REDESIGN: RDF-based Differential Signaling Framework for Precision Medicine Analytics

Zainab Al-Taie, MSc[1,2], Nattapon Thanintorn, MSc[2], Ilker Ersoy, PhD[2,1], Olha Kholod, MSc[1,2], Kristen Taylor, PhD[2], Richard Hammer, MD[2,1], Dmitriy Shin, PhD*[2,1]

1 - MU Informatics Institute, University of Missouri, Columbia, MO
2 - Department of Pathology and Anatomical Sciences, University of Missouri, Columbia, MO

Abstract
Pathway-based analysis holds promise to be instrumental in precision and personalized medicine analytics. However, the majority of pathway-based analysis methods utilize "fixed" or "rigid" data sets that limit their ability to account for complex biological inter-dependencies. Here, we present REDESIGN: RDF-based Differential Signaling Pathway informatics framework. The distinctive feature of the REDESIGN is that it is designed to run on "flexible" ontology-enabled data sets of curated signal transduction pathway maps to uncover high explanatory differential pathway mechanisms on gene-to-gene level. The experiments on two morphoproteomic cases demonstrated REDESIGN's capability to generate actionable hypotheses in precision/personalized medicine analytics.

Introduction

Precision medicine, set by President Obama as a new strategic initiative in medicine (1), has been gaining a lot of attention in biomedicine. However, it becomes increasingly evident that one-drug-one-target-one-disease paradigm is unable to address outstanding challenges in this new medical paradigm (2). High-throughput omics technologies, data integration and computational analytics among others have been considered to be crucial to move forward precision medicine ideas (3). Particularly, pathway-based methods (4,5) especially those that utilize curated knowledge bases (KB) such as KEGG (6) and Reactome (7) improve handling of the curse of dimensionality and noise in omics data by bringing analysis to the level of biological functions (8-11). Specifically, direct correlation of differential gene expression with a clinical phenotype usually results in a situation when gene sets, which are strongly correlated with the phenotype and reported by different studies, have little or no overlap. It happens because gene expression profiling studies are generally underpowered and each study group has different patient cohorts in which the same biological mechanisms may have been manifested by different genes. Despite efforts to address this limitation by integrating multiple datasets (12,13) the issue of the curse of dimensionality still persists because it is not feasible at the moment to compile a single data set with an order of magnitude of the genome. Early methods of Over Representation Analysis (ORA) (4,14) partially address this problem by considering affected biological pathways that are overrepresented by differentially expressed (DE) genes. However, due to the nature of statistical approaches used in ORA methods, levels of expression of DE genes is ignored. The importance of specific pathways is simply assessed by counting of how many DE genes are present in these pathways (enrichment). However, an enrichment score per se is not explanatory enough to evaluate the contribution of these genes to a particular biological state (14). Therefore, slight differential expression of genes that truly affect the phenotype in question might be discarded as noise. Functional Class Scoring (FCS) methods (4,14), such as Gene Set Enrichment Analysis (GSEA) (15), not only consider genes with large DE values but also account for sets of genes with possibly weak DE values but working in concert. For instance, GSEA defines a set of genes, e.g. representing a biological pathway, and assigns a large enrichment score for that pathway if the pathway's genes are found close to the top or bottom of the ranked DE gene list. The downside of GSEA and many other FSC methods is that they do not consider the role and position of genes in the pathways. For instance, genes located upstream (e.g. "driver genes") could have greater effect on the pathway that those found downstream (e.g. "passenger genes"). Pathway Topology (PT) approaches (4,14) take into consideration the location of genes in pathways. For example, Signaling Pathway Impact Analysis (SPIA) (16), computes a perturbation impact of each gene in accordance to its position in a pathway. PARADIGM pathway method (17) further improves pathway-based analytics by integrating multiple datasets derived from Comparative Genome Hybridization (CGH), and Single Nucleotide Polymorphism (SNP) technologies in addition to transcriptomics data. Different variations of ORA, FCS, TP and other pathway-based methods have been used to infer patient-specific pathway deregulations (18,19). What common in many of these methods, though, is the fact that they treat a biological pathway as a whole, assigning a metric of importance (enrichment) to the whole pathway

using various statistical approaches. However, in many applications, including precision/personalized medicine, it is critical to infer sub-pathway or gene-to-gene level mechanisms. Another important limitation, in our view, is that these methods are designed to run over "rigid" or "static" pathway networks' data. These datasets may not take into account inter-relationships among data entities, such as for instance the fact that distinctive network nodes may *de-facto* represent the same biological entities (e.g., alternative names or IDs), have functional similarity (e.g., the same binding domains), or belong to the same protein complex (i.e., part-whole relationship). Because of this "rigidity" many queries may not return desirable results, even though the information might be present in the network datasets in an implicit form.

Here, we introduce a high explanatory RDF-based differential signaling (REDESIGN) informatics framework that enables gene-to-gene level computational analysis of topological differences of signaling cascades. These differences can pinpoint places in signal transduction pathways and regulatory networks where a specific patient's case deviates from the canonical flow of the disease. This will not only help to design personalized combinatorial therapeutic regimens, but also to bring new hypotheses to basic research. The distinctive feature of this method is the fact that the analysis is performed over "flexible", RDF-formatted and ontology-enabled pathway network data. Such data, equipped with description logic inference, provide means to take into account "biological isomorphism", which makes differential pathway analysis more biologically relevant. As we will demonstrate further in text, the method has high potential to help realize precision medicine ideas in cancer theranostics.

Methods

REDESIGN framework utilizes an RDF-based "mashup" knowledge base of signal transduction pathways derived from Kyoto Encyclopedia of Genes and Genomes (KEGG), including pathways associated with signal transduction, cellular processes, organismal systems, human diseases, and drug development (6). We have previously demonstrated the utility of such KB to reduce complexity of biological knowledge for precision medicine analytics (20). In this work, we extended the KB to include more biologically relevant interaction types (see Table 1).

