of large signaling pathways [36-38] and were useful to answer several biological questions in signaling systems including identification of important molecules/proteins as well as alternative combinatorial drug targets for Glioma, Colon and Pancreatic cancer [39]. Among several computational tools, structural analysis using graph theoretical methods [40] and logical analysis using Boolean formalisms [39,41,42] have shown promising results for visual and/or topological interpretation of a very large complex network and identification of important target proteins, particularly when the kinetic parameters for pathway reactions are not available. Unfortunately, very few computational studies have been performed on Notch signaling pathway to study the dynamic of the network, but none of the studies addressed this kind of problems [43-45].

This article is specifically addressed to study the complexity of the human cell specific Notch signaling pathway and to identify as well as propose alternative solutions to control cancer situation. For this, we have curated a most up-to-date and comprehensive human cell specific Notch signaling pathway interactions data by assembling the experimental results found from numerous literatures and signaling pathway databases. Using this collated information, we have reconstructed a new Notch signaling pathway map with 115 molecules and 231 interactions or reactions, the largest number of molecular entities and their corresponding interactions to the best of our knowledge. Further, we have performed computational study using graph theoretical and logical analysis to model the reconstructed pathway and identify “Hub” proteins for alternative drug targets in place of GAMMA SECRETASE complex. Using the master logical model and varying the logical states of the input molecules of the pathway, we simulated different scenarios such as Normal Notch, Glioblastoma, GAMMA SECRETASE inhibition, and two proposed in-silico combinatorial drug treated scenarios. From these two analyses we have been able to identify some important proteins, which have high centrality values and could be used as probable drug targets. In order to validate our in-silico logical model, the expression results obtained from the logical analysis have been compared with the expressions profile of mRNA found in experimental study on human Glioblastoma cell line [30].

The proposed model is able to predict the expression of the proteins with significant accuracy and also been able to predict the expression of some other proteins whose expressions were not obtained successfully in the experimental study on Glioblastoma cell line. Also, we are able to match the expressions of some Notch pathway proteins for GAMMA SECRETASE treated Glioblastoma cell with our in-silico model. From our perturbation study in the proposed drug treated scenario we have predicted two minimal and novel combinations of proteins, which are useful to suppress the expressions of Notch target proteins partially or completely in Glioblastoma cell model and may be used as better therapeutic targets for cancer. Before performing further in-vitro and in-vivo experiments, this model and computational study of the Notch signaling pathway is not only helpful for identification of drug targets and biomarkers for cancer prognosis but also show new direction to the clinical and experimental pharmacologists in a faster and cost effective way to identify the probable, and safe drug targets in-silico.

Materials and Methods
Reconstruction of Notch Pathway Map
In order to reconstruct the human cell specific Notch signaling pathway, we searched around 28 available and popular databases on cell signaling, protein-protein interaction, cancer pathway, and microarray expressions (Table S1). Although the data available in these databases are mostly scattered and heterogeneously presented (Table S2), but this extensive and elaborative database searching gave us the basic information of the pathway, core proteins and the connections among its associated proteins/molecules, which are involved in the Notch signal transduction network and also its functional cross talks with other cell signaling pathways. Unfortunately there was no database available which gives complete and most up to date Notch pathway information along with cross talk molecules of other pathways. Therefore, in order to reconstruct a master pathway model of Notch signaling network, we used the core structure of Notch pathway available from the databases and collated additional information from different literatures and experimental reports (Table S1) to build a comprehensive, up to date and the largest human cell specific Notch signaling pathway to the best of our knowledge.

The Notch pathway shown in Figure 1 is completely based on manual curation of data from various databases and literature resources. In order to incorporate a new molecule or interaction, we set few criteria e.g., the newly inserted molecules should have at least one direct or indirect connection or interaction with the core Notch pathway molecules, all the newly inserted interactions should have at least one experimental evidence in a peer reviewed journal and all the molecules should be placed in the pathway map according to the specified locations i.e., extra-cellular and membrane region, cytoplasmic, nucleus and output. Different color codes were used to distinguish the protein molecules and interactions according to their locations and type of reactions respectively. The pathway map was drawn in CellDesigner Ver. 4.2, an open source “Systems Biology Marked Up Language” (SBML) based pathway illustrator software [46,47].

Pathway drawing and annotation
We first annotated the molecules of the pathway according to their sub cellular locations in the cell. In case of Notch signaling pathway we had to consider three sub-cellular locations: Extracellular and Membrane, Cytoplasm and Nucleus of Notch signal “Receiver Cell”. In order to reduce the complexity and create a simple pathway map, we considered only these three sub-cellular locations, but one can also consider other sub-cellular locations such as Golgi body, Mitochondria, Endoplasmic reticulum etc.. Moreover, as notch pathway is mostly activated by the ligands expressed by the neighboring cell, therefore we had to consider another cell membrane of Notch signal “Transmitter cell” to allocate the ligands [6]. In between these two membrane regions we also annotated a place for extracellular region.

