

Pancreatic cancer mutationscape: revealing the link between modular restructuring and intervention efficacy amidst common mutations

Daniel Plaughner, David Murrugarra

January 25, 2024

Abstract

Across cancer types, the prognosis for pancreatic cancer (PC) is among the worst and options for treatment are limited. There is increasing evidence that biological systems, including PC, are modular in both structure and function. Complex biological signaling networks such as gene regulatory networks (GRNs) are proving to be composed of subcategories that are interconnected and hierarchically ranked. These networks contain highly dynamic processes that ultimately dictate cellular function over time. In this work, we use an established Boolean multicellular signaling network of PC to show that the variance in topological rankings of the most phenotypically influential modules implies a strong relationship between structure and function. We further show that induction of mutations alters the modular structure, which analogously influences the aggression and controllability of the disease *in silico*. We finally present evidence that the impact and location of mutations with respect to PC modular structure directly corresponds to the efficacy of single agent treatments *in silico*, because topologically deep mutations require deep targets for control.

1 Introduction

Biologically, regulatory networks and protein-protein interaction networks are typically thought to be densely connected sub-regions of an overall sparse system [1]. Natural cellular functions such as signal transmission are carried out by so-called modules that are discrete entities with separable functionality from other modules [2]. For example, the ribosome is a module that is responsible for synthesizing proteins that is spatially isolated. A similar isolation is seen with the proteasome. Whereas, signaling systems through chemokines would be extended modules that are isolated through the binding of chemical signals to receptor proteins. These isolating features allow cells to achieve various objectives with minimal influence from cross-talk [2]. Yet, their connectivity allows complex guidance signals from one another.

More often, *in silico* models are being implemented in cancer research for the discovery of general principles and novel hypotheses that can guide the development of new treatments. Despite their potential, concrete examples of predictive models of cancer progression remain scarce. One reason is that most models have focused on single-cell type dynamics, ignoring the interactions between cancer cells and their local tumor microenvironment (TME). There have been a number of models that were used to study gene regulation at the single-cell scale, such as macrophage differentiation [3, 4, 5], T cell exhaustion [6], differentiation and plasticity of T helper cells [7, 8], and regulation of key genes in different tumor types [9, 10], including pancreatic cancers [11].

These models are all great steps towards control-based treatment optimization, but it has been demonstrated that the TME has a critical effect on the behaviour of cancer cells [12]. Ignoring the effect of cells and signals of the TME can generate confounding conclusions. For example, it was shown that in non-small cell lung cancer, the microenvironments of squamous tumors and adenocarcinomas are marked by differing recruitment of neutrophils and macrophages, respectively. *Ex vivo* experiments revealed the importance of the TME as a whole, especially when considering immunotherapy enhancement [13]. A similar observation has shown that removing pancreatic stellate cells from the TME *in silico* led to differing long-term outcomes because they form a protective layer around tumor cells [11].

To study the interplay of cancer cells with components of the TME, modelers developed multicellular models including cancer, stromal, immune, cytokines, and growth factors [14]. These models are typically multiscale integrating interactions at different scales, making it possible to simulate clinically relevant spatiotemporal scales and at the same time simulate the effect of molecular drugs on tumor progression [15, 16, 17, 18, 19, 20]. The high complexity of these models generates challenges for model validation such as the need to estimate too many model parameters. Moreover, a multi-scale model would likely provide more realistic simulations, but their complexity prevents the application of state of the art control techniques to find optimal therapeutic alterations [21]. Implementation of Boolean networks (BNs) provides a coarse grained description of signaling cascades without the need for tedious parameter fitting and can be simulated through stochastic discrete dynamical systems (SDDS) to streamline the modeling process and increase efficiency. These models have a well-studied and effective track record for capturing various biological system dynamics [22].

Pancreatic cancer is among the most lethal types of malignancies largely due to its difficulty to detect. The pancreas is located deep within body, and a standard doctor's exam will likely not reveal a tumor. Additionally, there is an absence of detecting and imaging techniques for early stage tumors. While PC only accounts for 3% of estimated new cases, it is fourth highest cause of cancer-related death in the United States [23], and there is only a 3% five year survival rate among its late stage patients (approximately 82% in Stages 3 and 4) [24, 25]. PC is widely known for its resistance to most traditional therapy protocols. According to the Pancreatic Cancer Action Network (PanCAN), most patients receive fluorouracil (5-FU) or a gemcitabine based treatments that are anti-metabolites targeting thymidylate synthase and ribonucleotide reductase, respectively [25]. In a prior work, we developed a multicellular model of pancreatic cancer (PC) based on a Boolean network approximation (Figure 1), and we used control strategies that direct the system from a diseased state to a healthy state by suppression (or expression) and disruption of specific signaling pathways. The model consists of pancreatic cancer cells (PCCs), pancreatic stellate cells (PSCs), cytokine molecules diffusing in the local microenvironment, and internal gene regulations for both cell types [11]. Throughout this writing, we will often denote PCC components with a subscript c and PSC components with subscript s . We then used the PC model to study the impact of four common mutations: KRAS, TP53, SMAD4, and CDKN2A [14, 26, 27].

Throughout, readers will find the following: Section 2 defines modularity and control, followed by a summary of the PC model dynamics and target efficacy in Section 3. We show that the modularity of GRNs is vulnerable to mutations (Section 4. 1), analogous perturbations to long-term dynamical outcomes occur that influences aggression and controllability (Sections 4. 1 and 4. 2), variance in topological rankings of the most phenotypically influential modules implies a strong relationship between structure and function (Section 4. 2), and we finally present evidence that the impact and location of mutations with respect to modular structure directly corresponds to the efficacy of single agent treatments *in silico* (Section 4. 3).

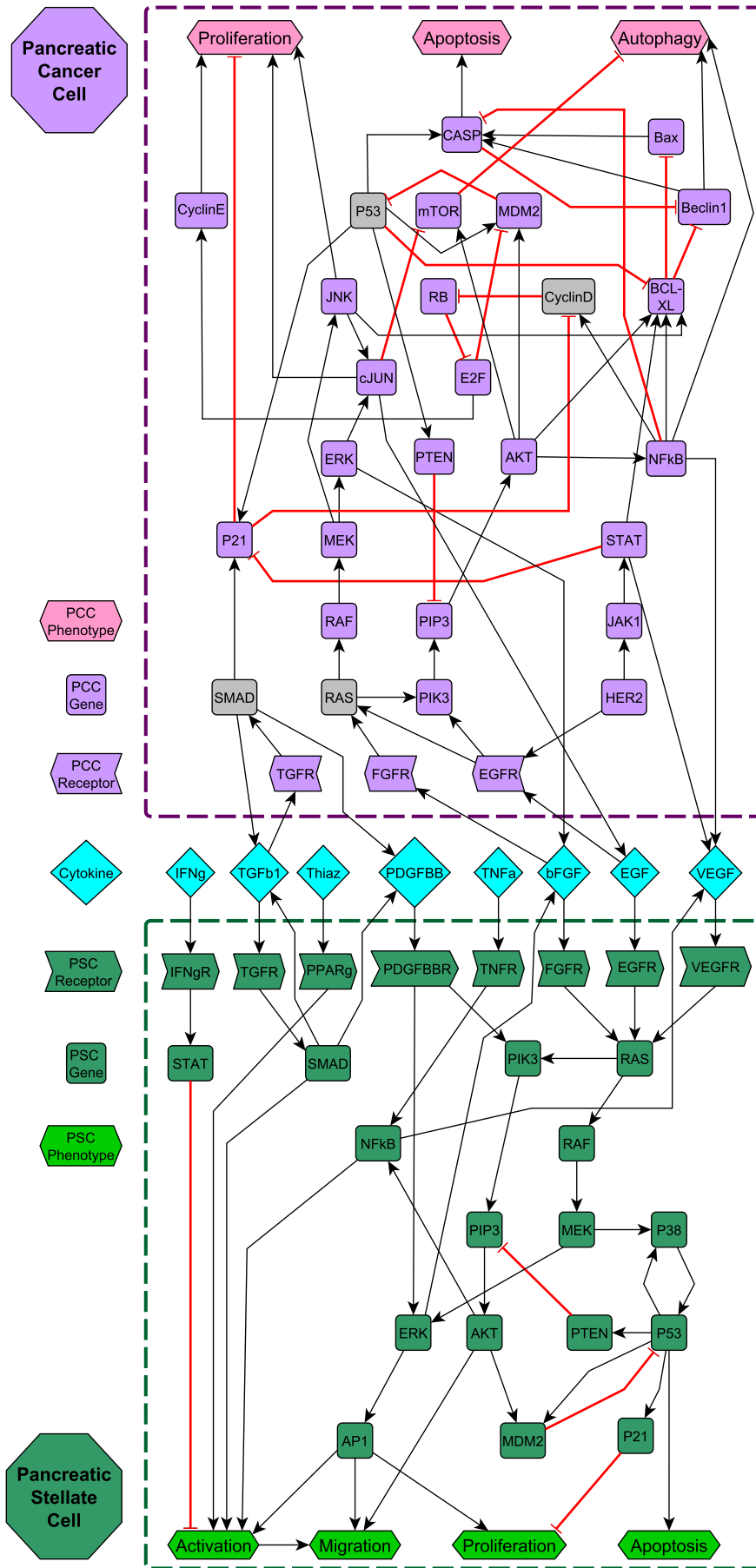


Figure 1:
Gene regulatory network model of pancreatic cancer.

Shapes and colors of nodes indicate their function and cell type (respectively), as shown in the legend. Black barbed arrows indicate signal expression, while red bar arrows indicate suppression. Grey nodes located in the PCC indicate prevalent mutant genes [11].

2 Network modularity

Systems biology can often build complicated structures from simpler building blocks, even though these simple blocks (i.e. modules) traditionally are not clearly defined. The concept of modularity detailed in [1] is structural by nature, in that, a *module* of a BN is a subnetwork in which the restriction of the network to the variables of a subgraph has a strongly connected wiring diagram. Further, the strongly connected components (SCCs) must be disjoint (nonoverlapping) with one-way connections between modules. This framework introduces both a structural and dynamic decomposition that encapsulates the dynamics of the whole system simply from the dynamics of its modules. Consequently, the decomposition yields a hierarchy among modules that can be used to specify controls. That is, by controlling key modules we are able to control the entire network.

More precisely, for a Boolean network F and subset of its variables S , we define a *subnetwork* of F as the restriction of F to S , denoted $F|_S$ (see the Appendix for an example of a restriction of a BN. For a formal definition of a restriction see [1]). For the Boolean network F with wiring diagram W , let W_1, \dots, W_m be the SCCs of W with variables S_i . The *modules* of F are $F|_{S_i}$, and setting $W_i \rightarrow W_j$ where there exists a node from W_i to W_j gives a directed acyclic graph $C = \{(i, j) | W_i \rightarrow W_j\}$ [1].

The dynamics of the state-space for Boolean network F are denoted as $\mathcal{D}(F)$, which is the collection of attractors. Further, if F is decomposable (say into subnetworks H and G), then we can write $F = H \times_P G$ which is called the *coupling* of H and G by scheme P . In the case where the dynamics of G are dependent on H , we call G *non-autonomous* denoted as \bar{G} . Then we adopt the following notation: let $A = A_1 \oplus A_2$ be an attractor of F with $A_1 \in \mathcal{D}(H)$ and $A_2 \in \mathcal{D}(G^{A_1})$ [1].

Lastly, a set of controls u stabilize a BN to attractor A if the only remaining attractor after inducing u is A . The decomposition strategy can be used to obtain controls for each module, that can then combine to control the entire system. That is, given a decomposable network $F = F_1 \times_P F_2$ and an attractor $A = A_1 \oplus A_2$ with $A_1 \in \mathcal{D}(F_1)$ and $A_2 \in \mathcal{D}(F_2^{A_1})$, assume u_1 is a control that stabilizes F_1 in A_1 , and u_2 is a control that stabilizes $F_2^{A_1}$ in A_2 . Then $u = (u_1, u_2)$ is a control that stabilizes F in A given that either A_1 or A_2 are steady states [1]. See an explicit example in the Appendix Section 8. 2

3 Model dynamics and target efficacy

Within the various mutation combinations of the PC model, we used our stochastic simulator based on the SDDS framework (see Section 8. 3 [28]) to derive aggressiveness scores from simulated long-term trajectory approximations (see Section 8. 1) [22]. The estimates in Figure 2 were tracked using phenotype expressions from only the PCC because the PSC is not considered malignant. Results showed that certain mutation combinations may indeed be more aggressive than others. We then performed statistical analysis on clinical gene expression data and derived survival curves that corroborated estimated aggressiveness scores. Phenotype control theory techniques revealed that sets of targets contained nodes within both the PCC and the PSC, highlighting PI3K and BAX as a strong combination [22]. Notice that cells in Figure 1 contain duplicate pathways. While targets are found in both the PSC and the PCC, targets such as PI3K_c and PI3K_s can be considered as one target biologically. This is because a PI3K inhibitor would act systemically rather than locally. Thus, when inducing controls in Figure 2, we assume systemic treatment. Here, we show a heatmap of projected mutation aggression compared to the application of target control. Note that as single agent targets, PI3K and BAX are not universally effective. However, their combination is an effective control across all mutation combinations. We believe this can be explained by topological module rankings, see Section 4. 1.

Uncontrolled Phenotype Aggression Scores																
Weight	N.I.	KRAS	TP53	CycD	SMAD	T-K	C-K	S-K	T-C	T-S	C-S	T-C-K	T-S-K	C-S-K	T-C-S	T-C-S-K
Same	1.09	1.12	3.03	1.56	1.14	3.05	1.56	0.71	1.92	3.72	1.55	1.95	3.81	1.58	1.92	1.93
High/Low	5.16	5.37	10.74	8.11	5.36	10.84	8.18	3.56	9.54	11.36	8.09	9.71	11.55	8.14	9.61	9.62
Low/High	2.79	2.80	7.51	1.70	2.82	7.49	1.67	2.63	2.06	11.08	1.65	2.07	11.36	1.72	2.00	2.05
Average	3.01	3.10	7.09	3.79	3.11	7.12	3.80	2.30	4.51	8.72	3.76	4.58	8.91	3.81	4.51	4.53
PI3K↓																
Same	-0.80	-0.78	0.08	-0.87	-0.82	0.07	-0.84	-0.78	0.04	0.09	-0.83	0.06	0.07	-0.83	0.04	0.04
High/Low	-0.68	-0.60	0.28	-0.78	-0.67	0.20	-0.70	-0.63	0.18	0.21	-0.69	0.22	0.22	-0.71	0.15	0.18
Low/High	-0.60	-0.51	0.24	-0.76	-0.65	0.30	-0.74	-0.47	0.16	0.41	-0.71	0.21	0.31	-0.71	0.20	0.15
Average	-0.70	-0.63	0.20	-0.80	-0.71	0.19	-0.76	-0.62	0.12	0.24	-0.75	0.16	0.20	-0.75	0.13	0.13
BAX↑																
Same	1.06	1.11	2.98	1.51	0.69	3.08	1.58	0.60	1.94	3.74	1.53	1.95	3.75	1.58	1.94	1.92
High/Low	5.23	5.36	10.70	8.04	3.60	10.85	8.26	3.34	9.70	11.43	8.13	9.70	11.44	8.28	9.73	9.60
Low/High	2.74	2.85	7.33	1.64	2.74	7.77	1.71	2.51	2.06	11.16	1.62	2.09	11.24	1.73	2.02	2.11
Average	3.01	3.11	7.00	3.73	2.34	7.23	3.85	2.15	4.57	8.78	3.76	4.58	8.81	3.86	4.57	4.55
PI3K↓ / BAX↑																
Same	-0.86	-0.86	-0.88	-0.90	-0.86	-0.91	-0.88	-0.91	-0.91	-0.88	-0.87	-0.89	-0.87	-0.89	-0.90	-0.91
High/Low	-0.65	-0.65	-0.75	-0.68	-0.60	-0.82	-0.66	-0.76	-0.75	-0.72	-0.63	-0.73	-0.69	-0.70	-0.80	-0.76
Low/High	-0.65	-0.62	-0.67	-0.80	-0.66	-0.76	-0.75	-0.76	-0.80	-0.65	-0.71	-0.73	-0.64	-0.76	-0.74	-0.82
Average	-0.72	-0.71	-0.77	-0.80	-0.71	-0.83	-0.76	-0.81	-0.82	-0.75	-0.74	-0.78	-0.73	-0.78	-0.82	-0.83

Figure 2: *Aggression Scores and Target Efficacy*. This figure shows aggression scores based on phenotype approximations by applying weights to trajectory probabilities. Here we include heat maps of aggression scores for each mutation combination (including the case when no mutation was induced (N.I.)), comparing cancer cell autophagy and proliferation while giving a negative weight to apoptosis. Row label “Same” indicates that the same weight was given to both autophagy and proliferation, “High/Low” indicates a high weight for autophagy but a low weight for proliferation, and “Low/High” indicates a low weight for autophagy but a high weight for proliferation. Scaling of the heat map ranges from orange (low score) to red (high score) based on the maximum and minimum values. Scaling shades of blue (cold) indicate non-aggressive or negative scores. See section 8.1 for more details.