Table 1. Modeled biological relationships.

RDF Predicate	Modeled Purpose	RDF Predicate	Modeled Purpose
activates	Molecular Interaction	misses_interaction_methylation_inhibition	Molecular Interaction
inhibits	Molecular Interaction	misses_interaction_methylation_activation	Molecular Interaction
binds_associates	Molecular Interaction	misses_interaction_glycosylation_inhibition	Molecular Interaction
dissociates	Molecular Interaction	misses_interaction_glycosylation_activation	Molecular Interaction
changes_state	Molecular Interaction	misses_interaction_deubiquitination_inhibition	Molecular Interaction
expresses	Molecular Interaction	misses_interaction_deubiquitination_activation	Molecular Interaction
represses	Molecular Interaction	misses_interaction_ubiquitination_inhibition	Molecular Interaction
indirectly_affects	Molecular Interaction	misses_interaction_ubiquitination_activation	Molecular Interaction
indirectly_affects_activates	Molecular Interaction	misses_interaction_dephosphorylation_inhibition	Molecular Interaction
indirectly_affects_inhibits	Molecular Interaction	misses_interaction_phosphorylation_activation	Molecular Interaction
phosphorylates_activates	Molecular Interaction	misses_interaction_phosphorylation	Molecular Interaction
phosphorylates_inhibits	Molecular Interaction	misses_interaction_methylation	Molecular Interaction
dephosphorylates_activates	Molecular Interaction	misses_interaction_inhibition_degradation	Molecular Interaction
dephosphorylates_inhibits	Molecular Interaction	misses_interaction_inhibition	Molecular Interaction
ubiquitinates_activates	Molecular Interaction	misses_interaction_indirect_effect	Molecular Interaction
ubiquitinates_inhibits	Molecular Interaction	misses_interaction_glycosylation	Molecular Interaction
deubiquitinates_activates	Molecular Interaction	misses_interaction_expression	Molecular Interaction
deubiquitinates_inhibits	Molecular Interaction	misses_interaction_phosphorylation_binding_association	Molecular Interaction
methylates_activates	Molecular Interaction	misses_interaction_dissociation_inhibition	Molecular Interaction
methylates_inhibits	Molecular Interaction	misses_interaction_dissociation_degradation	Molecular Interaction
glycosylates_activates	Molecular Interaction	misses_interaction_dissociation_activation	Molecular Interaction
glycosylates_inhibits	Molecular Interaction	misses_interaction_dissociation	Molecular Interaction
misses_interaction	Molecular Interaction	misses_interaction_dephosphorylation	Molecular Interaction
phosphorylates	Molecular Interaction	misses_interaction_state_change	Molecular Interaction
dephosphorylates	Molecular Interaction	misses_interaction_binding_association	Molecular Interaction
ubiquitinates	Molecular Interaction	misses_interaction_activation	Molecular Interaction
deubiquitinates	Molecular Interaction	misses_interaction_phosphorylation_inhibition	Molecular Interaction
glycosylates	Molecular Interaction	changes_conformation	Molecular Interaction
methylates	Molecular Interaction	contains	Biological Isomorphism
misses_interaction_deacetylation	Molecular Interaction	crosstalks_with	Pathway interaction
misses_interaction_acetylation	Molecular Interaction	participates_in	Biological process
misses_interaction_deubiquitination	Molecular Interaction	involved_in	Disease association
misses_interaction_ubiquitination	Molecular Interaction	is_part-Of	Pathway association
misses_interaction_repression	Molecular Interaction	sameAs	Biological isomorphism
misses_interaction_indirect_effect_inhibition	Molecular Interaction		
misses_interaction_indirect_effect_activation	Molecular Interaction		

The overall roadmap of the REDESIGN method, which is a two-step process, is depicted in Figure 1.

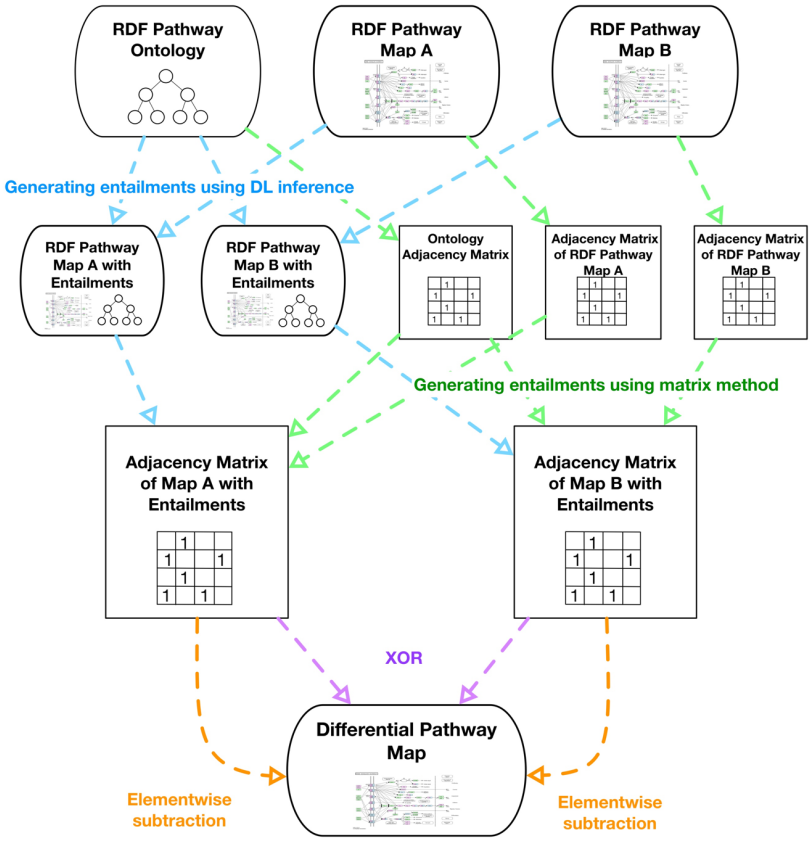

Figure 1. REDESIGN roadmap. Generating entailments using description logic (DL) inference marked in Blue. Generating entailments using the Matrix method marked in Green. Computing the differential pathway map using element-wise XOR marked in Purple and using element-wise subtraction marked in Orange.