In the Extracellular and Membrane region, we annotated 27 molecules including 4 Notch receptors (NOTCH1/2/3/4), 9 ligands...
Figure 1: Reconstructed human cell specific Notch signaling pathway. See "Pathway Drawing and Annotation" part of the materials and method section for details of the color scheme used.

(JAG1/2, DLL1/3/4, MAGPI/2, NOV, CNTN1) molecules, 6 proteolytic enzyme complex including TACE, GAMMA SECRETASE complex etc., and the truncated portions of four Notch receptors (NEXT1/2/3/4 and NECD1/2/3/4) [48-61]. All the proteins (except metalloprotease enzymes TACE and GAMMA SECRETASE) of this location were colored as “Green” to distinguish these proteins from all other proteins in the network. In order to differentiate the proteins from all the membrane associated receptor proteins, the shape of the receptors proteins were chosen “Hexagonal” (a standard SBML CellDesigner notation) [46,47]. The metalloprotease enzymes of this location were colored “Light blue”.

In the Cytoplasm region, we included total 35 molecules, out of which 5 molecules are metabolic compounds such as O-linked glucose,
different types of scenarios were created: Normal Notch Pathway was followed from previous study in this direction [39]. Total five of formation of logical equations of the signaling pathway molecules and its associated molecules. The entire logical model is available in the basis of biological and experimental knowledge about the pathway and signaling network and the corresponding parameter values are given A detail description and definitions of the construction of Notch built algorithms implemented in these software applications [69,70].

The Nuclear region was annotated mostly with the transcription factors, co-activators and co-repressor molecules and complexes found at the time of data curation. There were 23 such proteins (NICD1/2/3/4, CSL, SMAD3 etc.) and 2 transcription complexes (Co-activator and Co-repressor complex) considered [12-14]. The Co-activator complex (COA) and Co-repressor complex (COR) were colored as "Green" and "Orange", respectively. Other proteins were colored as "Dark Cyan". Also, in order to reduce the complexity we did not consider any gene or mRNA in this pathway map.

Moreover, except these three sub-cellular locations, we also annotated the Notch target proteins into a group called "Output". All the target proteins (total 28) in this group may belong to any sub-cellular locations depending on their functional activity. Few of them (e.g., HES and HEY proteins) are also the transcription factors for another genes [66]. All the proteins in this group are colored by "Yellow" color. We also linked these proteins with their phenotypic and functional activities (e.g., Transcription, Myelination, Cell Division, Anti-Apoptosis, Hypoxia etc.) [67,68]. Simultaneously, we also used specific color codes to distinguish the molecular reactions in the pathway on the basis of the type of the reactions. There were total 10 different types of reactions presented by colored arrows. Black: Physical interactions/ Complex Formation; Blue: Enzymatic reactions; Dotted Blue: Membrane Translocation; Brown: Proteolytic Cleavage; Red: Inhibition; Green: Activation; Dotted Green: Indirect Activation; Orange: Phosphorylation; Dotted Black: Nuclear Translocation; Violet: Protein Production, were considered at the time of drawing the pathway map.

Structural analysis

The structural and topological features of the reconstructed Notch pathway (Figure 1) were studied using ‘Graph theory’, which has shown promising results for visual and/ or topological interpretation of a very large complex network [39], despite its non-logical and static nature. The graph theoretical analysis was performed in open source software Gephi and igraph [69,70]. In order to identify the central nodes in the network, four types of centrality analysis were performed in CellNetAnalyzer [42,71,72], a brief description of which is presented in Table S2 of Supplementary file.

Results

Reconstructed Notch signaling network

Pathway statistics: In this work, we have reconstructed a comprehensive, most up to date and largest human cell specific Notch signaling pathway to the best of our knowledge. The entire Notch pathway (Figure 1) was annotated and reconstructed manually by collating the data from various literatures, experimental findings and biological databases. Although the basic core pathway is same as the pathway map available in the existing signaling databases, but to the best of our knowledge the newly reconstructed pathway map consists most up-to-date information of interaction data which is not available in any freely available major academic databases. In this reconstructed pathway we included 115 molecules (96 core and 19 cross talking pathway molecules including proteins and organic compounds) and 231 molecular interactions. Different types of molecular reactions such as Physical interaction, Enzymatic reactions, Phosphorylation, Protein production, Activation, Inhibition, Nuclear translacation etc., were also considered to construct the pathway map. A comparison between this reconstructed Notch pathway data (i.e., molecules and interactions) with the pathway information from other major biochemical signaling databases (e.g., KEGG, Biocarta, Netpath etc..) is presented in Table S2 of Supplementary file.