4 Results

The following key observations will be visited:

- Network “depth” of KRAS mutation can break the standard modular structure (4.1)
- TP53 always directly impacts the module responsible for PCC apoptosis (4.1)
- Certain modules may not determine phenotypic states (4.1)
- Signaling from mutations to modules directly influences aggression (4.2)
- Topological ranking gaps correspond to aggression projections (4.2)
- Single agent targets are insufficient to out-compete downstream mutations (4.3)

4.1 Mutations perturb modular structure

In [1], the wild-type PC model from Figure 1 was first analyzed for modularity and revealed that the system contained three non-trivial modules. The module with top hierarchical rank in Figure 3, highlighted in yellow, is an autocrine loop of five nodes. Module 2, highlighted in green, contains thirty seven nodes spanning deep within both the PCC and PSC. The final module, highlighted in grey, is a negative feedback loop of only two nodes (referred throughout as the ‘duplex module’) [1]. Modules and their components

are easily determined through a *condensation graph* (C) [29], defined as a directed, acyclic graph whose nodes represent the strongly connected components of the original graph (G). C is topologically sorted to flow downstream. Condensation determines the nodes and edges in C by the components and connectivity in G such that: C contains a node for each strongly connected component in G , and C contains an edge between node I and node J if there is an edge from any node in component I to any node in component J of G . For example, the overall wild-type modular structure from Figure 3 was achieved using the graphs in Figure 4. Notice that Figure 4a contains all modules, both trivial and non-trivial. However, Figure 4b is a reduced modular structure with only non-trivial nodes, including the cardinality of the modules as well as its original numbering from the condensation graph in curly brackets for easy reference. Figure 4a shows how phenotypes typically lie at the end (i.e. the bottom of a basin) because they are the dynamical endpoint: Apoptosis_s (24), Proliferation_s (26), Migration_s (28), Activation_s (27), Autophagy_c (21), Apoptosis_c (20), and Proliferation_c (23). Thus, the condensation ordering begins with central nodes and expands distally to phenotypes. We anticipate that, due to this structure, we must use targets sufficiently upstream of phenotypes to subvert the impact of mutations.

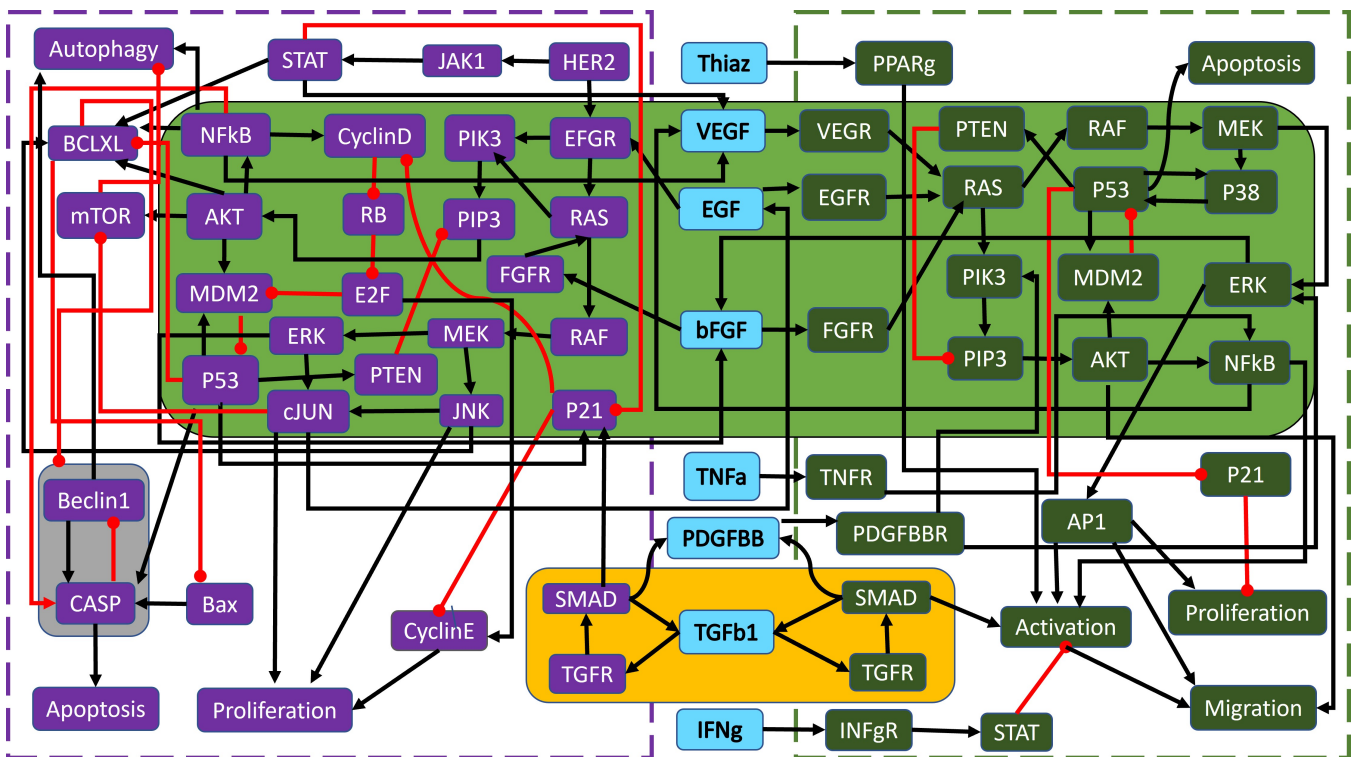


Figure 3: *Wild-type PC wiring diagram with modules*. Shown are the highlighted modules (yellow, green, and grey) for the wild-type PC model in Figure 1, adapted from [1]. A simplified structure can be seen in Figure 4b.

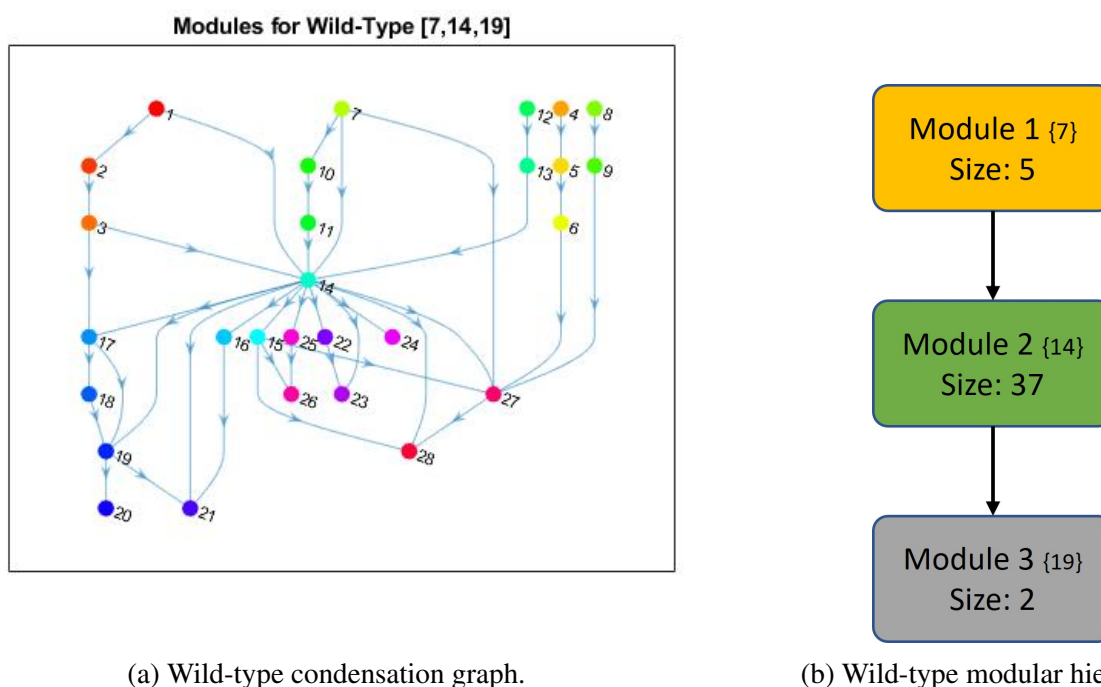
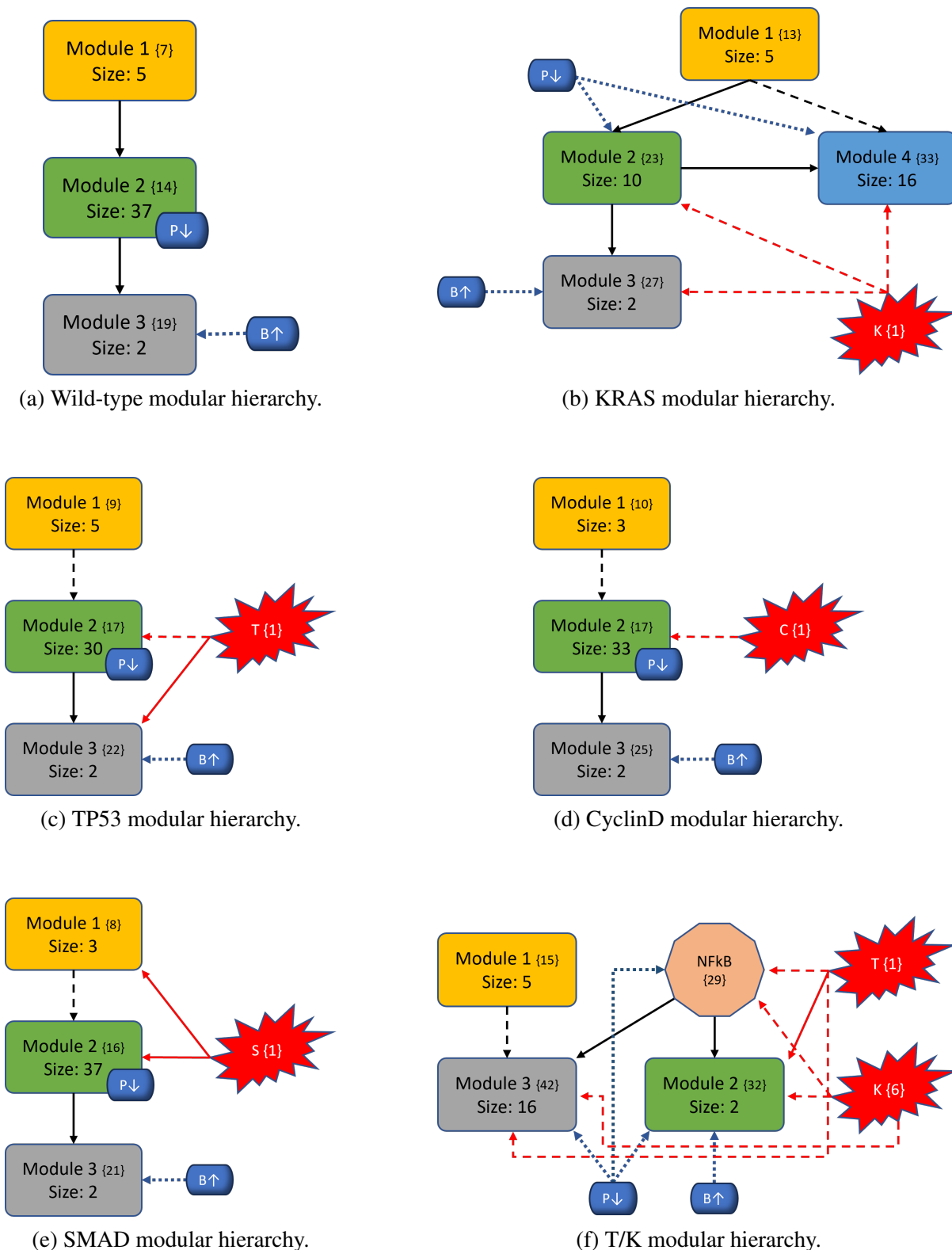


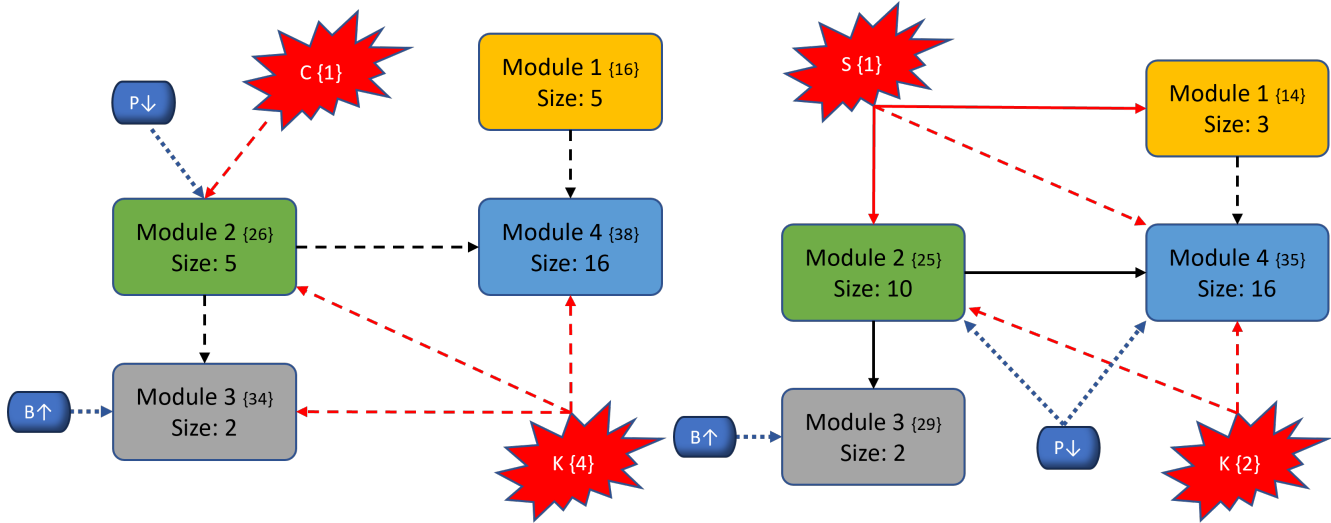
Figure 4: *Wild-type PC modular structure.* (a) Shown are the cumulative modules (trivial and nontrivial) ranked according to condensation, see Supplementary materials in Section 7. (b) The reduced modular structure with only nontrivial nodes, including the cardinality of the modules and its original condensation numbering.