REDESIGN Step 1: Given RDF network representations of pathway maps A and B, corresponding to two biological states, as well as an RDF ontology, consisting of biological relationships among pathway entities in maps A and B, we first extend information in maps A and B with inferred knowledge. For instance, given distinct pathway nodes ***A***, ***B***, ***C***, and ***D***, signaling interactions ***A-activates-B***, ***C-inhibits-D***, and an ontological relation ***A-aliasOf-C***, an additional signal transduction interaction ***A-inhibits-D*** will be added to the appropriate pathway map. We consider several types of such entailments, for instance, generated by RDF predicates "sameAs" (gene alias) and "contains" (ontological relation whole-part, e.g. protein complex/protein relationship). There are two ways in which REDESIGN performs the entailment process. The entailments can be generated by Description Logic (DL) inference using means of an underlying RDF store such as RDF/DL reasoners (21) or using matrix representation of RDF pathway maps. In the matrix method, RDF data of pathway and ontology maps are first represented by their adjacency matrices. In such an adjacency matrix, a cell holds a non-zero value only if corresponding RDF nodes have an edge. Here, we experimented with adjacency matrices that account and do not account for edge types. In the former case, a cell value of an adjacency matrix holds a numerical code of a corresponding edge. In the latter case, a cell value is 1 if a corresponding edge exists, and zero otherwise. REDESIGN then extends an adjacency matrix of each pathway map with relations derived from an ontology adjacency matrix using algorithmic steps shown in Figure 2.

REDESIGN Step 2: During the second step REDESIGN uses adjacency matrices of the RDF pathway maps with entailments to compute a differential pathway map. The differential pathway map consists of topological differences of the original pathway maps. The map is computed using either element-wise logical XOR operation or using element-wise subtraction method. In case with XOR, each cell of the differential adjacency matrix will have a non-zero value only if an edge exists only in one of the original maps for which the differential map is computed. In the subtraction method, we utilize an indicator function:

$$AdjDiffMap_{ij} = \begin{cases} 1, \left| AdjMapA_{ij} - AdjMapB_{ij} \right| > 0 \\ 0, otherwise \end{cases},$$

where $AdjDiffMap_{ij}$ is the *ij-th* cell of the adjacency matrix of the resulting differential map, $AdjMapA_{ij}$ and $AdjMapB_{ij}$, are the *ij-th* cells of adjacency matrices of original RDF pathway maps A and B with entailments correspondingly. The subtraction method takes into account types of edges. A cell value of the adjacency matrix of the resulting differential map will have a zero value only if the corresponding cell values in the original adjacency

matrices are equal, reflecting the equal type of edges in the original RDF pathway maps. Finally, the resulting differential RDF pathway map is reconstructed from its adjacency matrix and the types of its edges are obtained from the original RDF pathway maps A and B and visualized.

The current version of REDESIGN was implemented using AllegroGraph RDF Store (21), Java-based Jena (22) and Sesame (23) libraries. The RDF entailments were materialized and then transferred to the Neo4J (24) graph database for graph traversal operations and adjacency matrix computations.

REDESIGN "Toy" Example:

To demonstrate working logic of REDESIGN framework, let us look at an example shown on Panel A in Figure 3, which presents RDF pathway maps ***Map A*** and ***Map B*** corresponding to two biological conditions. The differential pathway map computed by REDESIGN is depicted in panel C, where color of the differential edges indicates which original map they came from.

If we take into account a simple ontological relationship of biological isomorphism of nodes **E** and **G** (panel B in Figure 3), then after generating entailments maps ***Map A*** and ***Map B*** will be augmented with inferential edges (marked in blue). The result of REDESIGN differential pathway analysis will then be different (panel D in Figure 3). It is obvious from this example that just one simple ontological relationship can lead to unexpected and more complex results.

Input:
RDF Pathway Network ***Map A*** = (N_A, E_A),
RDF Pathway Network ***Map B*** = (N_B, E_B),
RDF Pathway Ontology Network ***Map O*** = (N_O, E_O),
where for each network map, ***N*** and ***E*** are sets of nodes and edges correspondingly,
such that $\forall\ n_i, n_i \in N$ = RDF_SUBJECTS ∪ RDF_OBJECTS,
index(n_i) – index of n_i in ***N***,
$\forall\ e_i\ e_i \in E$ = RDF_PREDICATES, ***p_type***(e_i) – type of predicate e_i, and
$\forall\ t_i\ t_i$ = (***Subject***(t_i), ***Predicate***(t_i), ***Object***(t_i)) $\in T$, ***T*** set of all RDF triples in the corresponding RDF network map

Output:
RDF Pathway Network ***Map A with entailments***
RDF Pathway Network ***Map B with entailments***

Start
REM *Initializations of adjacency matrices*
Initialize adjacency matrix ***AdjMapA*** of ***Map A***, such that $\forall$***AdjMapA***$_{ij}$, ***AdjMapA***$_{ij}$ holds the predicate type of the corresponding edge $e \in E_A$

Initialize adjacency matrix ***AdjMapB*** of ***Map B***, such that $\forall$***AdjMapB***$_{ij}$, ***AdjMapB***$_{ij}$ holds the predicate type of the corresponding edge $e \in E_B$

Initialize adjacency matrix ***AdjMapO*** of ***Map O***, such that $\forall$***AdjMapO***$_{ij}$, ***AdjMapO***$_{ij}$ = 1, if ∃ the corresponding edge $e \in$ ***IsomorphicE***$_O$, where ***IsomorphicE***$_O \subset E_O$, is a set of predicates in RDF Pathway Ontology that encode for biological isomorphism