Description of the reconstructed notch pathway

Extracellular & membrane reactions: Several experimental results have identified that all the four NOTCH receptors can express in the cell membrane of a Notch "signal receiver" cell, whereas the membrane bound ligands JAG/DLL class proteins are expressed in the membrane of Notch signal "Transmitter cell" [6]. In the reconstructed pathway map (Figure 1), we have shown these proteins in the membrane regions of both cells. Moreover, besides these juxtracrine signaling property, Notch pathway can also be activated by the interactions between microfibrillar proteins MAGP1 and MAGP2, by CONTACTIN/ F3 (CNTN1) or by Nephroblastoma overexpressed (NOV) proteins [52-55] and were also shown separately in the extracellular regions in the pathway map. However, all these kind of ligand-receptors interactions are followed by the common proteolytic cleavage of NOTCH receptors and subsequent formation of Notch Extracellular Domain (NECD) and Notch Extracellular Truncated Protein (NEXT) [6]. These proteolytic cleavage reactions were shown by brown arrows in the figure emanating from NOTCH1/2/3/4 receptor proteins. A metalloprotease enzyme TACE helps in this type of reactions to cleave the Notch receptors [7]. The blue arrows that are connecting TACE and the four NOTCH1/2/3/4 proteins; were used to show this enzymatic reaction in the pathway map (Figure 1). Subsequently, NEXT1/2/3/4 is
again cleaved by another proteolytic enzyme GAMMA SECREATASE complex [8,32], which was also included in the Figure 1.

Reactions in cytoplasm: Followed by the GAMMA SECRETASE mediated reactions, four NEXT proteins produce four homologues of Notch Intracellular Domains (NICD1/2/3/4) which then translocate into the cytoplasm and further moves towards the nucleus [6]. During this travel through cytoplasmic region, NICD encounters with various activator (RAS, GSK_3BETA, WDR12) and inhibitor (DVL, JIP1) proteins [62,63,73,74]. RAS pathway proteins activate NICD1 through some intermediate proteins which were not included in the pathway map for the sake of simplicity. This activation process was shown by dotted green arrow in Figure 1. Moreover the cytoplasmic region (or specifically the Golgi body) is also the place for post translation modification of Notch precursor proteins (NOTCH1_PRE, NOTCH2_PRE, NOTCH3_PRE and NOTCH4_PRE) before they express in the cell membrane. In these reactions, notch precursors pass through several glycosylation or fucosylation reactions by Glucose, Galactose, Fucose and the enzymes POGlut_1, FRINGE, GASE, PO Fut_1 etc.. These post translational modifications of Notch precursors increase the specificity of ligand receptors interactions, so that it can easily recognize and interact with Notch ligands [75-78]. On the other hand Xylosylatin by Xylose with the help of the enzyme Xylosyltransferase (XYLE) is also observed in several cases which in turn reduce the specificity of notch ligand bindings [75]. All these enzyme-substrate reactions were included in the pathway map (Figure 1) and the connections were represented by blue colored arrows.

Reactions in nucleus: After encountering with several modifications and interactions by nuclear proteins, the activated NICD1/2/3/4 enters and starts the transcription process in the nucleus [6,62,63,73,74]. NICD initiates its transcription by binding with another transcription factor CSL, which in general forms a transcription repressor complex with another transcription Co-repressor complex (COR). It is a complex of SMRT, SAP30, HDAC, CIR, SIN3A proteins in the nucleus [11-14]. On the other hand, there is another protein complex which acts as a transcription co-activator of CSL to transcribe Notch target genes/ proteins such as HES1, HES5, HEY1, HEY2, HEYL, BCL2, P65, NOTCH1/2/3/4 etc., (colored as Yellow) [15-21]. Besides this CSL/ NICD/COR/COA mediated transcription, it is also reported that Notch pathway can also be activated through CONTACTIN/F3 (CNTN1) mediated interaction, which requires the involvements of DTX1 as a pathway can also be activated through CONTACTIN/F3 (CNTN1) and inhibitor (DVL, JIP1) proteins [62,63,73,74]. This activation process was shown by dotted green arrow in Figure 1. Moreover the cytoplasmic region (or specifically the Golgi body) is also the place for post translation modification of Notch precursor proteins (NOTCH1_PRE, NOTCH2_PRE, NOTCH3_PRE and NOTCH4_PRE) before they express in the cell membrane. In these reactions, notch precursors pass through several glycosylation or fucosylation reactions by Glucose, Galactose, Fucose and the enzymes POGlut_1, FRINGE, GASE, PO Fut_1 etc.. These post translational modifications of Notch precursors increase the specificity of ligand receptors interactions, so that it can easily recognize and interact with Notch ligands [75-78]. On the other hand Xylosylatin by Xylose with the help of the enzyme Xylosyltransferase (XYLE) is also observed in several cases which in turn reduce the specificity of notch ligand bindings [75]. All these enzyme-substrate reactions were included in the pathway map (Figure 1) and the connections were represented by blue colored arrows.