To better understand the impact of mutations on the modular structure, we constructed modules in same manner as the wild-type after induced mutations. Mutations break SCCs and change the modularity structure both in magnitude of the modules (total nodes within a module) and amount of modules. In general, we maintain the three basic structures shown in Figures 5a, 5b, and 5f. Some cases hold to the same structure as the wild-type, but notably, every instance of KRAS induction yields split modules. We posit that this phenomenon is due to the location of KRAS in regard to its “depth” within the largest central module. Notice that Figure 3 shows SMAD, p53, and CyclinD are all located distally within their associated modules (Module 1 and all others in Module 2, respectively). Therefore, inducing their mutation merely shrinks the module size. When KRAS is induced, it is located more centrally within its original module and results in a broken component. For example, Figure 5b shows a gained module that only contains elements from the PSC that have split from their strongly connected PCC counterparts. Additionally, mutations do not always influence modules directly (e.g. Figure 5c), and there are scenarios where the mutation does not impact any modules at all (e.g. Figure 5o). Mutations yield anywhere from 28 to 51 modules in total (including trivial nodes), but non-trivial modules remain at either three or four in total (see Figures 8a and 8b in Appendix 8. 4).

We give a comprehensive breakdown of module structures for each mutation combination in Figure 5. Black arrows indicate communication between modules, red arrows indicate influence from mutations, dashed arrows illustrate indirect communication, and the PI3K (P↓) / BAX (B↑) targets are shown in blue with respect to their location inside or outside modules. Modules are numbered according to their hierarchical rankings from the condensation graphs, original identifiers from the condensation graph are in brackets (to easily reference graphs in Appendix 8. 4), and magnitude of each module are also included. More details are available in the supplementary materials (see Section 7).

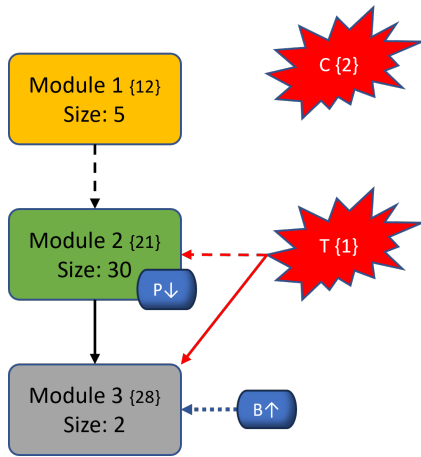
Figure 5: *Variants of PC modular structure*. Black arrows indicate communication between modules, red arrows indicate direct influence from mutations, and dashed arrows indicate indirect communication. Modules are ranked in order, and curly brackets contain the original numeration from the condensation graph. Lastly, the PI3K (P↓) / BAX (B↑) targets are shown in blue with respect to their location inside or outside modules.



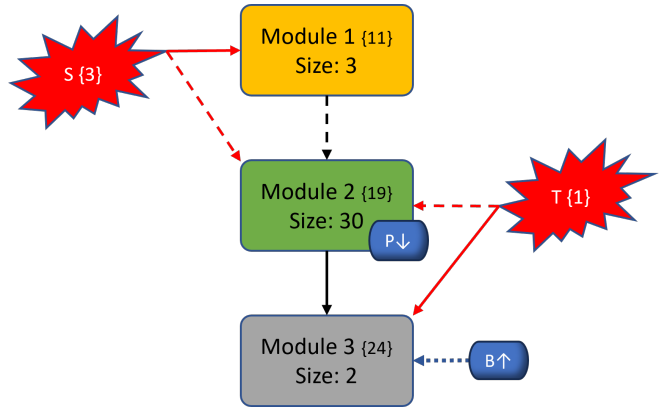


(g) C/K modular hierarchy.

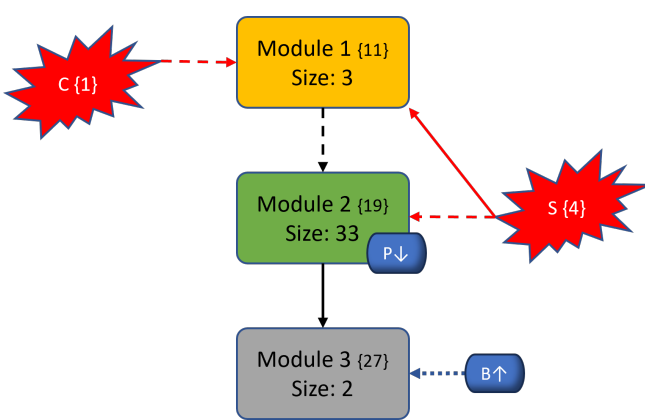
(h) S/K modular hierarchy.



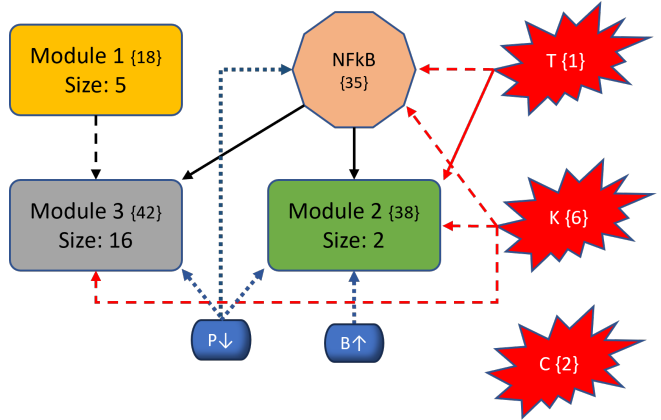
(i) T/C modular hierarchy.



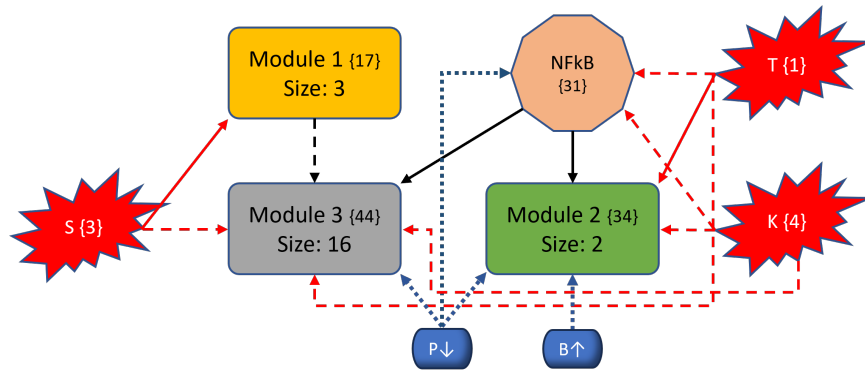
(j) T/S modular hierarchy.



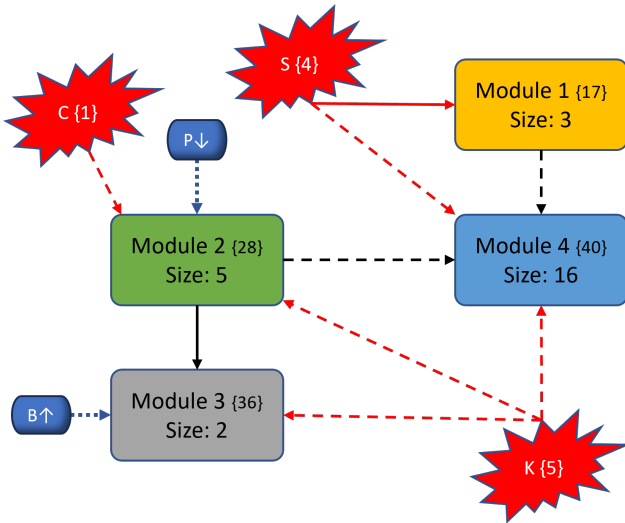
(k) C/S modular hierarchy.



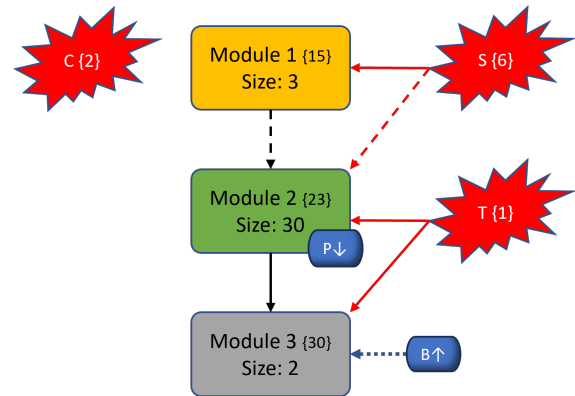
(l) T/C/K modular hierarchy.



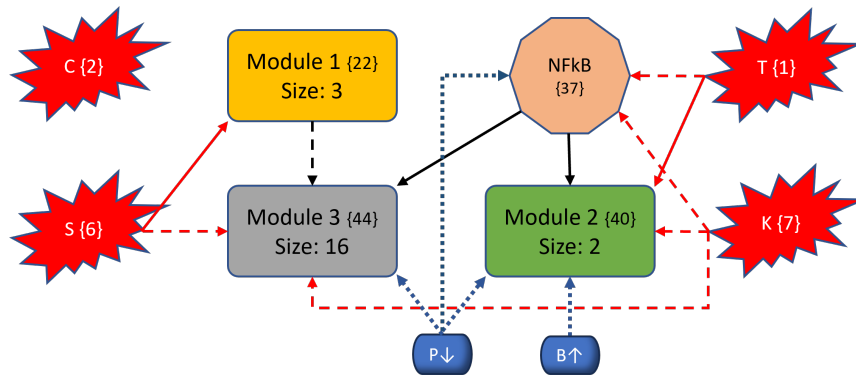
(m) T/S/K modular hierarchy.



(n) C/S/K modular hierarchy.



(o) T/C/S modular hierarchy.



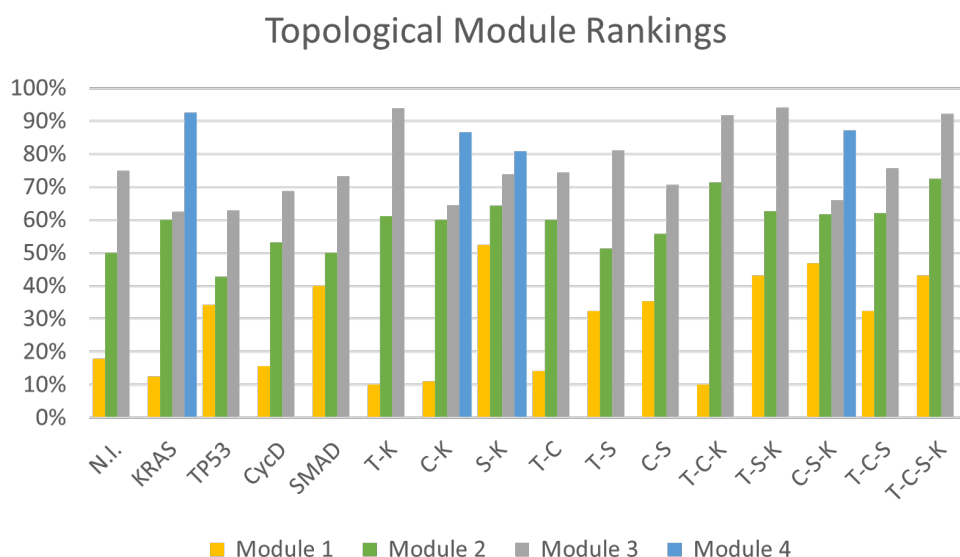
(p) T/C/S/K modular hierarchy.

4.2 Modular structure correlates with aggression

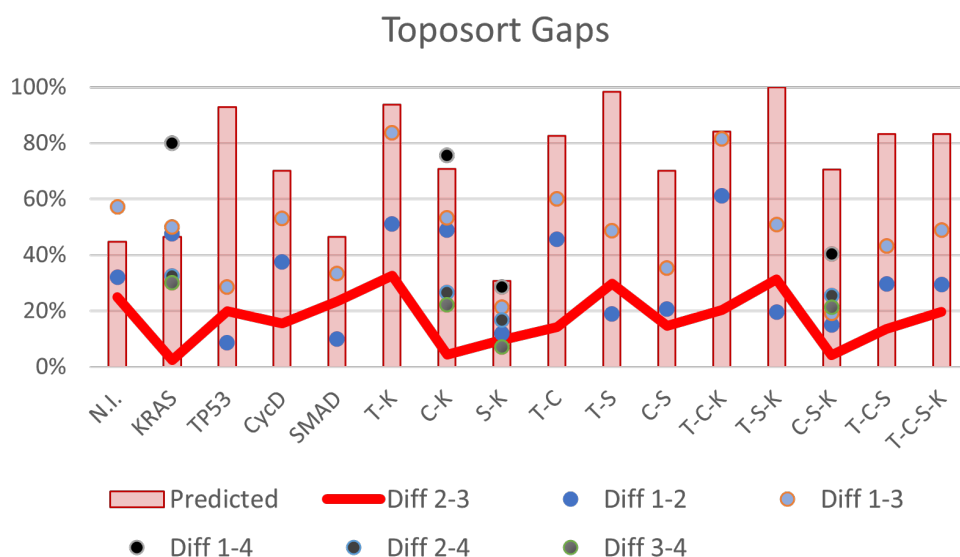
The most aggressive scores from Figure 2 are TP53, T/K, T/S, and T/S/K. The modularity structures for TP53 (Figure 5c) and T/S (Figure 5j) are unique, in that, the TP53 component always directly interacts with the duplex module that is responsible for apoptosis. Further, the structures for T/K (Figure 5f) and T/S/K (Figure 5m) are unique because (1) the duplex module is isolated from the standard downstream flow (2) the TP53 component influences all three downstream modules (versus only the trivial and duplex modules in T/C/K and T/C/S/K). This added layer of influence is what we believe drives a more aggressive prediction.

We further ranked the modules using a topological sorting algorithm called “toposort” from MATLAB [29], applied to the acyclic condensation graphs from Section 4.1. The topological ordering of an acyclic graph is an ordering of the nodes in the graph such that each node appears before its successors. The algorithm is based on a depth-first search, where a node is added to the beginning of the list after considering all of its descendants and returns a new order of nodes such that $i < j$ for every edge ($\text{ORDER}(i)$, $\text{ORDER}(j)$) in the original graph G . For example, consider a directed graph whose nodes represent the courses one must take in school and whose edges represent dependencies that certain courses must be completed before others. For such a graph, the topological sorting of the graph nodes produces a valid sequence in which the tasks could be performed [29].

Figure 6 shows the importance of modules and their ordering. Figure 6a indicates the topological ranking for each module across every mutation combination using ‘toposort’. Rankings are calculated according to the modules overall percentile score (e.g. a rank of 5 among 28 total modules yields 18%). Thus, rankings of a higher percentile would indicate a greater impact on final phenotypic state. An even more compelling correlation is found in Figure 6b, which shows the percentage difference between the rankings of modules. Recall that Module 4 only contains elements from the PSC, and we track aggressiveness from the PCC. Thus, we see that the gap between Modules 2 and 3 (i.e. the modules communicating PCC to phenotypes) as the most critical. The mutation combinations where the gap is the greatest are indeed the combinations predicted to be the most aggressive.



(a) Toposort overall module ranks.



(b) Toposort gaps between modules ranks.

Figure 6: *Toposort analysis*. (a) Shown are the overall topological rankings for each module among all mutation combinations. (b) Differences among each module ranking shows the variance (or distance) between modules. Gaps between Modules 2 and 3 are shown with a line graph since they best correlate with the predicted aggressive mutation combinations.