REM *Generating entailments using matrix method*
For each $t_A \in T_A$ do:
For each column *j* in ***AdjMapO*** do:
if ***AdjMapO***(***index***(***Object***(t_A)), *j*) == 1
then ***AdjMapA***(***index***(***Subject***(t_A)), *j*) = ***p_type***(***Predicate***(t_A))
if ***AdjMapO***(***index***(***Subject***(t_A)), *j*) == 1
then ***AdjMapA***(*j*, ***index***(***Object***(t_A))) = ***p_type***(***Predicate***(t_A))

For each $t_B \in T_B$ do:
For each column *j* in ***AdjMapO*** do:
if ***AdjMapO***(***index***(***Object***(t_B)), *j*) == 1
then ***AdjMapB***(***index***(***Subject***(t_B)), *j*) = ***p_type***(***Predicate***(t_B))
if ***AdjMapO***(***index***(***Subject***(t_B)), *j*) == 1
then ***AdjMapB***(*j*, ***index***(***Object***(t_B))) = ***p_type***(***Predicate***(t_B))

End

Figure 2. Matrix method for generating entailments in RDF pathway network maps.

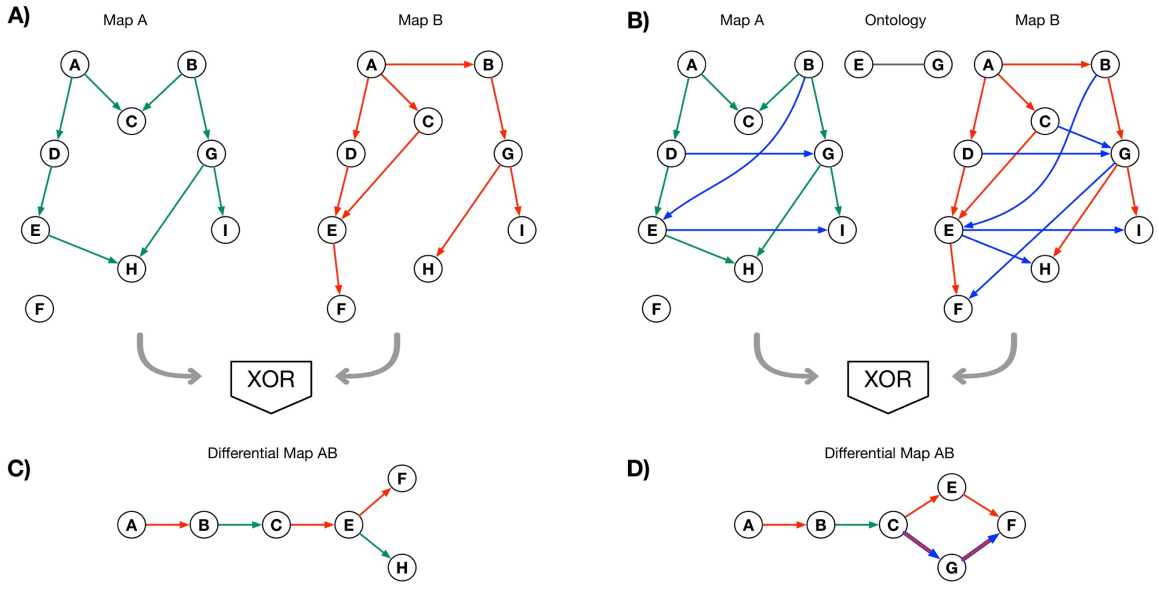

Figure 3. "Toy example" of REDESIGN inferential analysis.

Results and Discussion

To test the utility of REDESIGN framework for precision medicine analytics we have conducted two experiments that are presented below.

Validation of REDESIGN on Primary GBM and Secondary GBM WHO Grade IV

Glioblastoma Muliforme (GBM) is the most aggressive type of primary brain tumors in humans that affects central nervous system (CNS) and has low curable rate. Development of GBM involves multiple genetic alterations and aberrant signaling pathways activation. Among the signaling pathways in GBM, recent studies suggest that NF-kB signaling pathway is one of the main oncogenic pathways in promoting tumor formation and response to therapy (25). Like other types of cancers, *NF-kB* is constitutively expressed at high level in malignant glioma (GBM IV), leading to multiple aspects of aberrant activities involved in anti-apoptosis, cell proliferation, angiogenesis, disease recurrence, and resistance to therapy. In order to validate our proof of concept, we tested our differential pathway analysis algorithm on the curated signaling pathways of GBM to see if the algorithm reveals biologically useful information as the morphoproteomic diagrams reported by the pathologists. Morphoproteomics approach has been used by pathologists to elucidate constitutive activation of *NF-kB* pathway and its signaling cascades in malignancy and therapy of patients with glioblastoma multiforme (26).

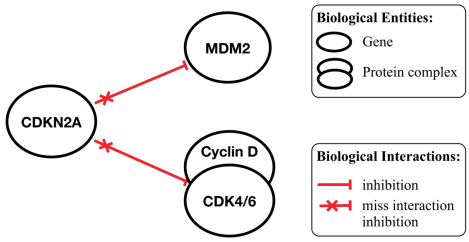

Figure 4. The network depicts no inhibition of MDM2 and CyclinD-CDK4/6 complexes by CDKN2A.