Reactions in nucleus: After encountering with several modifications and interactions by nuclear proteins, the activated NICD1/2/3/4 enters and starts the transcription process in the nucleus [6,62,63,73,74]. NICD initiates its transcription by binding with another transcription factor CSL, which in general forms a transcription repressor complex with another transcription Co-repressor complex (COR). It is a complex of SMRT, SAP30, HDAC, CIR, SIN3A proteins in the nucleus [11-14]. On the other hand, there is another protein complex which acts as a transcription co-activator of CSL to transcribe Notch target genes/ proteins such as HES1, HES5, HEY1, HEY2, HEYL, BCL2, P65, NOTCH1/2/3/4 etc., (colored as Yellow) [15-21]. Besides this CSL/ NICD/COR/COA mediated transcription, it is also reported that Notch pathway can also be activated through CONTACTIN/F3 (CNTN1) mediated interaction, which requires the involvements of DTX1 as a transcription co-activator to produce the proteins MAG. This protein is involved in the oligodendrocyte maturation and myelination [54].

Cross talks with other pathways: Notch pathway has cross connections with different signaling pathways such as, JAK/STAT, PTEN/P13K/akt, RAS/MAPK, TGF/SMAD3, CYCLIN/CDK, HYPOXIA/HIF1A, BCL2/IAP/ ANTI-APOPTOSIS, P65/P50/NFKB etc., which widen the scope of the study [10,17,22,24,44,64,65, 79-84]. We calculated all these network parameter values for each proteins/molecules that form “Hub” molecules in the network, we used the connectivity and centrality measurement parameters of the network such as Degree, Closeness, Betweenness, and Eigenvector centrality [39,86-88]. We calculated all these network parameter values for each protein of the network and extracted the significant proteins that had the parameter values higher than the corresponding average values. The values of network parameters are available in the “Supplementary” file. In Table 1, we have presented only the extracted and important proteins from the network analysis of Notch signaling pathway.

The extracted proteins enlisted in Table 1 helped us to identify the Hubs as well as the important centrally situated proteins from the network. In case of In-Degree, we identified almost all the four types of Notch receptors, Notch precursors and Notch Intracellular domains proteins were showing high In-Degree values compared to the other proteins in the network (more than the average value, 1.97). Again, out of those Notch receptor proteins, NOTCH1 had high values compared to the other homologues. Similarly, NOTCH1_PRECORSOR and NICD1 showed high values compared to their corresponding homologues present in the network. It signifies the importance of NOTCH1 compared to all other homologues as higher number of incoming connections or interactions are regulating this protein in the network (Figure S1 in Supplementary file, for In-Degree values).
In case of Out-Degree data, the nuclear protein CSL had highest number of Out-Degree value in the network (more than the average value 1.97). Moreover, most of the ligands as well as the enzymes, including GAMMA_SECRETASE, and the Notch post translational modifier enzymes, such as POGLUT_1, POOFUT_1, GASE, had the significant number of Out-Degree values in the network, which also signifies that activation of Notch Pathway mostly occur by the activation of these molecules in the network. On the other hand there are also some inhibitor molecules or complex, such as, Co-repressor complex (COR), HDAC, SMRT and the phosphorylated form of P53 (inhibitors of NUC_NICD1/2/3/4), which also have significant number of Out-Degree values in the network. Interestingly, only HIF1A and NRARP from the output molecules of this network had significant Out-Degree value in the network, which signifies higher number of signaling cascades passing through the shortest path of other two nodes) in the network, higher value of Betweenness centrality values of these proteins were also increased. This result also signifies that certain perturbations or mutations of these proteins will cause worst effect than the other proteins having lower closeness centrality values.

After finding the hub or important proteins of the network on the basis of Degree, Eigenvector and Closeness centrality, we measured another kind of important centrality parameter called Betweenness centrality (Figure S6 in Supplementary file). This parameter does not assign the importance of a molecule in a signaling network on the basis of number of connections but identifies the molecules on the basis of their position (the situation of a node which lies in between the shortest path of other two nodes) in the network, higher value of which signifies higher number of signaling cascades passing through a particular node implying that all biochemical reaction cascades in general prefer the shortest route to relay the signal in much more cost effective way [89,90]. As expected we found that CSL had highest Betweenness centrality value in the network as the production of all the output proteins are mediated by this protein. But surprisingly we found that NIDC1 had higher Betweenness centrality value compared to its other homologues (i.e., NICD2/3/4). In order to search the reason for this different result, we checked its connectivity profile and found that unlike the other three homologues of these proteins, NIDC1 had extra three upstream regulators proteins: RAS, JIP1 and WDR12 as well as P53 protein in downstream. It is also connected with its nuclear counterpart NUC_NICD1, which has additional downstream target genes (e.g., BCL2, FLIP, IAP, P21, P65, PSO_C_Rel, REL_B) for transcription than its counterparts NUC_NICD2/3/4. Hence, more number of other proteins in the network. As a result the closeness centrality values of these proteins were also increased. This result also signifies that certain perturbations or mutations of these proteins will cause worst effect than the other proteins having lower closeness centrality values.