4.3 Mutations impact target efficacy

In [1], target analysis was performed on the wild-type PC model in Figure 3. For an overview of some of the most prominent control techniques, see [21]. Using apoptosis as the desired attractor, one could control the entire system by regulating a total of three nodes within the top two modules. The module with top hierarchical rank (yellow) is an autocrine loop with two fixed points and two 3-cycles. Therefore, the Feedback Vertex Set [30] strategy yielded control of Module 1 by pinning $TGF\beta 1$. Next, Module 2

(green) was searched for targets using algebraic methods [31] and revealed that pinning KRAS in both the PCC and PSC stabilized the module. The final module (grey) becomes constant after the initial three pinnings. That is, the entire wild-type system was found to be controllable by a single growth factor and systemic KRAS inhibition, both within the first two modules [1].

However, KRAS is well-known for being unmanageable even though progress is being made towards its targetability [32]. Therefore, targeting PI3K and BAX may be more achievable [22]. We showed in Figure 2 that the single agent controls of PI3K and BAX are insufficient to achieve PCC apoptosis across all mutation combinations. This is because TP53 directly influences the duplex module responsible for apoptosis signaling (Figure 5) Since PI3K is upstream of TP53, it must have help to circumvent or out-compete the mutation (through help from BAX). Such a combination is effective topologically because PI3K is heavily involved in communications between multiple modules (either directly inside influential modules or indirect communication to many modules), and BAX can override the deepest mutation signals.

5 Conclusion

Systems biology is continually searching for general principles and tools for their identification. The approach of network modularity adds to the repertoire of techniques that provide structural and dynamical analysis across complex biological systems. We have shown here that the given definition of modules captures the drastic perturbing effect that mutational occurrence can have on normal biological mechanisms. A major limitation of most other approaches to modularity is their focus on a static representation of GRNs. Clearly, living organisms are dynamic and need to be modeled as such [1, 33]. That is precisely what we advocate here. However, it is important to note that the decomposition presented in [1] does not preclude the existence of emergent properties. Each module is a complex dynamical system in itself. As we have shown, modules perturb other downstream modules, and their emergent properties propagate to other modules.

What we have provided herein uses a novel strategy of network modularity [1] to investigate the link between structure and function as it relates to the PC regulatory network. Namely, this link was shown to be strongly related to the variance in topological rankings of the most phenotypically influential modules. We have identified that the location of mutations, with respect to network depth and module positioning, expressly influence aggression and controllability. Thereby, presenting evidence that the impact and location of mutations with respect to modular structure directly corresponds to the efficacy of single agent treatments *in silico*. These cumulative results help provide more clarity on the impact of mutations in PC and posit the viability of using network modularity to study dynamical systems.

6 Acknowledgments

DP was supported by the NIH Training Grant T32CA165990. DM was partially supported by a Collaboration grant (850896) from the Simons Foundation.

7 Data Availability

All data and code used for running simulations, statistical analysis, and plotting is available on a GitHub repository at https://github.com/drplaughter/PCC_Mutations.

References

- [1] C. Kadelka, M. Wheeler, A. Veliz-Cuba, D. Murrugarra, and R. Laubenbacher. Modularity of biological systems: a link between structure and function. *J R Soc Interface*, 20(207):20230505, Oct 2023.
- [2] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray. From molecular to modular cell biology. *Nature*, 402(6761 Suppl):47–52, Dec 1999.
- [3] Alessandro Palma, Abdul Salam Jarrah, Paolo Tieri, Gianni Cesareni, and Filippo Castiglione. Gene regulatory network modeling of macrophage differentiation corroborates the continuum hypothesis of polarization states. *Frontiers in physiology*, 9:1659, 2018.
- [4] Julia Rex, Ute Albrecht, Christian Ehling, Maria Thomas, Ulrich M Zanger, Oliver Sawodny, Dieter Häussinger, Michael Ederer, Ronny Feuer, and Johannes G Bode. Model-based characterization of inflammatory gene expression patterns of activated macrophages. *PLoS computational biology*, 12(7):e1005018, 2016.
- [5] Filippo Castiglione, Paolo Tieri, Alessandro Palma, and Abdul Salam Jarrah. Statistical ensemble of gene regulatory networks of macrophage differentiation. *BMC bioinformatics*, 17(19):119–128, 2016.
- [6] Hamid Bolouri, Mary Young, Joshua Beilke, Rebecca Johnson, Brian Fox, Lu Huang, Cristina Costa Santini, Christopher Mark Hill, Anne-Renee van der Vuurst de Vries, Paul T Shannon, et al. Integrative network modeling reveals mechanisms underlying t cell exhaustion. *Scientific reports*, 10(1):1–15, 2020.
- [7] Luis Mendoza and Ioannis Xenarios. A method for the generation of standardized qualitative dynamical systems of regulatory networks. *Theoretical Biology and Medical Modelling*, 3(1):1–18, 2006.
- [8] Paolo Tieri, Vinca Prana, Teresa Colombo, Daniele Santoni, and Filippo Castiglione. Multi-scale simulation of t helper lymphocyte differentiation. In *Brazilian Symposium on Bioinformatics*, pages 123–134. Springer, 2014.
- [9] David Murrugarra and Boris Aguilar. *Algebraic and Combinatorial Computational Biology*, chapter 5, pages 149–150. Academic Press, 2018.
- [10] Minsoo Choi, Jue Shi, Sung Hoon Jung, Xi Chen, and Kwang-Hyun Cho. Attractor landscape analysis reveals feedback loops in the p53 network that control the cellular response to dna damage. *Science signaling*, 5(251):ra83–ra83, 2012.
- [11] Daniel Plaughter and D. Murrugarra. Modeling the pancreatic cancer microenvironment in search of control targets. *Bulletin of Mathematical Biology*, 83(115), 10 2021.
- [12] Haswanth Vundavilli, Aniruddha Datta, Chao Sima, Jianping Hua, Rosana Lopes, and Michael Bitner. In silico design and experimental validation of combination therapy for pancreatic cancer. *IEEE/ACM transactions on computational biology and bioinformatics*, 17(3):1010–1018, 2018.
- [13] Tanner J. DuCote, Xiulong Song, Kassandra J. Naughton, Fan Chen, Daniel R. Plaughter, Avery R. Childress, Abigail R. Edgin, Xufeng Qu, Jinze Liu, Jinpeng Liu, Fei Li, Kwok-Kin Wong, and Christine F. Brainson. Ezh2 inhibition promotes tumor immunogenicity in lung squamous cell carcinomas. *bioRxiv*, 2023.

- [14] Boris Aguilar, David L Gibbs, David J Reiss, Mark McConnell, Samuel A Danziger, Andrew Dervan, Matthew Trotter, Douglas Bassett, Robert Hershberg, Alexander V Ratushny, and Ilya Shmulevich. A generalizable data-driven multicellular model of pancreatic ductal adenocarcinoma. *Gigascience*, 9(7), 07 2020.
- [15] Andrea Padoan, Mario Plebani, and Daniela Basso. Inflammation and pancreatic cancer: Focus on metabolism, cytokines, and immunity. *International Journal of Molecular Sciences*, 20:676, 02 2019.
- [16] Jörg Kleeff, Philipp Beckhove, Irene Esposito, Stephan Herzig, Peter E. Huber, J. Matthias Löhr, and Helmut Friess. Pancreatic cancer microenvironment. *International Journal of Cancer*, 121(4):699–705, 2007.
- [17] Jesse Gore and Murray Korc. Pancreatic cancer stroma: Friend or foe? *Cancer cell*, 25:711–712, 06 2014.
- [18] Mert Erkan, Carolin Reiser-Erkan, Christoph Michalski, and Jörg Kleeff. Tumor microenvironment and progression of pancreatic cancer. *Experimental oncology*, 32:128–31, 09 2010.
- [19] Buckminster Farrow, Daniel Albo, and David H. Berger. The role of the tumor microenvironment in the progression of pancreatic cancer. *Journal of Surgical Research*, 149(2):319–328, 2008.
- [20] Christine Feig, Aarthi Gopinathan, Albrecht Neesse, Derek S. Chan, Natalie Cook, and David A. Tuveson. The pancreas cancer microenvironment. *Clinical Cancer Research*, 18(16):4266–4276, 2012.
- [21] D. Plaughter and D. Murrugarra. Phenotype Control techniques for Boolean gene regulatory networks. *Bull Math Biol*, 85(10):89, Aug 2023.
- [22] Daniel Plaughter, Boris Aguilar, and David Murrugarra. Uncovering potential interventions for pancreatic cancer patients via mathematical modeling. *Journal of Theoretical Biology*, 548:111197, 2022.
- [23] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal. Cancer statistics, 2023. *CA Cancer J Clin*, 73(1):17–48, Jan 2023.
- [24] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca Siegel, Lindsey Torre, and Ahmedin Jemal. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries: Global cancer statistics 2018. *CA: A Cancer Journal for Clinicians*, 68, 09 2018.
- [25] M. J. Pishvaian, R. J. Bender, D. Halverson, L. Rahib, A. E. Hendifar, S. Mikhail, V. Chung, V. J. Picozzi, D. Sohal, E. M. Blais, K. Mason, E. E. Lyons, L. M. Matrisian, J. R. Brody, S. Madhavan, and E. F. Petricoin. Molecular Profiling of Patients with Pancreatic Cancer: Initial Results from the Know Your Tumor Initiative. *Clin Cancer Res*, 24(20):5018–5027, Oct 2018.
- [26] Chu Wu, Ping-Lian Yang, Bingxue Liu, and Yunlian Tang. Is there a cdkn2a-centric network in pancreatic ductal adenocarcinoma? *OncoTargets and therapy*, 13:2551 – 2562, 2020.
- [27] Kirsten L. Bryant, Joseph D. Mancias, Alec C. Kimmelman, and Channing J. Der. Kras: feeding pancreatic cancer proliferation. *Trends in Biochemical Sciences*, 39(2):91–100, 2014.

- [28] David Murrugarra, Alan Veliz-Cuba, Boris Aguilar, Seda Arat, and Reinhard Laubenbacher. Modeling stochasticity and variability in gene regulatory networks. *EURASIP Journal on Bioinformatics and Systems Biology*, 2012(1):5, 2012.
- [29] MATLAB. *version 9.12.0.2039608 (R2022a) Update 5*. The MathWorks Inc., Natick, Massachusetts, 2022.
- [30] Jorge Gomez Tejada Zañudo, Gang Yang, and Réka Albert. Structure-based control of complex networks with nonlinear dynamics. *Proc Natl Acad Sci U S A*, 114(28):7234–7239, 07 2017.
- [31] Luis Sordo Vieira, Reinhard C Laubenbacher, and David Murrugarra. Control of intracellular molecular networks using algebraic methods. *Bulletin of mathematical biology*, 82(1):1–22, 2020.
- [32] A. Rosenzweig, J. Berlin, S. Chari, H. Kindler, L. Matrisian, A. Mayoral, J. Mills, N. Nissen, V. Picozzi, F. Zelada-Arenas, and J. Fleming. Management of Patients With Pancreatic Cancer Using the Right Track Model. *Oncologist*, 28(7):584–595, Jul 2023.
- [33] A. Jimenez, J. Cotterell, A. Munteanu, and J. Sharpe. A spectrum of modularity in multi-functional gene circuits. *Mol Syst Biol*, 13(4):925, Apr 2017.
- [34] Shenghong Yang, Xiaoxu Wang, Gianmarco Contino, Marc Liesa, Ergun Sahin, Haoqiang Ying, Alexandra Bause, Yinghua Li, Jayne M. Stommel, Giacomo Dell’Antonio, Josef Mautner, Giovanni Tonon, Marcia Haigis, Orian S. Shirihai, Claudio Doglioni, Nabeel Bardeesy, and Alec C. Kimmelman. Pancreatic cancers require autophagy for tumor growth. *Genes & Development*, 25(7):717–729, 2011.
- [35] Chul Won Yun and Sang Hun Lee. The roles of autophagy in cancer. *International Journal of Molecular Sciences*, 19(11), 2018.
- [36] Atsushi Mochizuki, Bernold Fiedler, Gen Kurosawa, and Daisuke Saito. Dynamics and control at feedback vertex sets. ii: A faithful monitor to determine the diversity of molecular activities in regulatory networks. *Journal of Theoretical Biology*, 335:130 – 146, 2013.

8 Appendix

8.1 Aggression scores

Uncontrolled Trajectory Approximations																
	N.I.	KRAS	TP53	CycD	SMAD	T-K	C-K	S-K	T-C	T-S	C-S	T-C-K	T-S-K	C-S-K	T-C-S	T-C-S-K
Aut _c	0.509	0.531	0.964	0.818	0.528	0.974	0.827	0.356	0.953	0.955	0.817	0.97	0.967	0.82	0.962	0.961
Apo _c	0.359	0.356	0.021	0.109	0.343	0.011	0.118	0.479	0.025	0.035	0.105	0.018	0.013	0.1	0.028	0.017
Pro _c	0.213	0.209	0.56	0.017	0.211	0.555	0.013	0.24	0.018	0.92	0.012	0.015	0.944	0.018	0.01	0.014

Table 1: *Expression Approximations*. This table records the approximate phenotype expressions for the PCC in Figure 1. Given 1,000 random initial states, these results show trajectory approximations after 300 time steps (i.e. function updates) with 1% noise.

We derived aggressiveness scores for each mutation combination using long-term trajectory approximations. Simulations were run using 1000 random initializations, 300 time steps, and 1% noise to achieve an approximate probability of phenotype expression. In Table, 1, we see that the non-induced (N.I.) system showed levels of 51% autophagy, 36% apoptosis, and 21% proliferation. The heat maps in Figure 2 are sorted with column-wise mutation groups and used to compare cancer cell autophagy and proliferation while giving a negative weight ($\omega = -1$) to apoptosis. The row label “Same” indicates that the same weight was given to both autophagy and proliferation (used value $\omega = 2$ for both), “High/Low” indicates a high weight for autophagy ($\omega = 10$) but a low weight for proliferation ($\omega = 2$), and “Low/High” indicates a low weight for autophagy ($\omega = 2$) but a high weight for proliferation ($\omega = 10$). Thus, scores were calculated using:

$$\text{Score} = \text{Aut}_c \times \omega_1 + \text{Pro}_c \times \omega_2 + \text{Apop}_c \times (-1) \quad \text{where} \quad \omega_{1,2} \in \{2, 10\}$$

Scaling of the heat map ranges orange (low score) to red (high score) based on the maximum and minimum values in each table. However, blue shading (i.e. cold) indicates a negative score, which is interpreted as successful depletion of aggression. See [22] sections 2.3 and 4.4 for more details.

Lastly, we justify the positive weight given to autophagy, which is a natural process where cells heal themselves. The cell will break down any damaged or unnecessary components, and it will reallocate the nutrients from these processes to those that are essential. However, studies have shown that autophagy is required for pancreatic tumor growth [34]. Autophagy can help tumors overcome conditions such as hypoxia and nutrient deprivation. Within tumors, cells can exist under hypoxic conditions. If activated autophagy is then suppressed by deletion of Beclin 1, studies have shown increased cell death. It has also been observed that autophagy is increased in KRAS mutated cells, and aids in survival of the cancer cells while experiencing nutrient starvation. Further, animal studies have shown that autophagy contributes to tumor-cell survival by enhancing stress tolerance and supplying nutrients to meet the metabolic demands of tumors. Once suppression of autophagy occurred, there was an observance of tumor-cell death [35].

Note: our aggression scores are based on combinations autophagy, apoptosis, and proliferation, merely one method among many for estimating aggression. Moreover, the attractor analysis (see [11, 22] indicated that certain mutation combinations yield a large basin for attractors with both autophagy and proliferation expression. It is likely that modular structure alone is not enough to determine aggression and target cardinality. Rather, it should be used alongside other analyses.