First, we ran our experiment on the signaling pathway maps of primary GBM and secondary GBM WHO Grade IV. The goal of this task was to identify differential pathways in both signaling maps, which could potentially be main drivers in development of advanced stages of glioma. Since activation of *NF-kB* and loss of *PTEN* (involved in virtually all types of cancers) are present in both signaling pathways of GBM, the algorithm does not include the two genes in the networks. The morphoproteomic findings suggest several GBM oncogenic pathways in which *EGFR* amplification and *MDM2* were involved. Figure 4 shows that *CDKN2A* (alias of *INK4A*) gene does not inhibit the activity of *MDM2* resulting in MDM2 up-regulation. Consequently, aberrant regulation of *MDM2* leads to NF-kB overexpression at the transcriptional and protein level. NF-kB overexpression due to the loss of *PTEN* and *MDM2* amplifications results in chemo-resistance, disruption of programmed cell death, and angiogenesis (27). In addition, our algorithm allowed us to capture the aberrant interactions between *CDKN2A* (alias of *INK4A*) and Cyclin D-CDK4/6 complexes shown in Figure 4. As a result, these complexes could not initiate the phosphorylation of the tumor suppressor protein RB resulting in the dissociation of E2F from RB-E2F complexes. Without the dissociation of RB-E2F complexes, loss of phosphorylated RB leads to genomic instability and lack of E2F expression results in anti-apoptosis and uncontrolled cell proliferation (28,29). This demonstrates that our differential pathway analysis framework confirms the previous morphoproteomic findings showing signal cascades of constitutively activated *NF-kB* (26,30,31). Moreover, these results depict potential genes candidates contributing to development of malignant glioma, which are *MDM2*, *Cyclin D*, and *CDK4/6*. Thus, our approach opens a new avenue for further investigation of aberrant gene functions and gene-to-gene interactions in experimental and computational settings in precision medicine.

Uncovering alternative therapeutic targets for GBM patients with tyrosine kinase inhibitor resistance

Next, we tested REDESIGN framework on the signaling pathways of secondary GBM WHO Grade IV and GBM with tyrosine kinase inhibitor (TKI) resistance in order to see if the algorithm enables us to uncover alternative therapeutic targets for the GMB patients with TKI resistance. The *EGFR* gene is often amplified and mutated in virtually all types of cancers, including malignant glioma. EGFR overexpression is reported to promote DNA synthesis via tyrosine kinase involvement, resulting in progression of malignant brain tumors. Thus, EGFR is one of the therapeutic targets for GBM patients. Recent studies report that Gefitinib, one of EGFR tyrosine kinase inhibitors, is often administered to the patients with glioma in order to block signal transduction pathways implicated in the activity of cell proliferation and tumor growth activities within tumor cells (32-34). However, the study reported by Heimberger et. al demonstrated that Gitfinib inhibited growth of tumors that were highly expressing wild-type EGFR, but showed inability to inhibit tumor cells expressing the EGFR variant III (EGFRvIII), approximately a 70% of reduction in therapeutic efficacy tested *in vivo* (35-38). The morphoproteomic findings of GMB WHO Grade IV demonstrated expression of EGFR/EGFRvIII on the cell surface of the malignant glial cells and expression of PKC-alpha as well. In addition, activation of mTOR/Akt pathway was evidenced by nuclear compartmentalization of p-mTOR and p-Akt. The findings also reported the overexpression of vascular endothelial

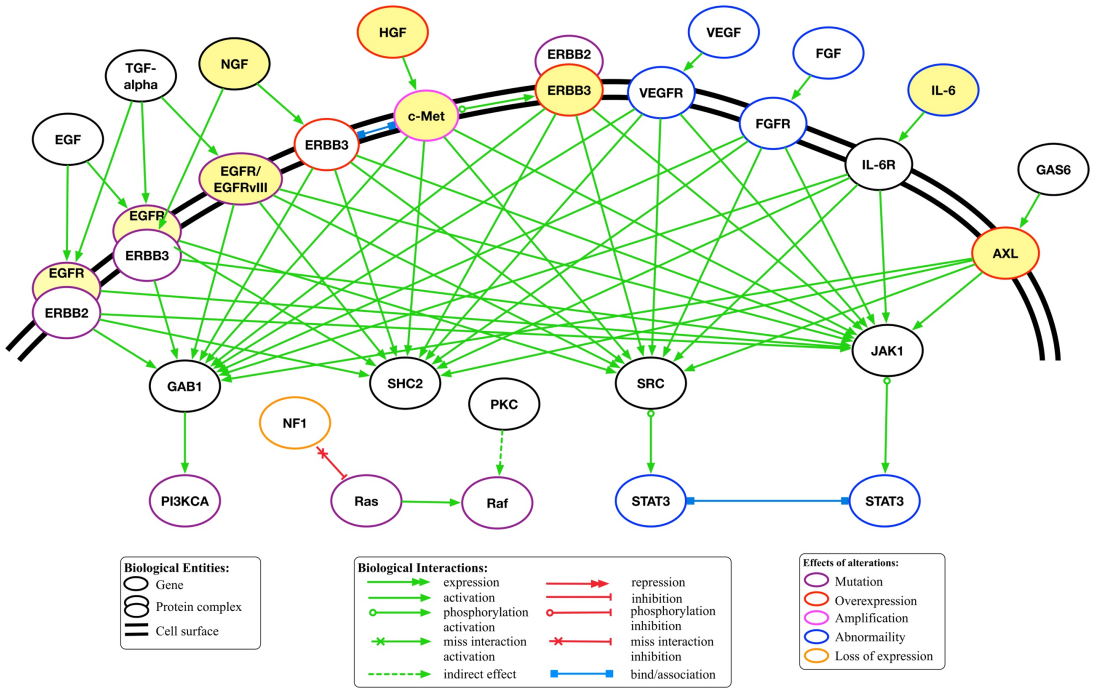

Figure 5. The network of receptor tyrosine kinases resulting in tyrosine kinase inhibitor resistance.

growth factor (VEGF) and amplification of *AKT3* gene (39). Based on these literature findings, our results show differential signaling cascades between the pathway maps of secondary GBM WHO Grade IV and GBM with TKI resistance. Figure 5 depicts a network of receptor tyrosine kinases (RTKs), including EGFR/EGFRvIII, ErbB, NGF, HGF, IL-6, AXL, c-Met, which are often amplified or overexpressed in glioma patients resulting in tyrosine kinase inhibitor resistance. Previous studies also revealed that activation of alternative pathways via HGF, AXL, and c-Met and aberrance of the downstream pathways via K-Ras mutation, PTEN loss, BCL2-like, and BIM-deletion is reported as mechanisms leading to resistance to EGFR-TKI therapy (34,40,41). Moreover, the resulted networks coincide with morphoproteomic diagrams reported in (42) for *PKC-alpha*, *mTOR*, *AKT3*. As it can be seen from Figure 6, the Akt3 network shows the signal transduction leading to aberrant activity of cell proliferation, survival, motility, and angiogenesis within tumor cells. The *BAD*, *BIM*, *BAX*, *BCL-2,* and *BCL-XL* genes are reported to play an important role in apoptosis (43,44).