Although this analysis was useful to identify the important proteins from such a large complex network and helped us to identify the probable drug targets to suppress the activity of maximum Notch target proteins, but using this simple technique we were unable to identify the exact proteins that are directly or indirectly influenced by these identified proteins, which is one of the limitations of the graph theoretical analysis. In order to overcome this drawback and to test the effect of mutation or deregulation of important proteins in the network.
under certain circumstances as well as to identify the new biomarkers of Notch pathway, we performed logical or semi-dynamic analysis of the reconstructed network.

**Logical analysis of Notch signaling pathway:** Logical analysis of Notch signaling network was performed to simulate the pathway activity and the expression of pathway proteins in Normal, Glioblastoma cell specific, Gamma Secreatase inhibitor treatment and two proposed drug treated scenarios, and also to see the logical relationship that exist among the proteins in the newly reconstructed Notch pathway and to analyze their regulations and expression patterns that vary according to the normal, disease and drug treated scenarios. As mentioned earlier, the highly interactive Notch pathway and its regulations and cross talks with other cancer forming pathways make this pathway more delicate and sensitive in cancer and tumor pathology. As certain mutation or change of a protein in this pathway can cause cancer, similarly without knowing the exact regulations targeting a protein in this pathway to treat cancer also can cause severe side effects, as it happens in case of Gamma Secreatase inhibition [33]. Therefore selective targeting with the knowledge of precise consequence is required and hence logical relationships/model among all the proteins in the network could be useful to identify the probable and safe drug targets.

The entire logical analysis of Notch pathway was performed using the logical relationships presented in the Table S3 as a master logical model. In the following subsections we present the outcome from the model simulations for different scenarios.

**Model validation:** The expression scenarios generated in the simulation for each protein in the pathway is shown in Figure 2, where GBE represents the expression of notch pathway proteins found in mRNA expression profile of Gliblastoma cell line collected from EBI-ARRAYEXPRESS database [30]. The rest of the columns (GBS, NNS, GSI, TS1 and TS2) depict the in-silico simulation results for five different types of scenarios. Figure 2A presents the expression of the input proteins of our model, whereas Figure 2B depicts the expression and simulation results of the intermediate and output proteins of our logical model. While comparing the simulation result of GBS with GBE, we found that there were total 54 input proteins and 62 intermediate and output proteins in our Notch pathway model.

![Figure 2: Expression of each protein of Notch signaling pathway in five different scenarios. mRNA expression profile of Gliblastoma Cell Line (GBE) [30], Glioblastoma (GBS), Normal Notch Pathway (NNS), Gamma Secreatase Inhibition (GSI), and two proposed drug treated scenarios; TS2: NICD1 and MAML combinatorial inhibition, and TS1: NICD1 and HIF1A combinatorial inhibition. (A) The expression of the input proteins (B) The expression and simulation results of the intermediate and output proteins](image-url)
previously done cancer stem cell and molecular biology experiments GBS and GSI scenarios with the experimental results obtained from reduced by administering the GAMMA SECRETASE inhibition in GBS downstream activated proteins of several Notch pathway activator mentioned genes/proteins in GBS and GSI scenarios (Figure 3A). 

(GSI) by comparing the number of upstream activators of the above inhibitor responded cell populations. Same result is also found from GAMMA SECRETASE inhibited cell populations as compared to the 

hes1, maml, dll3, jag2, etc., were found active in the non-responder 17 genes including the Notch pathway genes such as notch1, notch3, pathway proteins in Glioblastoma cell line [105]. It shows that around experimental findings of DAPT, BMS-708163 and RO4929097 (known enzyme in Glioblastoma cell line, the number of upstream activator and inhibitor molecules were reducing significantly as compared 

to the GBS (Figure 3A and Figure 3C). In order to validate this simulation result with experimental data, we considered the previous treatment experiments so that excess Notch pathway activity could be blocked. Also the feasibility of targeting/inhibiting this membrane associated metallo-protease enzyme makes it as a common choice for all clinical and experimental biologists [9,32].