8.2 Mathematical framework of modules

Consider the network in Figure 7a, which can be written as

$$F(x_1, x_2, x_3, x_4, x_5, x_6) = (x_3, x_1, x_2, x_1x_6, x_4, x_5).$$

Subnetworks are defined according to the dependencies of variables encoded by the wiring diagram [1]. For example, the subnetwork $F|_{\{x_4, x_5, x_6\}} = (x_1x_6, x_4, x_5)$ is the restriction of F to $\{x_4, x_5, x_6\}$ with external parameter x_1 . From the given wiring diagram, we derive two SCCs where Module 1 (red in 7b) flows into Module 2 (green in 7b). That is, $F = F_1 \times F_2$ with

$$\begin{aligned} F_1(x_1, x_2, x_3) &= (x_3, x_1, x_2) \\ F_2(x_4, x_5, x_6) &= (x_6, x_4, x_5) \\ \overline{F_2}(x_4, x_5, x_6) &= (x_1x_6, x_4, x_5) \\ \mathcal{D}(F_1) &= \{000, 111, [001, 100, 010], [011, 101, 110]\}. \end{aligned}$$

Suppose we aim to stabilize the system into $y = 000000$. First we see that either $x_1 = 0$, $x_2 = 0$ or $x_3 = 0$ stabilize Module 1 (i.e. F_1) to $A_1 = 000$ by applying the Feedback Vertex Set method [30, 36]. Likewise, $x_4 = 0$, $x_5 = 0$ or $x_6 = 0$ stabilize Module 2 (i.e. $F_2^{A_1}$) to $A_2 = 000$. Thus, we conclude that $u = (x_1 = 0, x_6 = 0)$ achieves the desired result.

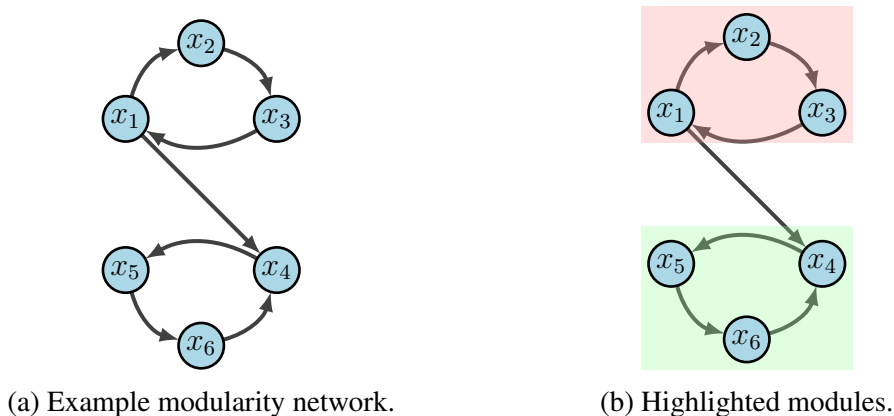


Figure 7: Modularity example.

8.3 Stochastic discrete dynamical systems

Synchronous updating schedules produce deterministic dynamics, where all nodes are updated simultaneously (i.e. in sync). The stochastic discrete dynamical systems (SDDS) framework developed by [9] incorporates Markov chain tools to study long-term dynamics of Boolean networks. By definition, an SDDS of the variables (x_1, x_2, \dots, x_n) is a collection of n triples denoted $\hat{F} = \{f_k, p_k^\uparrow, p_k^\downarrow\}_{k=1}^n$, where for $k = 1, \dots, n$,

- $f_k : \{0, 1\}^n \rightarrow \{0, 1\}$ is the update function for x_k
- $p_k^\uparrow \in [0, 1]$ is the activation propensity
- $p_k^\downarrow \in [0, 1]$ is the deactivation propensity

Consider the state-space S , consisting of all possible states of the system. If $x = (x_1, \dots, x_n) \in S$ and $y = (y_1, \dots, y_n) \in S$, then the probability of transitioning from x to y is

$$a_{x,y} = \prod_{i=1}^n P(x_i \rightarrow y_i)$$

where

$$P(x_i \rightarrow f_i(x)) = \begin{cases} p_k^\uparrow, & \text{if } x_i < f_i(x) \\ p_k^\downarrow, & \text{if } x_i > f_i(x) \\ 1, & \text{if } x_i = f_i(x) \end{cases} \quad \text{and} \quad P(x_i \rightarrow x_i) = \begin{cases} 1 - p_k^\uparrow, & \text{if } x_i < f_i(x) \\ 1 - p_k^\downarrow, & \text{if } x_i > f_i(x) \\ 1, & \text{if } x_i = f_i(x) \end{cases}$$

Then $P(x_i \rightarrow y_i) = 0$ for any $y_i \notin \{x_i, f_i(x)\}$. When propensities are set to $p = 1$, we have a traditional BN [11]. With this framework, we built a simulator that takes random initial states as inputs and then tracks the trajectory of each node through time. Long-term phenotype expression probabilities can then be estimated, as well as network dynamics with (and without) controls.

8.4 Tables and Graphs

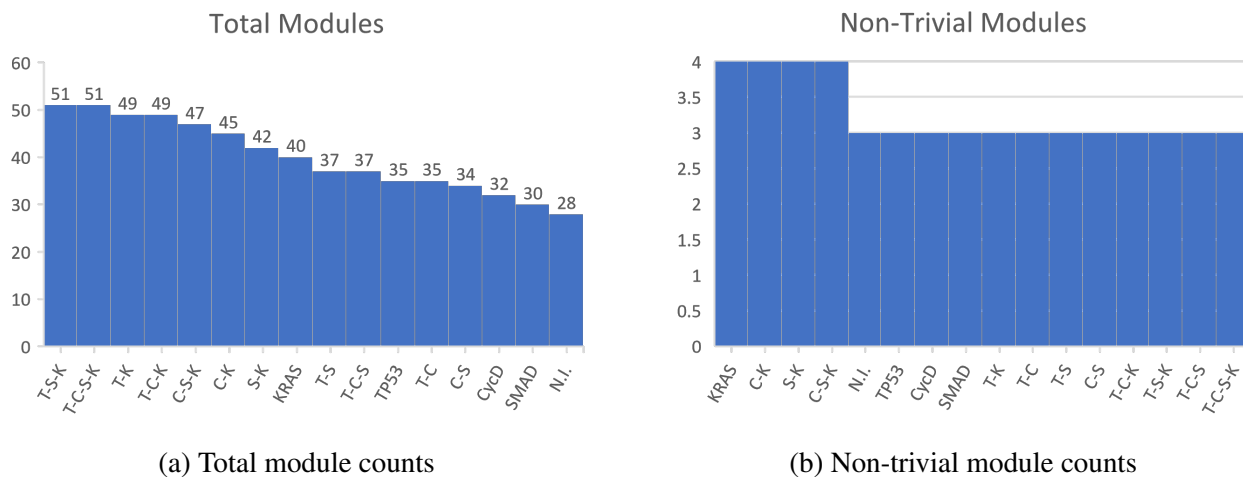
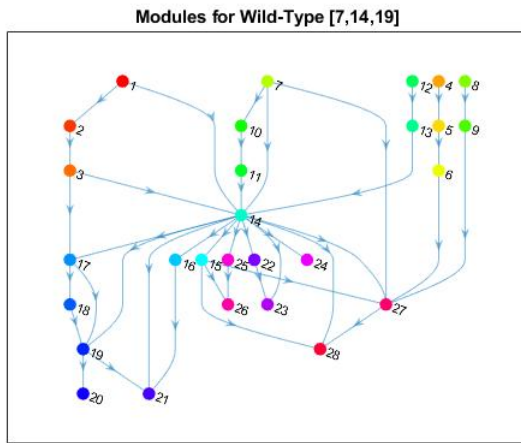
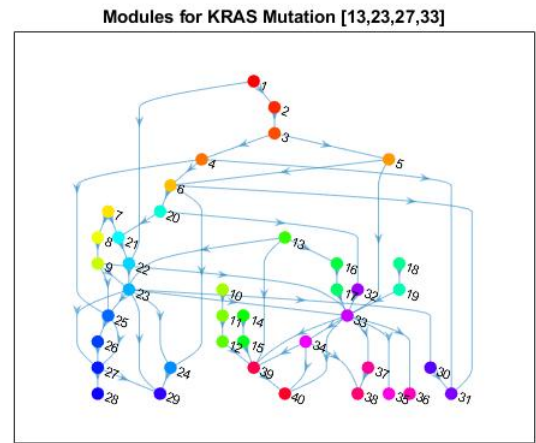


Figure 8: *Module counts*

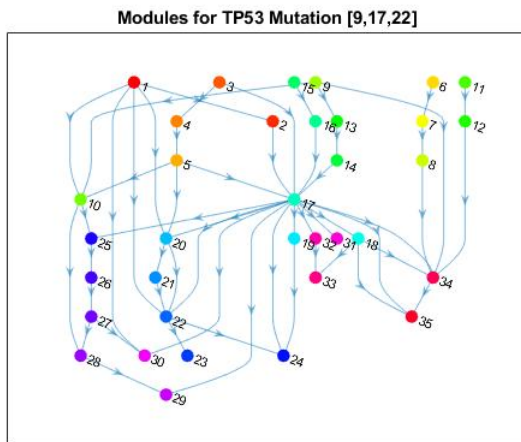
Figure 9: *PC condensation graphs*. Included are all condensation graphs for each mutation combination. These are directed, acyclic graphs that are topologically sorted, and whose nodes represent the strongly connected components of Figure 1. Colors of nodes are based on the components they represent, and node numbers correspond to bin numbers (see supplementary materials through Section 7).



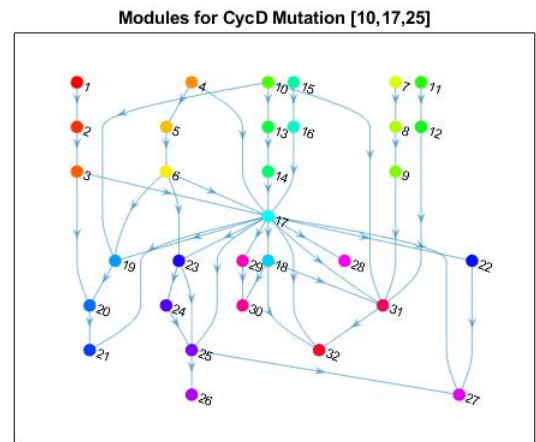
(a) Wild-type condensation graph.



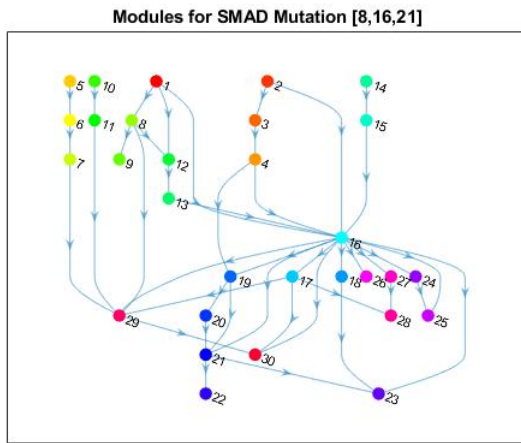
(b) KRAS condensation graph.



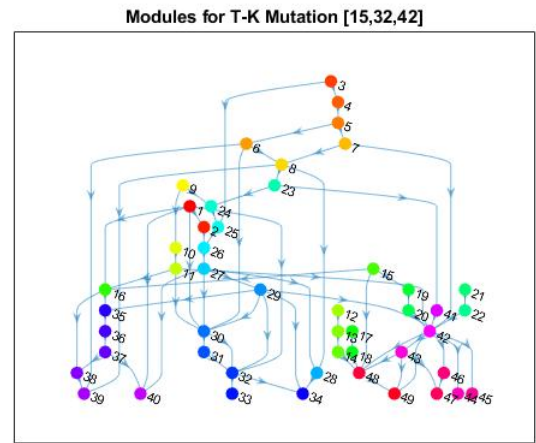
(c) TP53 condensation graph.



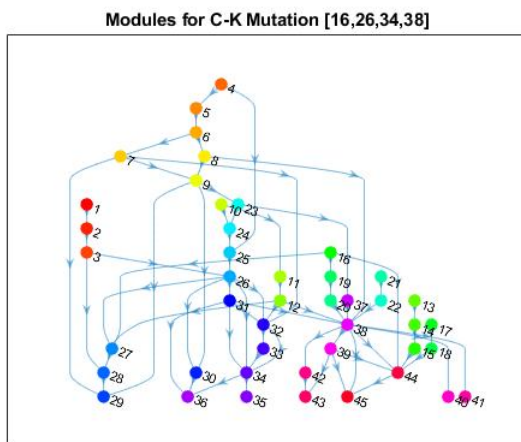
(d) CyclinD condensation graph.



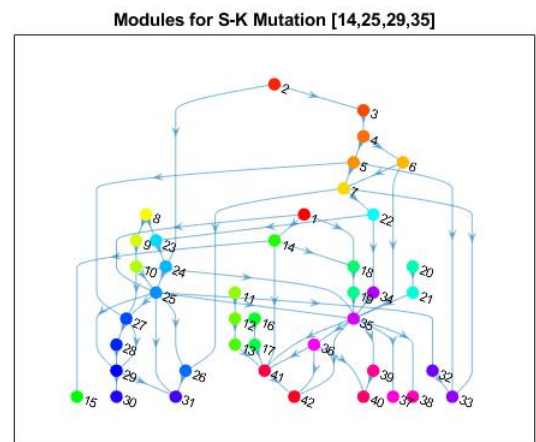
(e) SMAD condensation graph.



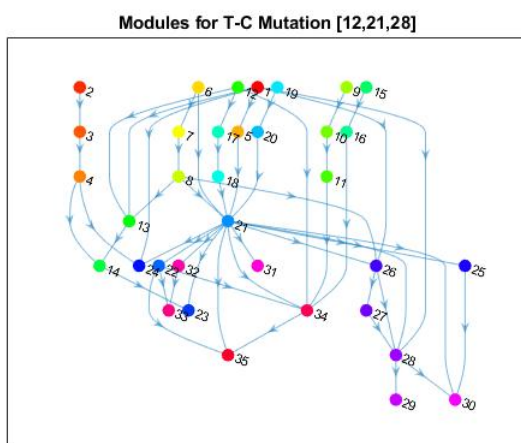
(f) T/K condensation graph.



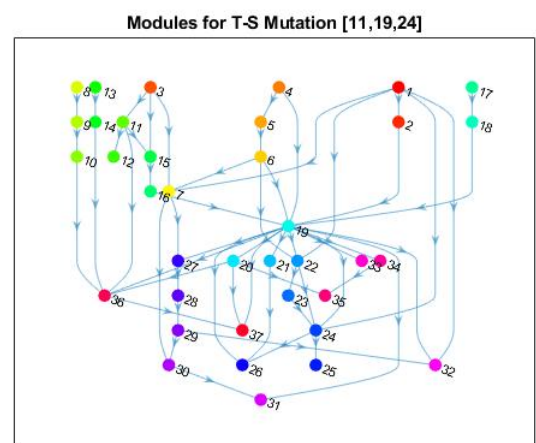
(g) C/K condensation graph.



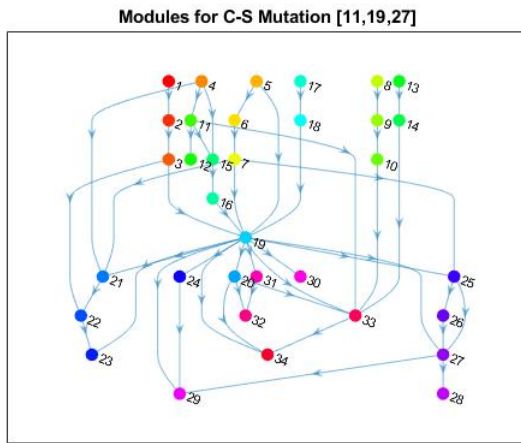
(h) S/K condensation graph.