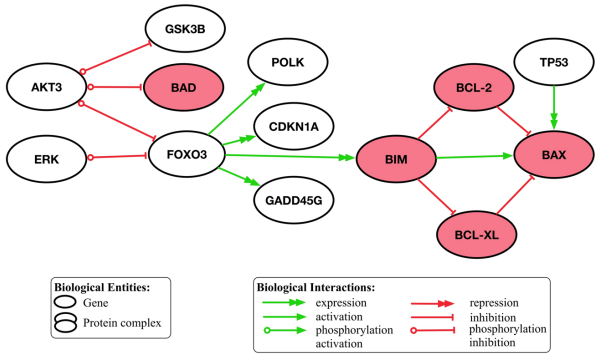

Figure 6. The Akt3 network and downstream pathways resulting in uncontrolled proliferation, increased survival, and anti-apoptosis.

Figure 7 shows the network of *mTOR* signal transduction that cascade to p70S6K, S6, eIF-4EGBP, and eIF-4E. Brown et.al demonstrated the morphoproteomic findings that correlative expression of phosphorylated mTOR and phosphorylated p70S6K was observed in invasive head and neck tumors (45).

Our results also confirm the morph oproteomic studies revealing how the pathologists treated the EGFR-inhibitor-resistance GBM patient with an adjuvant program using a combination of anti-tumor agents shown in Table 2 (42).

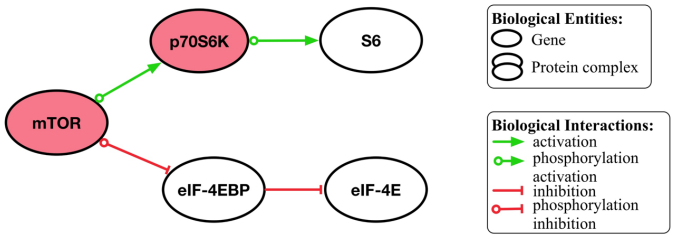

Figure 7. The network of mTOR resulting in metastasis in head and neck tumors.

Table 2. Potential anti-neoplastic agents and therapeutic targets for adjuvant treatment.

Potential Anti-Neoplastic Agents	Potential Therapeutic Targets
Lapatinib	EGFR/EGFRvIII
Temozolomide	EGFR/EGFRvIII, p-Akt, p53
Metformin	EGFR/EGFRvIII, mir-201, mir-26, mir-26a, SIRT1, RICTOR
Niacinamide	p53, SIRT1
Plerixafor	CXCR4, CXCL12/SDF-1

Figure 8 provides the same evidence for alternative therapeutic targets as the morphoproteomic findings reported by the group of pathologists. Therefore, our approach is deemed to support in uncovering alternative therapeutic targets that interrupt the signaling pathways of TKI resistance into the therapy of GBM patients.

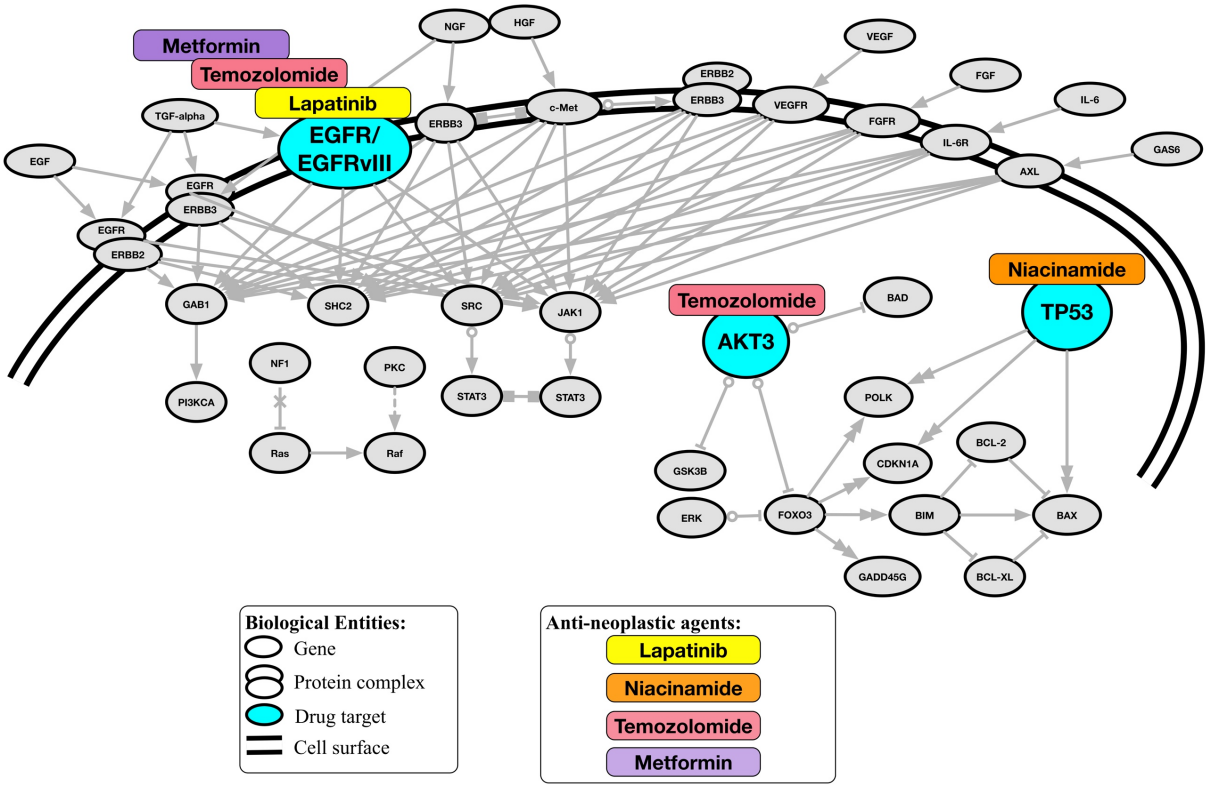

Figure 8. The network depicts alternative therapeutic targets and downstream pathways.