GAMMA SECRETASE inhibited Glioblastoma cell model (GSI) was created and simulated by considering the logical state of this protein as "0" or "OFF" in our Glioblastoma cell model scenario (GBS). We observed that after inhibiting the GAMMA SECRETASE enzyme in Glioblastoma cell line, the number of upstream activator and inhibitor molecules were reducing significantly as compared to the GBS (Figure 3A and Figure 3C). In order to validate this simulation result with experimental data, we considered the previous experimental findings of DAPT, BMS-708163 and RO4929097 (known GAMMA SECRETASE inhibitors) treated expressions profile of Notch pathway proteins in Glioblastoma cell line [105]. It shows that around 17 genes including the Notch pathway genes such as notch1, notch3, hes1, maml, dll3, jag2, etc., were found active in the non-responder GAMMA SECRETASE inhibited cell populations as compared to the inhibitor responded cell populations. Same result is also found from our in-silico simulation of GAMMA SECRETASE Inhibitor scenario (GSI) by comparing the number of upstream activators of the above mentioned genes/proteins in GBS and GSI scenarios (Figure 3A). Simultaneously we also observed from the simulation result that the downstream activated proteins of several Notch pathway activators (e.g., JAG1/2, DLL1/3/4, MAGP1, NICD1 etc.) were getting reduced by administering the GAMMA SECRETASE inhibition in GBS cell line (Figure 3B). These results clearly validate our logical model for GBS and GSI scenarios with the experimental results obtained from previously done cancer stem cell and molecular biology experiments [105].

Comparison between Normal and Glioblastoma scenario: In order to observe the change of the expressions of different Notch pathway proteins in Normal (NNS) and Glioblastoma scenarios (GBS), we simulated our model for both the scenarios and compared to extract and identify the pathway molecules that were showing significant fluctuations in both scenarios. The logical states of the input proteins were considered as same as shown in Figure 2 and the expression levels are provided in Table S4 and Table S5 of Supplementary file. By calculating the dependency matrices for both the scenarios, we were able to identify significant variations of Upstream Activators, Upstream Inhibitors, Downstream Activated and Inhibited proteins for the proteins reported in the X-axis of Figure 3.

Here, the simulation result of Normal Notch pathway scenario (NNS) served as a control to measure the change in the expression level of Notch pathway molecules in Glioblastoma scenario. In this analysis we correlated the expression level of a protein with its total number of Upstream Activators and Inhibitors molecules. We found that different types of proteins from different sub-cellular locations were showing significant changes (Figure 3A and Figure 3C) in Glioblastoma scenario (GBS) compared to the Normal Notch pathway scenario (NNS). Considering the fact that the expression of a protein in a signaling network is directly proportional to its number of upstream activator molecules and inversely proportional to its upstream inhibitor molecules [39], we were able to extract the proteins which were causing Glioblastoma by mutating the Notch signal and its associated molecules. Mostly all the Notch target proteins (Onco-proteins of Glioblastoma such as MAG, BCL2, MYOD etc.,) of GBS had higher number of upstream activators as compared to the NNS (Figure 3C). On the other hand the total number of inhibitor molecules (Upstream inhibitor) acting on the Notch target proteins of NNS scenario was comparably higher than the GBS scenario. Higher the inhibition by upstream inhibitors on Notch target proteins, lower their expressions and thus controls several phenotypic expressions such as cell proliferation, myelination, anti-apoptosis etc.,. This meticulous regulation gets perturbed by several oncogenic factors in Glioblastoma scenario and the inhibitions on the output proteins are withdrawn [22-24,63]. The number of downstream activated proteins of Notch pathway (GAMMA_SECRETASE, WDR12, NICD1, NICD4, EP300, MAML, SKIP, HAT etc.,) was also showing the variations as compared to the NNS (Figure 3B). It was also revealed that the downstream activator molecules of HES1, HES5 and HIF1A were increased from 0 to around 50 in GBS scenario. Surely this finding gave us a preliminary direction to identify the probable drug targets of Glioblastoma treatment. It is also worth mentioning that all these above mentioned molecules are the transcription activators or co-activators of Notch pathway and their up-regulation may be the main reason to increase the activations/productions of Notch target proteins in GBS. However, the number of downstream inhibited molecules of most of the molecules were not showing significant variations in all the five scenarios (Figure 3D), except for NUC_NCD1, FBW7, CDK8, COR and HEY1. Moreover, there was no significant variation in the number of downstream inhibited molecules for GBS, NNS and GSI scenarios.