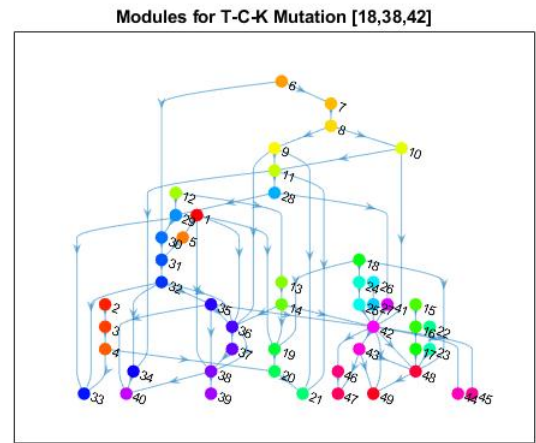
(i) T/C condensation graph.



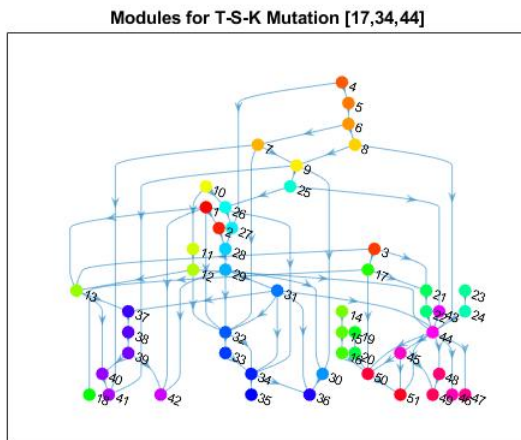
(j) T/S condensation graph.



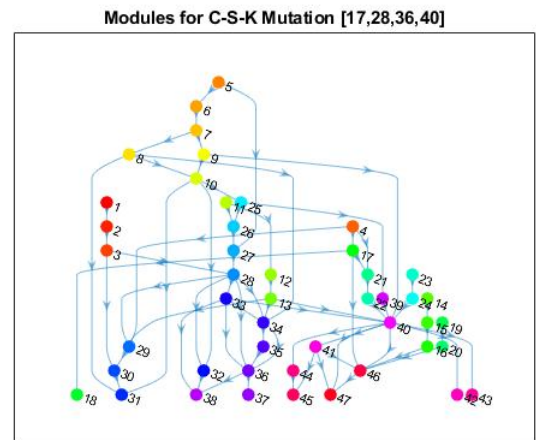
(k) C/S condensation graph.



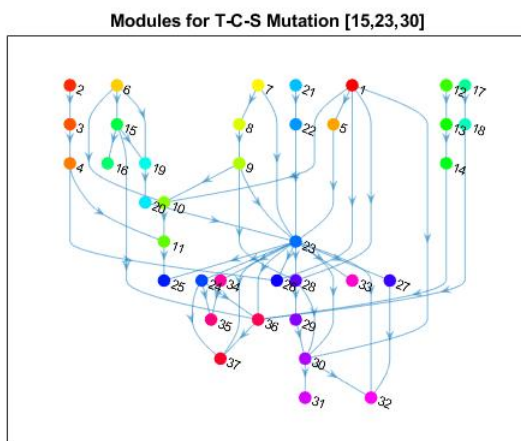
(l) T/C/K condensation graph.



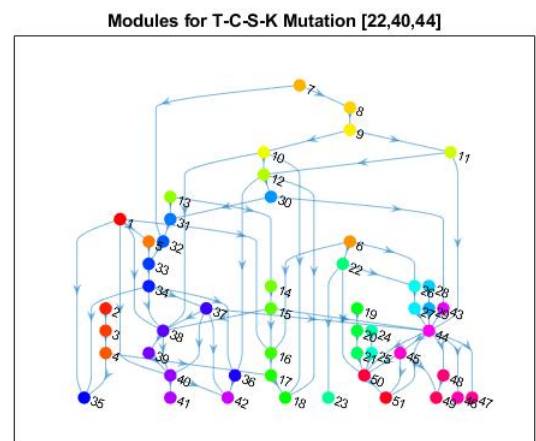
(m) T/S/K condensation graph.



(n) C/S/K condensation graph.



(o) T/C/S condensation graph.



(p) T/C/S/K condensation graph.

Module 1 {10}
Size: 3

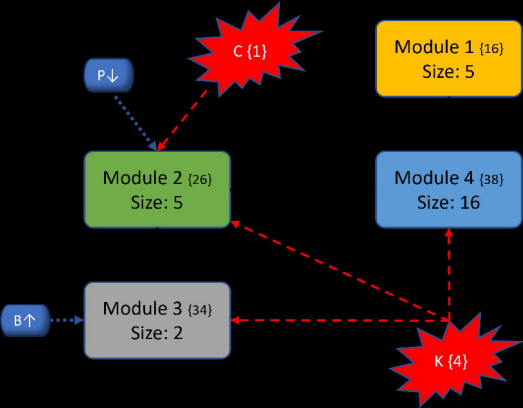
Module 2 {17}
Size: 33

Module 3 {25}
Size: 2

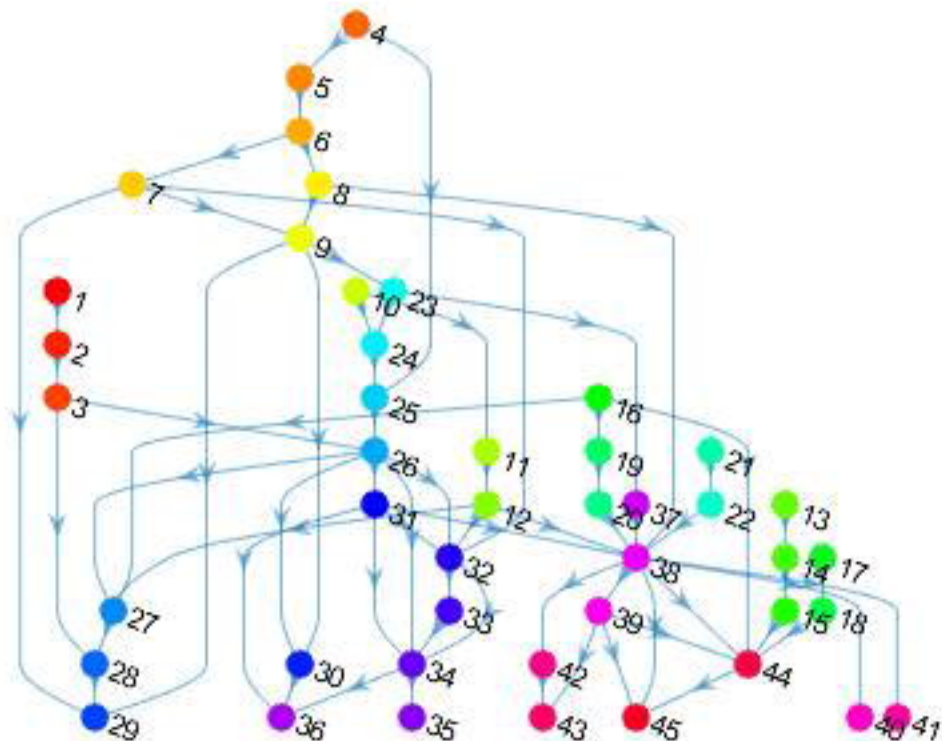
P↓

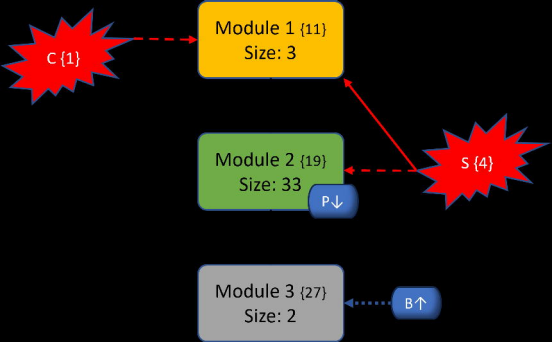
C {1}

B↑

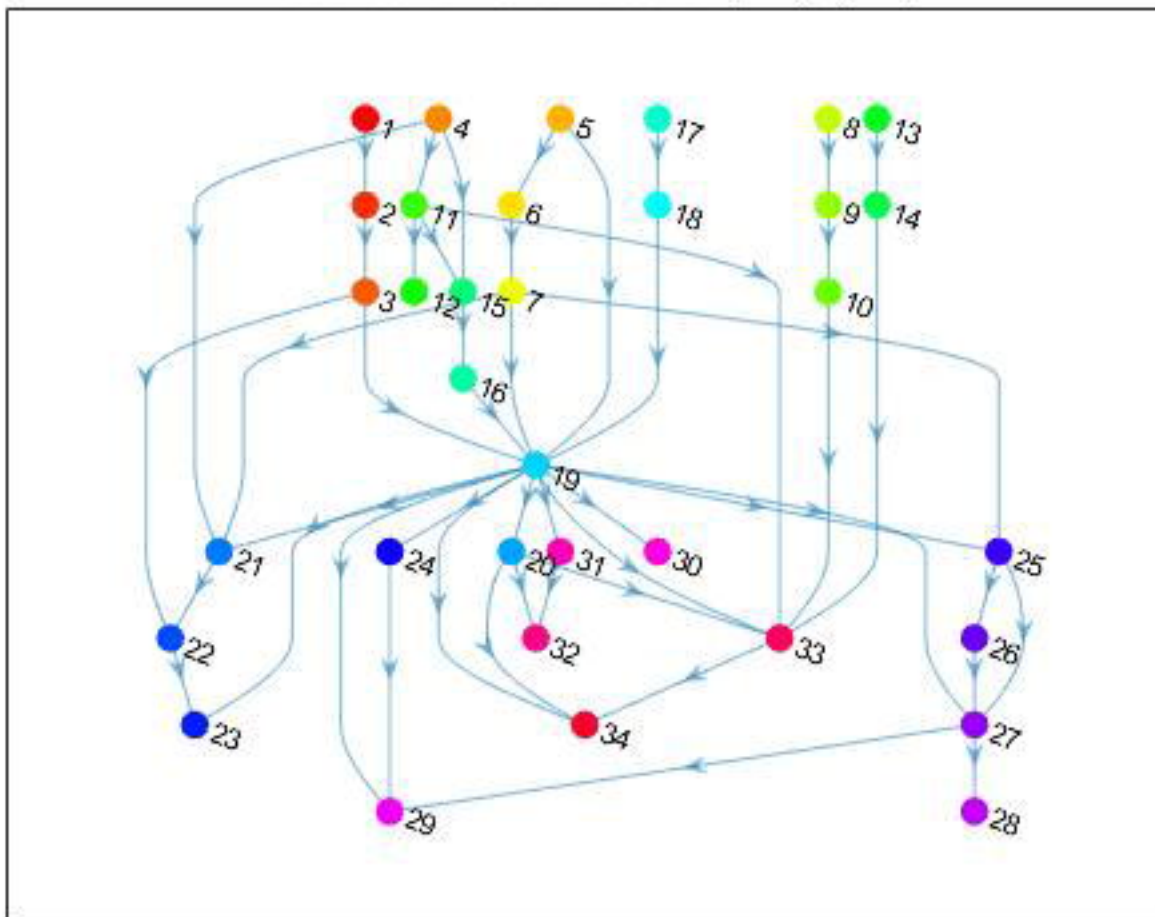


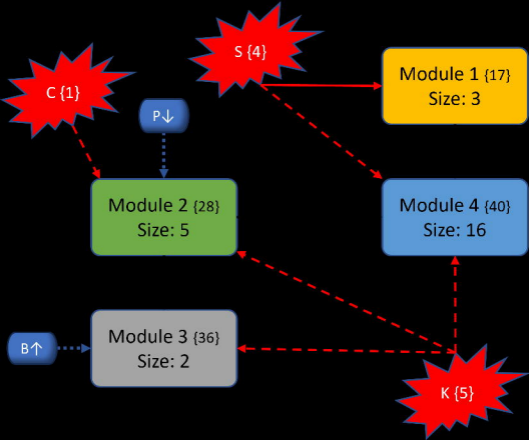
Modules for C-K Mutation [16,26,34,38]



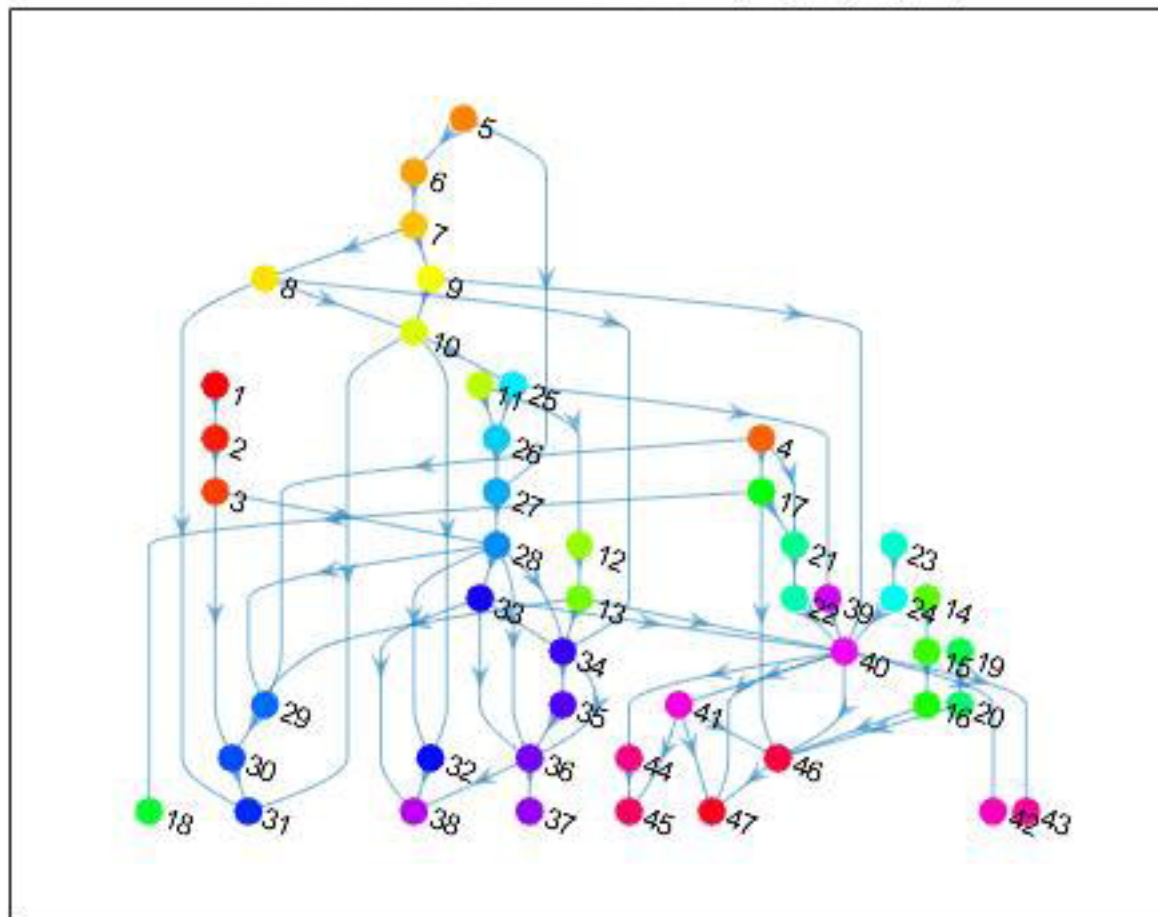


Modules for C-S Mutation [11,19,27]