Conclusion

Here, we presented REDESIGN: RDF-based Differential Signaling Pathway informatics framework. The distinctive feature of the REDESIGN is that it is designed to run on "flexible" ontology-enabled data sets of curated signal transduction pathway maps to uncover high explanatory differential pathway mechanisms on gene-to-gene level. Preliminary validation of REDESIGN using retrospective studies with two morphoproteomic cases demonstrated REDESIGN's utility to generate actionable hypotheses in precision/personalized medicine analytics. However, more validation is needed to assess predictive power of REDESIGN. In future work, we plan to extend REDESIGN to include a variety of ontological relationships reflecting biological isomorphism, such as functional domain similarity and epistatic mechanisms.

References

1. FACT SHEET: President Obama's Precision Medicine Initiative. Available from: http://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative
2. Xie L, Ge X, Tan H, Xie L, Zhang Y, Hart T, et al. Towards structural systems pharmacology to study complex diseases and personalized medicine. Nussinov R, editor. PLoS Comp Biol. Public Library of Science; 2014 May;10(5):e1003554.
3. Lussier YA, Li H, Pouladi N, Li Q. Accelerating precision biology and medicine with computational biology and bioinformatics. Genome Biol. BioMed Central Ltd; 2014 Sep 5;15(9):450.
4. Jin L, Zuo X-Y, Su W-Y, Zhao X-L, Yuan M-Q, Han L-Z, et al. Pathway-based analysis tools for complex diseases: a review. Genomics Proteomics Bioinformatics [Internet]. 2014 Oct;12(5):210–20. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=25462153&retmode=ref&cmd=prlinks
5. la Fuente de A. From 'differential expression'to 'differential networking'–identification of dysfunctional regulatory networks in diseases. Trends in genetics. 2010;26(7):326–33.
6. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000;28(1):27–30.
7. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, et al. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res. 2011;39(Database issue):D691–7.
8. Shin D, Arthur G, Popescu M, Korkin D, Shyu C-R. Uncovering influence links in molecular knowledge networks to streamline personalized medicine. J Biomed Inform. 2014 Dec;52:394–405.
9. Ji J, Yuan Z, Zhang X, Li F, Xu J, Liu Y, et al. Detection for pathway effect contributing to disease in systems epidemiology with a case–control design. BMJ open. 2015;5(1):e006721.
10. Ogata H, Goto S, Fujibuchi W, Kanehisa M. Computation with the KEGG pathway database. BioSystems. 1998;47(1):119–28.
11. Lee S, Park Y, Kim S. MIDAS: MIning Differentially Activated Subpaths of KEGG pathways from multi-class RNA-seq data. Methods. 2017.
12. Turnbull AK, Kitchen RR, Larionov AA, Renshaw L, Dixon JM, Sims AH. Direct integration of intensity-level data from Affymetrix and Illumina microarrays improves statistical power for robust reanalysis. BMC Med Genomics. BioMed Central; 2012;5(1):35.
13. Lê Cao K-A, Rohart F, McHugh L, Korn O, Wells CA. YuGene: a simple approach to scale gene expression data derived from different platforms for integrated analyses. Genomics. 2014 Apr;103(4):239–51.
14. Ten years of pathway analysis: current approaches and outstanding challenges. 2012 Feb;8(2):e1002375.
15. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.
16. Tarca AL, Draghici S, Khatri P, Hassan SS, Mittal P, Kim J-S, et al. A novel signaling pathway impact analysis. Bioinformatics. Oxford University Press; 2009 Jan 1;25(1):75–82.
17. Vaske CJ, Benz SC, Sanborn JZ, Earl D, Szeto C, Zhu J, et al. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. Bioinformatics. 2010 Jun 15;26(12):i237–45.
18. Ahn T, Lee E, Huh N, Park T. Personalized identification of altered pathways in cancer using accumulated normal tissue data. Bioinformatics. Oxford University Press; 2014 Sep 1;30(17):i422–9.
19. Drier Y, Sheffer M, Domany E. Pathway-based personalized analysis of cancer. Proc Natl Acad Sci U S A. 2013 Apr 16;110(16):6388–93.

20. Thanintorn N, Wang J, Ersoy I, Al-Taie Z, Jiang Y, Wang D, et al. RDF SKETCH MAPS - KNOWLEDGE COMPLEXITY REDUCTION FOR PRECISION MEDICINE ANALYTICS. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing. 2016;21:417–28.

21. Aasman J. Allegro graph: RDF triple database. 2006.

22. McBride B. Jena: a semantic Web toolkit. IEEE Internet Comput. 2002 Nov;6(6):55–9.

23. Broekstra J, Kampman A, van Harmelen F. Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema. In: Horrocks I, Hendler J, editors. The Semantic Web — ISWC 2002. Berlin, Heidelberg: Springer Berlin Heidelberg; 2002. pp. 54–68. (Lecture Notes in Computer Science; vol. 2342).

24. Vicknair C, Macias M, Zhao Z, Nan X, Chen Y, Wilkins D. A comparison of a graph database and a relational database. New York, New York, USA: ACM Press; 2010. p. 1.