In-silico identification of alternative drug treated scenarios: In order to find out the alternative drug targets, in place of GAMMA SECRETASE inhibitors/drugs to treat Glioblastoma tumor cell line, we examined several sole and combinatorial targets (mostly proteins) of our Notch pathway model. Identification of such combinations was not easier, especially when the numbers of probable options were very high in such a large signaling network. Selection of the target molecules from a signaling network depends on its topological structure (i.e., connectivity, centrality) as well its logical relationship with all other molecules in the network. From our graph theoretical analysis we were able to identify few important proteins like ADAM/TACE, CML, NICD1, MAML, HIF1A, NRARP, HES1, HES5 etc., which had high centrality values within the network (Table 1). Furthermore, on the basis of the biological feasibility and the evidence of being used as targets in previous experiments, we were able to filter out the proteins like ADAM/TACE, NICD1, MAML, HIF1A, DLL4, as probable drug targets for our analysis. It should be noted that almost all these identified proteins also show the variations of high expression level in GBS Scenario (Figure 3C). Thus by integrating the results from network, logical and reported experimental results, our in-silico modeling approach was able to extract the possible combinations of few proteins to treat Glioblastoma scenario. This finding was useful to bring down the number of probable drug targets from more than 100 proteins of the network to 4 or 5 proteins.
Figure 3: Comparison between normal, glioblastoma, gamma secretase inhibition and two proposed drug target scenarios. NNS: Normal Notch Pathway; GBS: Glioblastoma Scenario; GSI: Gamma Secretase Inhibition; TS1: NICD1 and HIF1A combinatorial inhibition; TS2: NICD1 and MAML combinatorial inhibition. (A) Represents number of upstream activator molecules (Y-axis) activating the molecules (X-axis) representing significant variations (B) Represents number of downstream activated molecules (Y-axis) activated by the molecules (X-axis) representing significant variations (C) Represents number of upstream inhibitor molecules (Y-axis) inhibiting the molecules (X-axis) representing significant variations (D) Represents number of downstream inhibited molecules (Y-axis) inhibited by the molecules (X-axis) representing significant variations.
Using these identified lead targets, we further targeted each protein as a sole target, but unfortunately none of the target gave significant result (result not shown). Eventually, we had to look for the combinations of the proteins from all the lead target molecules. We observed that the synergistic effect of targeting the multiple proteins in combinations was giving good result as compared to the results we were getting from the individual in-silico targeting. We tried several possible combinations while doing this in-silico drug treated perturbation analysis and the best results found from that analysis was presented in TS1 and TS2 scenarios of Figure 3. TS1 represents NICD1 and HIF1A combinatorial drug treated scenario whereas TS2 represents the NICD1 and MAML treated scenario. In TS1 scenario, we were able to suppress partially but comparably lower expressions of Notch target onco-proteins (BCL2, HES1, MAG, IAP etc…) as compared to Glioblastoma as well as GAMMA SECRETASE Inhibitor scenario (GSI). On the other hand, in TS2 scenario, the expressions of the target onco-proteins were completely suppressed (Figure 3). Thus, both the partial inhibition and complete suppression can be achieved by using TS1 and TS2 scenarios, respectively.

**Discussion**

The involvement of the Notch signaling pathway in cancer stem cell generations have been reported in many molecular and cancer stem cell experiments [1-3]. Inhibition of this pathway by several drug targets in various tumor cell lines including Glioblastoma, Oligodendrocyte growth is also been tested [32,35]. Synthesis of several small molecule drugs or inhibitors of this pathway prove the significance and importance of this pathway to the pharmaceutical experiments of cancer pathology. Unfortunately despite several attempts to suppress the cancer progression by inhibiting the molecules of this pathway, the fruitful results are yet to come. Recently developed several GAMMA SECRETASE inhibitors were also found to have several toxic effects [33,34]. Hence a better understanding of the complete map of Notch signaling is useful before developing any drug targets against cancer, as it is a highly regulated as well as sensitive signaling pathway connected with several other signaling and metabolic pathways in cell. Unfortunately, there is no complete map available which can give information about its regulations with other cross talking molecules. Therefore in our work, we first reconstructed the entire Notch signaling map by curating the data from all the available resources especially from literatures and signaling database. In this analysis, we reconstructed a new notch signaling network with 115 molecules (including protein complex, metabolites etc…) and 231 reactions (e.g., Phosphorylation, physical interactions, proteolytic cleavage etc…). To the best of our knowledge this is largest human cell specific Notch signaling map. This pathway and its associated molecules can also be used as biomarkers for the detection of cancer and the identification of drug targets for further in-vitro and in-vivo analysis.