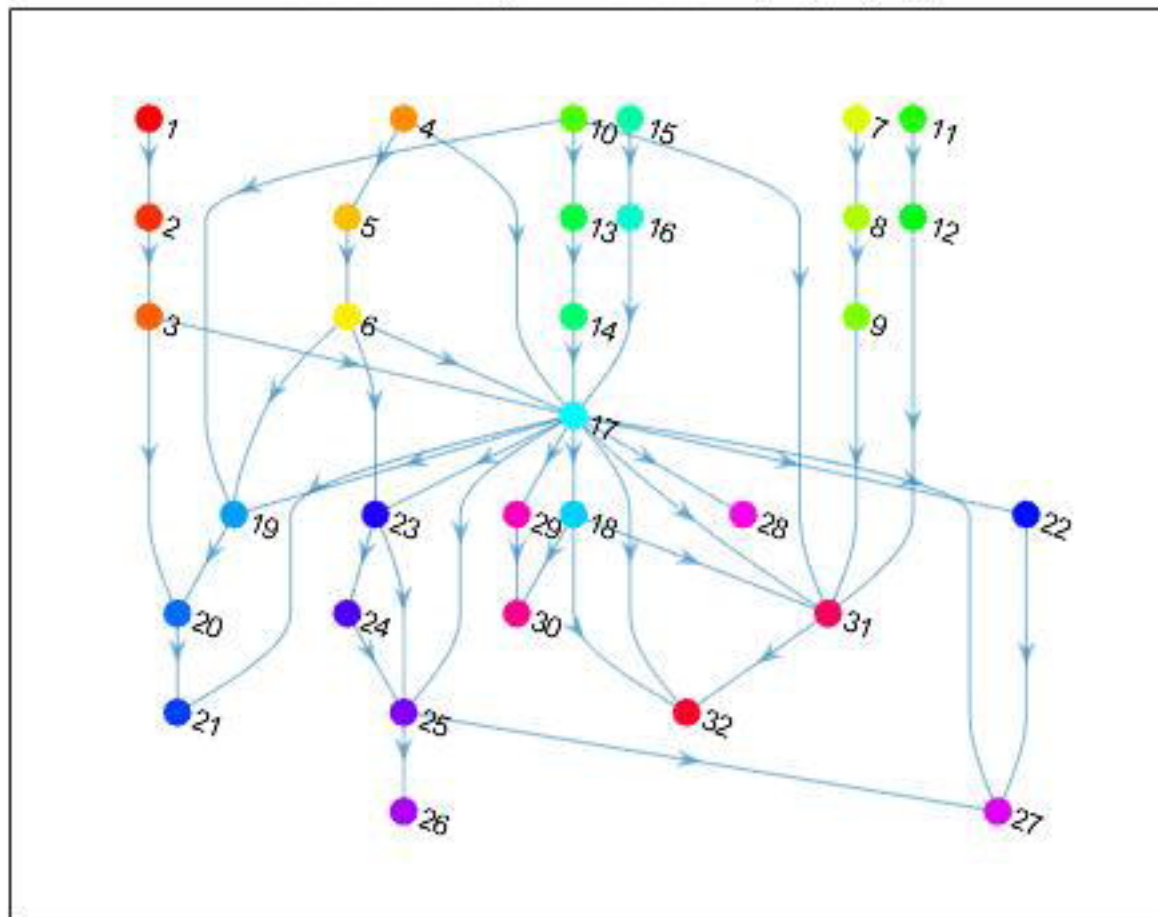




Modules for C-S-K Mutation [17,28,36,40]



Modules for CycD Mutation [10,17,25]



Uncontrolled Phenotype Aggression Scores

1.09	1.12	3.03	1.56	1.14	3.05	1.56	0.71	1.92	3.72	1.55	1.95	3.81	1.58	1.92	1.93
5.16	5.37	10.74	8.11	5.36	10.84	8.18	3.56	9.54	11.36	8.09	9.71	11.55	8.14	9.61	9.62
2.79	2.80	7.51	1.70	2.82	7.49	1.67	2.63	2.06	11.08	1.65	2.07	11.36	1.72	2.00	2.05
3.01	3.10	7.09	3.79	3.11	7.12	3.80	2.30	4.51	8.72	3.76	4.58	8.91	3.81	4.51	4.53

PI3K↓

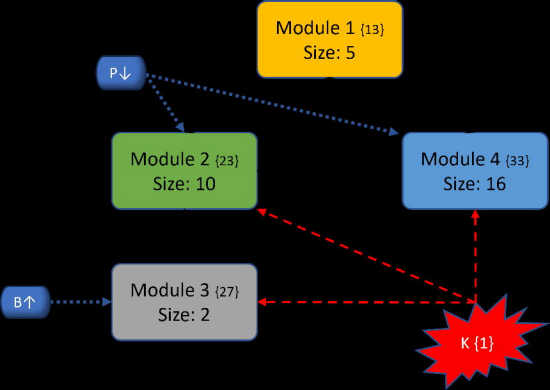
-0.80	-0.78	0.08	-0.87	-0.82	0.07	-0.84	-0.78	0.04	0.09	-0.83	0.06	0.07	-0.83	0.04	0.04
-0.68	-0.60	0.28	-0.78	-0.67	0.20	-0.70	-0.63	0.18	0.21	-0.69	0.22	0.22	-0.71	0.15	0.18
-0.60	-0.51	0.24	-0.76	-0.65	0.30	-0.74	-0.47	0.16	0.41	-0.71	0.21	0.31	-0.71	0.20	0.15
-0.70	-0.63	0.20	-0.80	-0.71	0.19	-0.76	-0.62	0.12	0.24	-0.75	0.16	0.20	-0.75	0.13	0.13

BAX↑

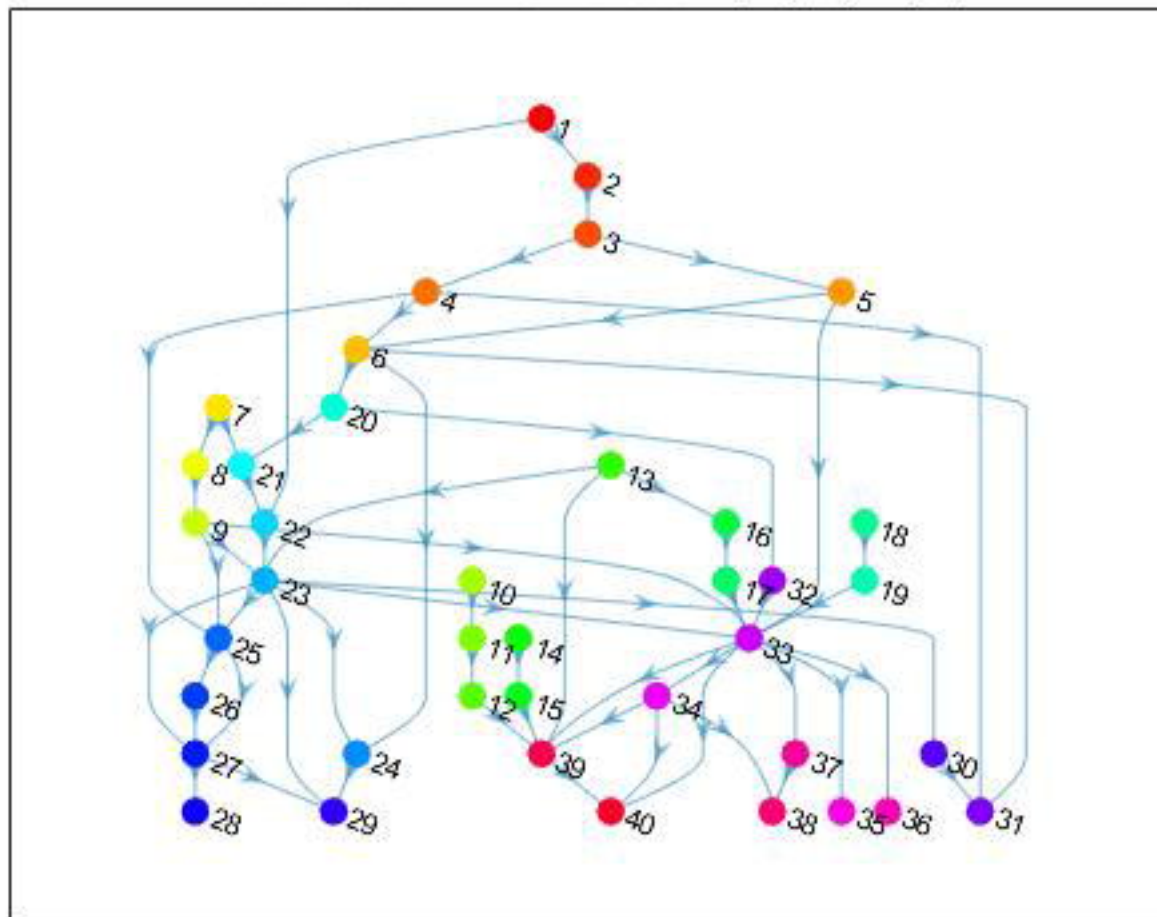
1.06	1.11	2.98	1.51	0.69	3.08	1.58	0.60	1.94	3.74	1.53	1.95	3.75	1.58	1.94	1.92
5.23	5.36	10.70	8.04	3.60	10.85	8.26	3.34	9.70	11.43	8.13	9.70	11.44	8.28	9.73	9.60
2.74	2.85	7.33	1.64	2.74	7.77	1.71	2.51	2.06	11.16	1.62	2.09	11.24	1.73	2.02	2.11
3.01	3.11	7.00	3.73	2.34	7.23	3.85	2.15	4.57	8.78	3.76	4.58	8.81	3.86	4.57	4.55

PI3K↓ / BAX↑

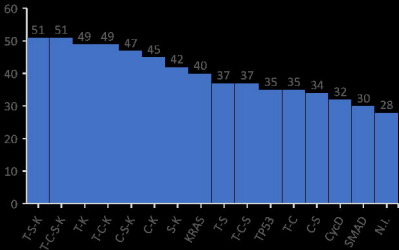
-0.86	-0.86	-0.88	-0.90	-0.86	-0.91	-0.88	-0.91	-0.91	-0.88	-0.87	-0.89	-0.87	-0.89	-0.90	-0.91
-0.65	-0.65	-0.75	-0.68	-0.60	-0.82	-0.66	-0.76	-0.75	-0.72	-0.63	-0.73	-0.69	-0.70	-0.80	-0.76
-0.65	-0.62	-0.67	-0.80	-0.66	-0.76	-0.75	-0.76	-0.80	-0.65	-0.71	-0.73	-0.64	-0.76	-0.74	-0.82
-0.72	-0.71	-0.77	-0.80	-0.71	-0.83	-0.76	-0.81	-0.82	-0.75	-0.74	-0.78	-0.73	-0.78	-0.82	-0.83



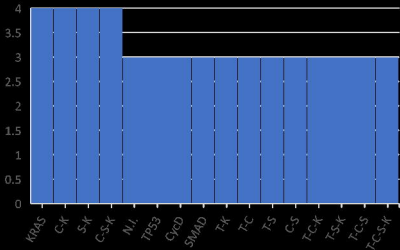
Modules for KRAS Mutation [13,23,27,33]