25. Puliyappadamba VT, Hatanpaa KJ, Chakraborty S, Habib AA. The role of NF-κB in the pathogenesis of glioma. Molecular & Cellular Oncology. Taylor & Francis; 2014 Dec 23;1(3):e963478.

26. Brown RE, Law A. Morphoproteomic demonstration of constitutive nuclear factor-kappaB activation in glioblastoma multiforme with genomic correlates and therapeutic implications. Ann Clin Lab Sci. 2006 ed. 2006/11/28; 2006;36(4):421–6.

27. Kim HJ, Hawke N, Baldwin AS. NF-κB and IKK as therapeutic targets in cancer. Cell Death and Differentiation. 2006 Feb;13(5):738–47.

28. Watanabe G, Albanese C, Lee RJ, Reutens A, Vairo G, Henglein B, et al. Inhibition of cyclin D1 kinase activity is associated with E2F-mediated inhibition of cyclin D1 promoter activity through E2F and Sp1. Mol Cell Biol. American Society for Microbiology (ASM); 1998 Jun;18(6):3212–22.

29. Musgrove EA, Caldon CE, Barraclough J, Stone A, Sutherland RL. Cyclin D as a therapeutic target in cancer. Nat Rev Cancer. Nature Publishing Group; 2011 Jul 7;11(8):558–72.

30. Rios A, Hsu SH, Blanco A, Buryanek J, Day AL, McGuire MF, et al. Durable response of glioblastoma to adjuvant therapy consisting of temozolomide and a weekly dose of AMD3100 (plerixafor), a CXCR4 inhibitor, together with lapatinib, metformin and niacinamide. Oncoscience. 2016;3(5-6):156–63.

31. Brown RE. Morphogenomics and morphoproteomics: a role for anatomic pathology in personalized medicine. Arch Pathol Lab Med. 2009 ed. 2009/04/28; 2009 Apr;133(4):568–79.

32. Reardon DA, Quinn JA, Vredenburgh JJ, Gururangan S, Friedman AH, Desjardins A, et al. Phase 1 trial of gefitinib plus sirolimus in adults with recurrent malignant glioma. Clin Cancer Res. American Association for Cancer Research; 2006 Feb 1;12(3 Pt 1):860–8.

33. Rich JN, Reardon DA, Peery T, Dowell JM, Quinn JA, Penne KL, et al. Phase II trial of gefitinib in recurrent glioblastoma. J Clin Oncol. American Society of Clinical Oncology; 2004 Jan 1;22(1):133–42.

34. Hegi ME, Diserens AC, Bady P, Kamoshima Y, Kouwenhoven MCM, Delorenzi M, et al. Pathway Analysis of Glioblastoma Tissue after Preoperative Treatment with the EGFR Tyrosine Kinase Inhibitor Gefitinib--A Phase II Trial. Molecular Cancer Therapeutics. American Association for Cancer Research; 2011 Jun;10(6):1102–12.

35. Learn CA, Hartzell TL, Wikstrand CJ, Archer GE, Rich JN, Friedman AH, et al. Resistance to tyrosine kinase inhibition by mutant epidermal growth factor receptor variant III contributes to the neoplastic phenotype of glioblastoma multiforme. Clin Cancer Res. 2004 May 1;10(9):3216–24.

36. Heimberger AB, Learn CA, Archer GE, McLendon RE, Chewning TA, Tuck FL, et al. Brain tumors in mice are susceptible to blockade of epidermal growth factor receptor (EGFR) with the oral, specific, EGFR-tyrosine kinase inhibitor ZD1839 (iressa). Clin Cancer Res. 2002 Nov;8(11):3496–502.

37. Huang PH, Xu AM, White FM. Oncogenic EGFR signaling networks in glioma. Sci Signal. 2009.

38. Reardon DA, Wen PY. Therapeutic advances in the treatment of glioblastoma: rationale and potential role of targeted agents. The oncologist. 2006.

39. Brown RE, McGuire MF. Oncogenesis recapitulates embryogenesis via the hypoxia pathway:

morphoproteomics and biomedical analytics provide proof of concept and therapeutic options. Ann Clin Lab Sci. 2012;42(3):243–57.

40. Learn CA, Hartzell TL, Wikstrand CJ, Archer GE. Resistance to tyrosine kinase inhibition by mutant epidermal growth factor receptor variant III contributes to the neoplastic phenotype of glioblastoma multiforme. Clinical Cancer …. 2004.

41. Garofalo M, Romano G, Di Leva G, Nuovo G. EGFR and MET receptor tyrosine kinase-altered microRNA expression induces tumorigenesis and gefitinib resistance in lung cancers. Nature. 2012.

42. Rios A, Hsu SH, Blanco A, Buryanek J, Day AL, McGuire MF, et al. Durable response of glioblastoma to adjuvant therapy consisting of temozolomide and a weekly dose of AMD3100 (plerixafor), a CXCR4 inhibitor, together with lapatinib, metformin and niacinamide. Oncoscience. 2016;3(5-6):156–63.

43. Cheng EH, Wei MC, Weiler S, Flavell RA, Mak TW, Lindsten T, et al. BCL-2, BCL-X(L) sequester BH3 domain-only molecules preventing BAX- and BAK-mediated mitochondrial apoptosis. Molecular cell. 2001 Sep;8(3):705–11.

44. Zong WX, Lindsten T, Ross AJ, MacGregor GR, Thompson CB. BH3-only proteins that bind pro-survival Bcl-2 family members fail to induce apoptosis in the absence of Bax and Bak. Genes Dev. Cold Spring Harbor Lab; 2001 Jun 15;15(12):1481–6.

45. Brown RE, Zhang PL, Lun M, Zhu S, Pellitteri PK, Riefkohl W, et al. Morphoproteomic and pharmacoproteomic rationale for mTOR effectors as therapeutic targets in head and neck squamous cell carcinoma. Ann Clin Lab Sci. 2006 ed. 2006/09/05; 2006;36(3):273–82.