Using this reconstructed network, we further analyzed structurally the network topology by graph theoretical model and were able to identify the centrally located proteins forming the “Hub” in the network. We found that the nuclear transcription factor CSL, STAT3 had the higher eigenvector centrality value in the network along with HES1 and HESS. Similarly the Eigenvector centrality of NICD1/2/3/4 was also increased due to its connections with the output proteins NRARP, implying that these proteins not only have high number of connections but also are connected with other highly prestigious nodes that possess higher number of connections in the network. We were also able to identify the connection of several cross talking molecules, such as, HIF1A from Hypoxia, PTEN from PI3K/AKT pathway, P53, RAS etc., with the core molecules of Notch pathway and creation of either positive or negative feedback loops in the network. Moreover, CSL, NRARP, HIF1A, STAT3 were also showing high Closeness centrality, where the feedback connections of these proteins with the other important proteins such as NICD1/2/3/4 or HES1/5 in the network gave the access of these proteins to regulate more number of other proteins in the network, resulting in an increased closeness centrality values of these proteins. This also signifies that certain perturbations or mutations of these proteins will cause worst effect than the other proteins having lower closeness centrality value. The higher Betweenness centrality value found for NICD1 compare to its other homologues (i.e., NICD2/3/4) shows that unlike the other three homologues of these proteins, NICD1 had extra three upstream regulators proteins: RAS, JIP1 and WDR12 as well as P53 protein in downstream. It is also connected with its nuclear counterpart NUC_NICD1, which has additional downstream target genes (e.g., BCL2, FLIP, IAP, P21, P65, P50, C_REL, REL_B) for transcription than its counterparts NUC_NICD2/3/4, implying enhancement of Betweenness centrality value as more numbers of shortest paths intersect this protein. This structural and topological analysis helps us to identify the probable drug targets to suppress the activity of maximum Notch target proteins.

For facilitating clinical and experimental pharmacologists to perform further in-vivo and in-vitro experiments in real cancer cell model and to test the effect of mutation or deregulation of important proteins in the network under certain circumstances as well as to identify the new biomarkers of Notch pathway, we followed a semi-dynamic computational approach, logical analysis. In the logical analysis we modeled the entire pathway reactions by logical equations and created five scenarios: Normal (NNS), Glioblastoma (GBS), GAMMA SECRETASE inhibition (GSI), Treatment scenarios by inhibiting NICD1 and HIF1A (TS1) and by inhibiting NICD1 and MAML (TS2). We also validated our model with the mRNA expression levels measured in human Glioblastoma cell line [30]. Further we were also able to verify the expression or activation of GSI with the reported experimental findings [105]. From our in-silico simulation of GSI by comparing the number of upstream activators genes/proteins in GBS and GSI scenarios, we observed that the downstream activated proteins of several Notch pathway activator proteins (e.g., JAG1/2, DLL1/3/4, MAGP1, NICD1 etc…) were getting reduced by administering the GAMMA SECRETASE inhibition in GBS cell line, which not only prove the reported experimental findings [105] but also validate our computational study.

Moreover, by comparing the NNS and GBS, we were also able to identify the proteins which were abnormally getting activated or inhibited in Glioblastoma cell line compared to the normal scenario. Along with our network analysis, these findings gave us the opportunity to filter out the possible drug target molecules from out of 115 molecules of the pathway. We identified several probable targets through sole or combinations of proteins by perturbing the logical states of GBS model. Though the sole perturbation did not show promising results, but targeting these proteins in combinations showed promising result in suppressing the expressions of several Notch target proteins. The synergistic effect imposed by using the combinatorial drug targets has improved the result in several folds. Among these possible combinations we observed NICD1 and HIF1A (TS1) are suitable for the partial blocking of Notch pathway activity whereas inhibition of NICD1 and MAML (TS2) is useful to completely suppress the pathway activity. It can also be concluded that depending...
on the critical situations of Glioblastoma, one can use any of the above combinations to suppress the growth of Glioblastoma.

Through our reconstruction and computational study of the human cell specific Notch signaling pathway we get an insight and complete understanding of the interactions between the signaling proteins in the pathway along with identification of alternative drug targets for Glioblastoma, where the pathway is known to become mutated. The computational strategy and alternative drug targets predicted from our analysis may help to achieve more accurate therapeutic strategy, not only for Glioblastoma but also for other diseased conditions. From our analysis the perturbation effects of minimal combination of proteins and identification of new combinatorial drug targets or pathway signatures provide a more sensible strategy for finding therapeutic targets for cancer. Our findings will be useful for the experimental and clinical pharmacologists to select the biomarkers for cancer prognosis and will show new direction in human cell specific signaling pathway study.

Acknowledgement
We thank Mr. Shomeek Chowdhury, Summer Research Fellow, Indian Academy of Sciences, for collating few data for pathway reconstruction. We also thank The Director, CSIR-National Chemical Laboratory, Pune for providing the required infrastructures related to this work under CSIR-NCL grant MLP026226. The work is funded by the Department of Biotechnology, Govt. of India, Grant No. BT/PR13689/BID/07/363/2010.

References