Total Modules



Non-Trivial Modules



Module 1 {8}
Size: 3

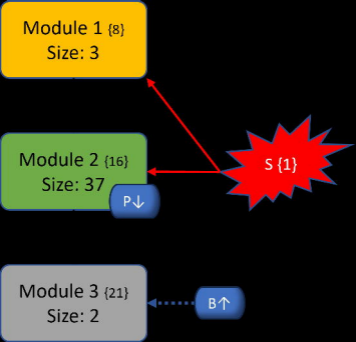
Module 2 {16}
Size: 37

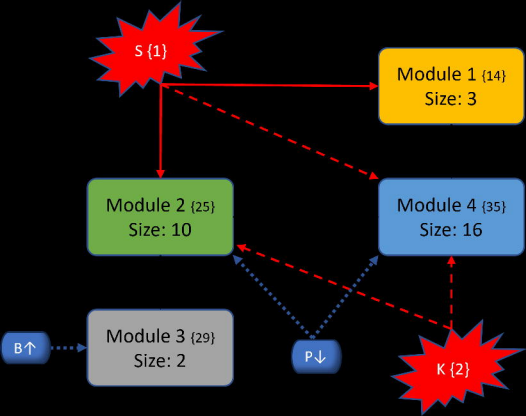
Module 3 {21}
Size: 2

P↓

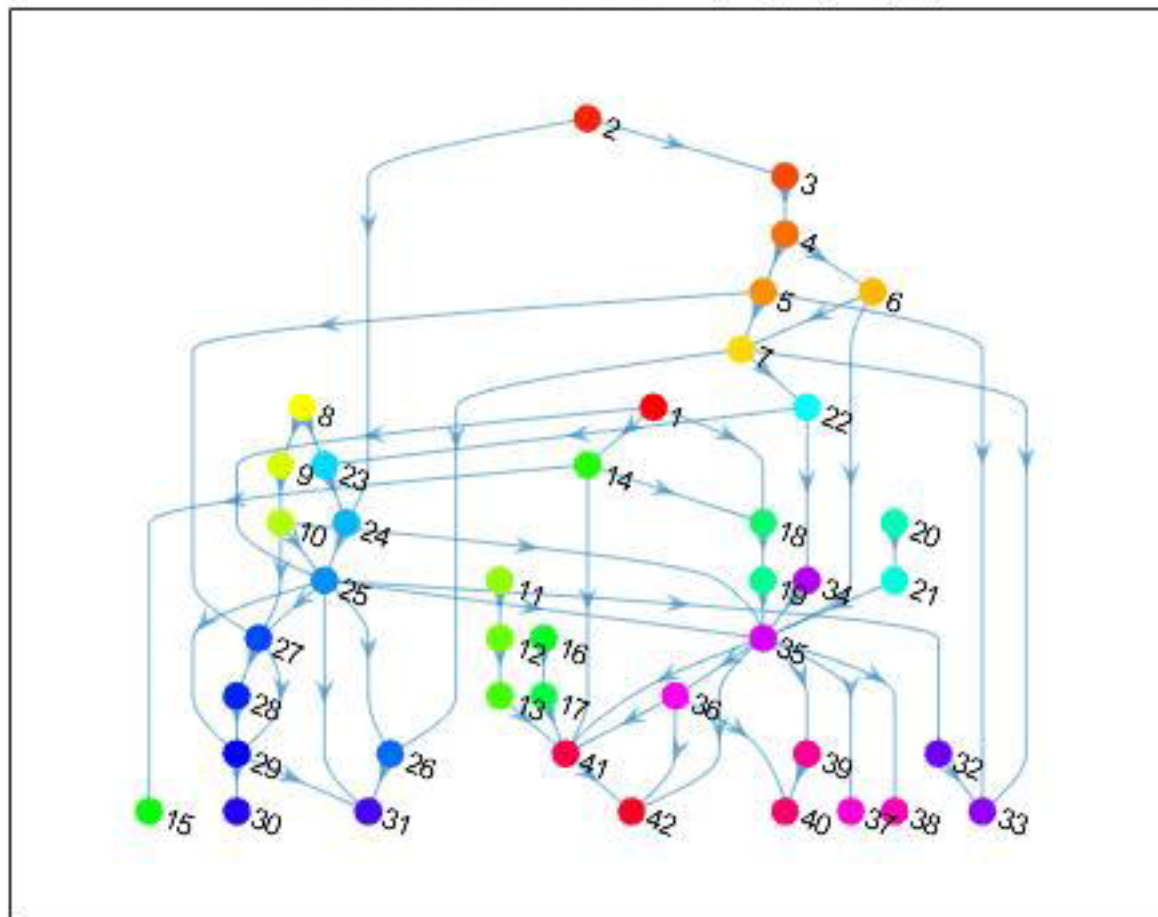
B↑

S {1}

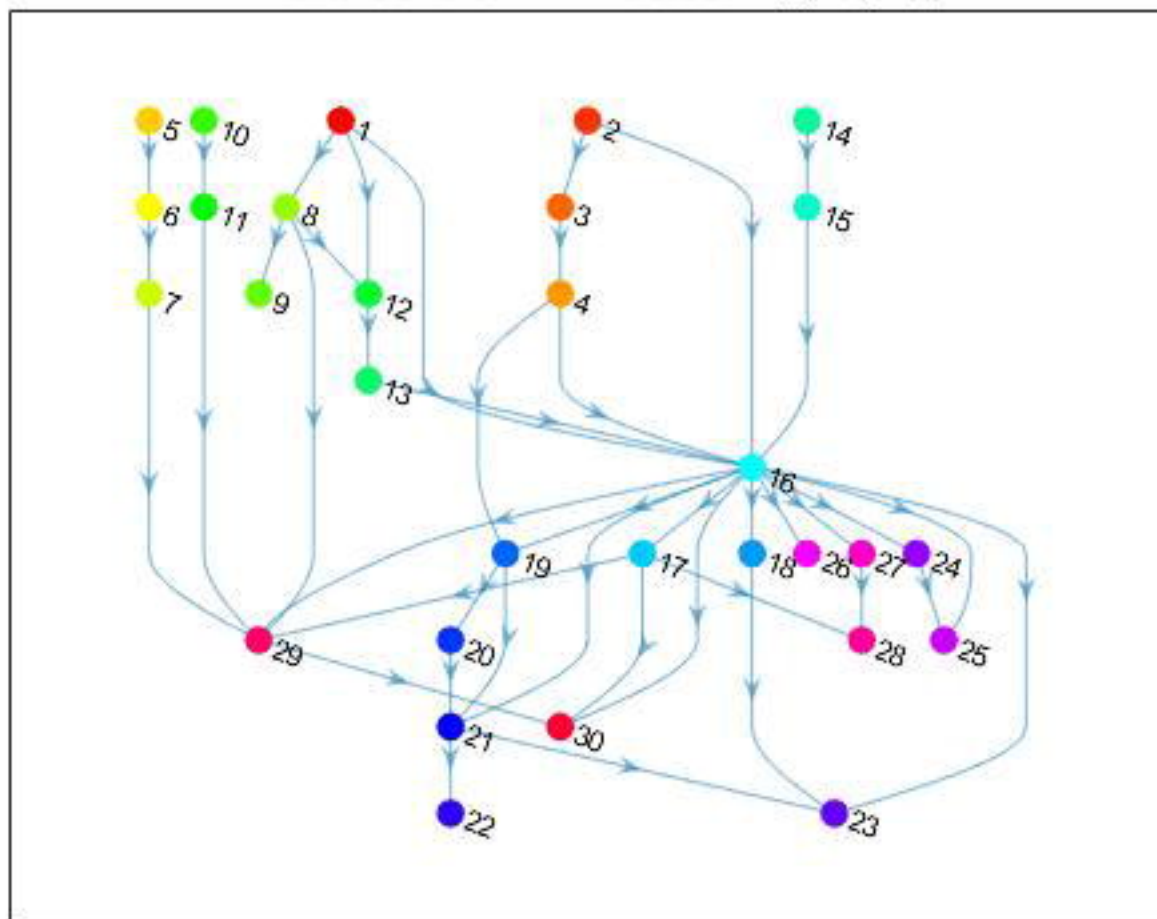




Modules for S-K Mutation [14,25,29,35]



Modules for SMAD Mutation [8,16,21]



Module 1 {9}
Size: 5

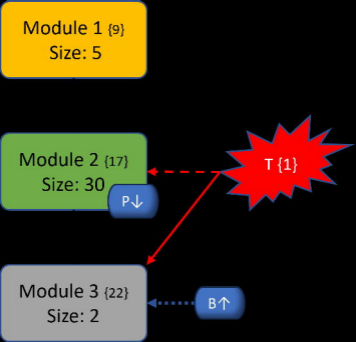
Module 2 {17}
Size: 30

Module 3 {22}
Size: 2

P↓

B↑

T {1}



Module 1 {12}
Size: 5

C {2}

Module 2 {21}
Size: 30

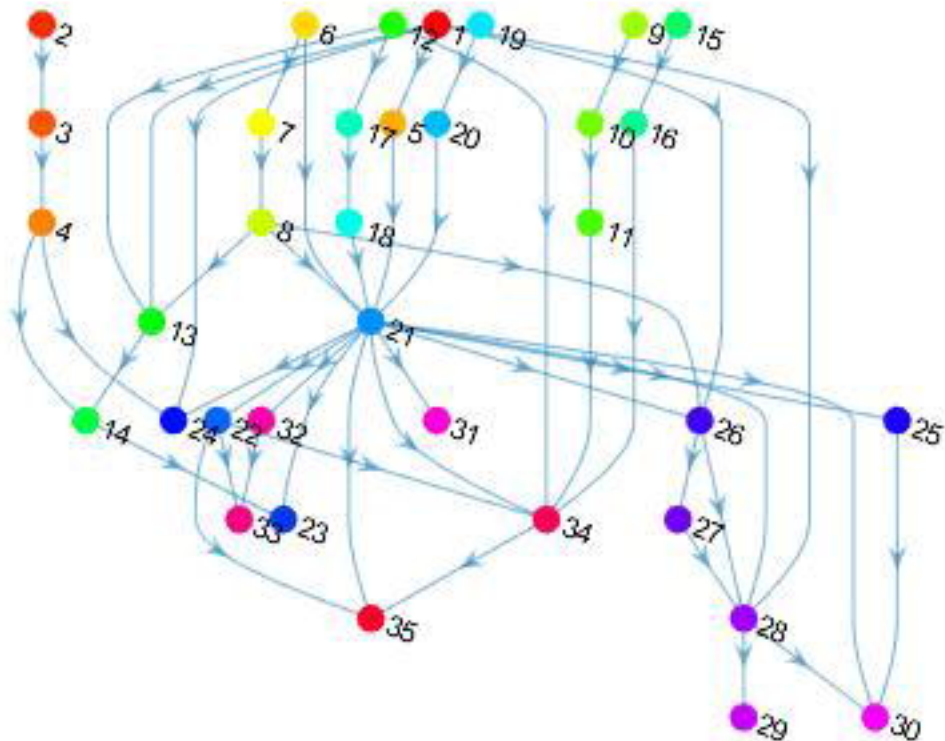
T {1}

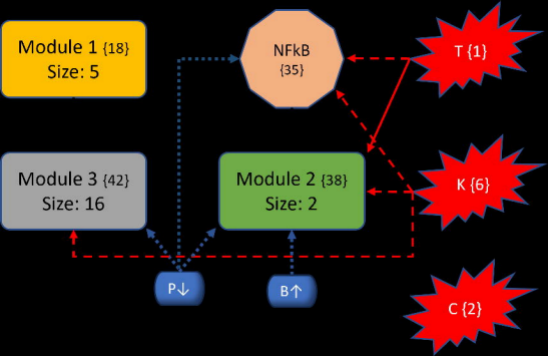
Module 3 {28}
Size: 2

P↓

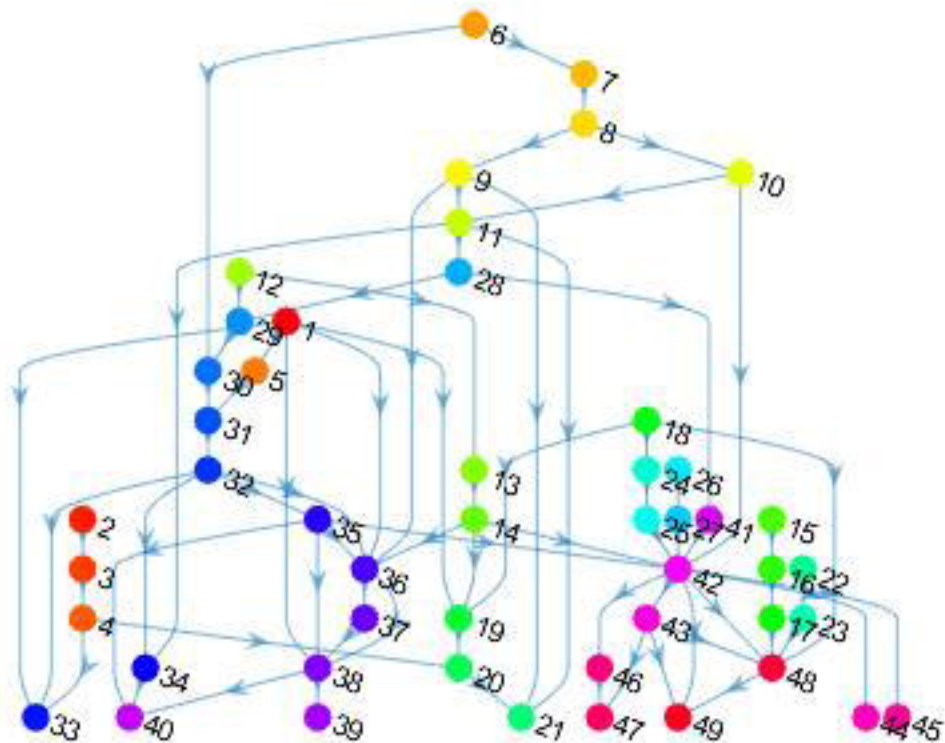
B↑

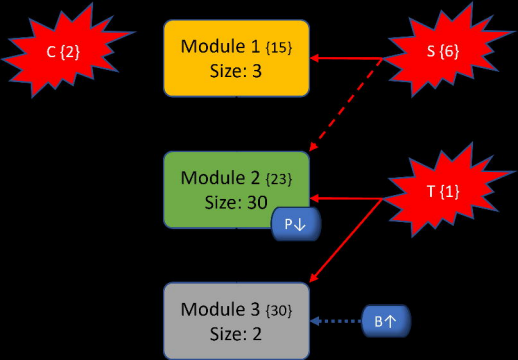
Modules for T-C Mutation [12,21,28]



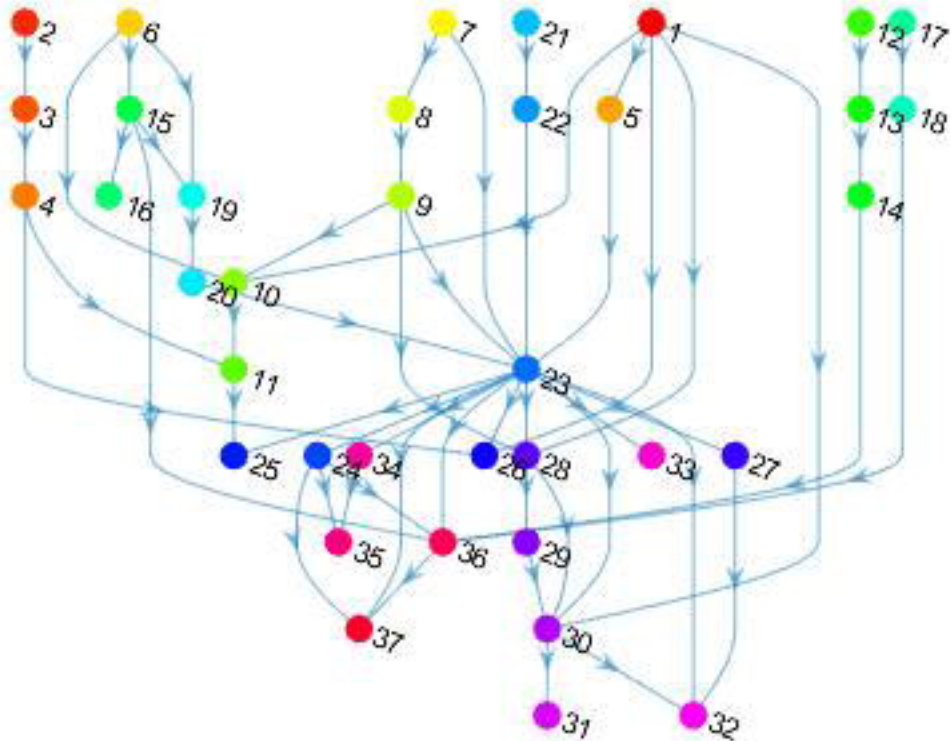


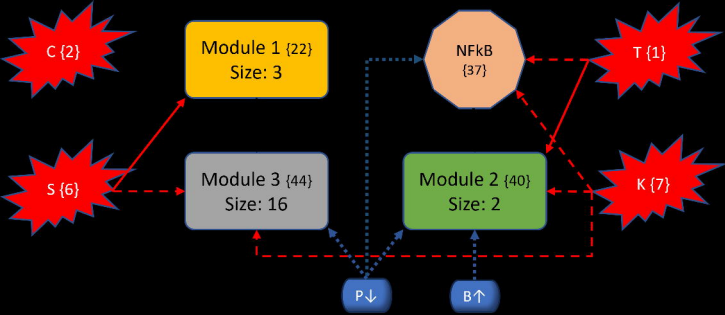
Modules for T-C-K Mutation [18,38,42]



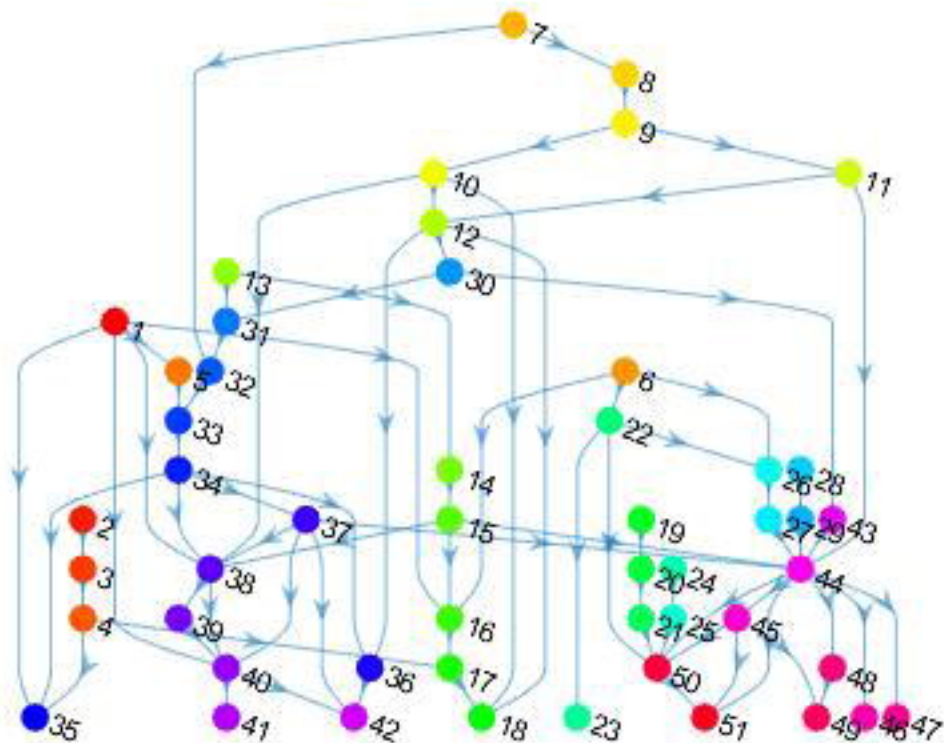


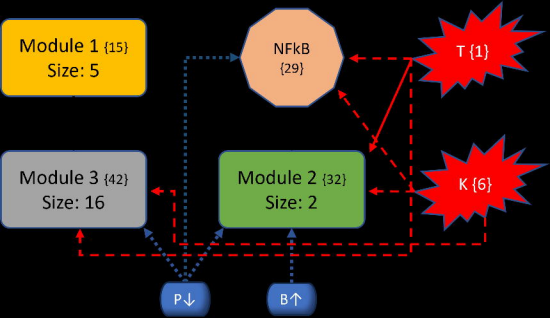
Modules for T-C-S Mutation [15,23,30]



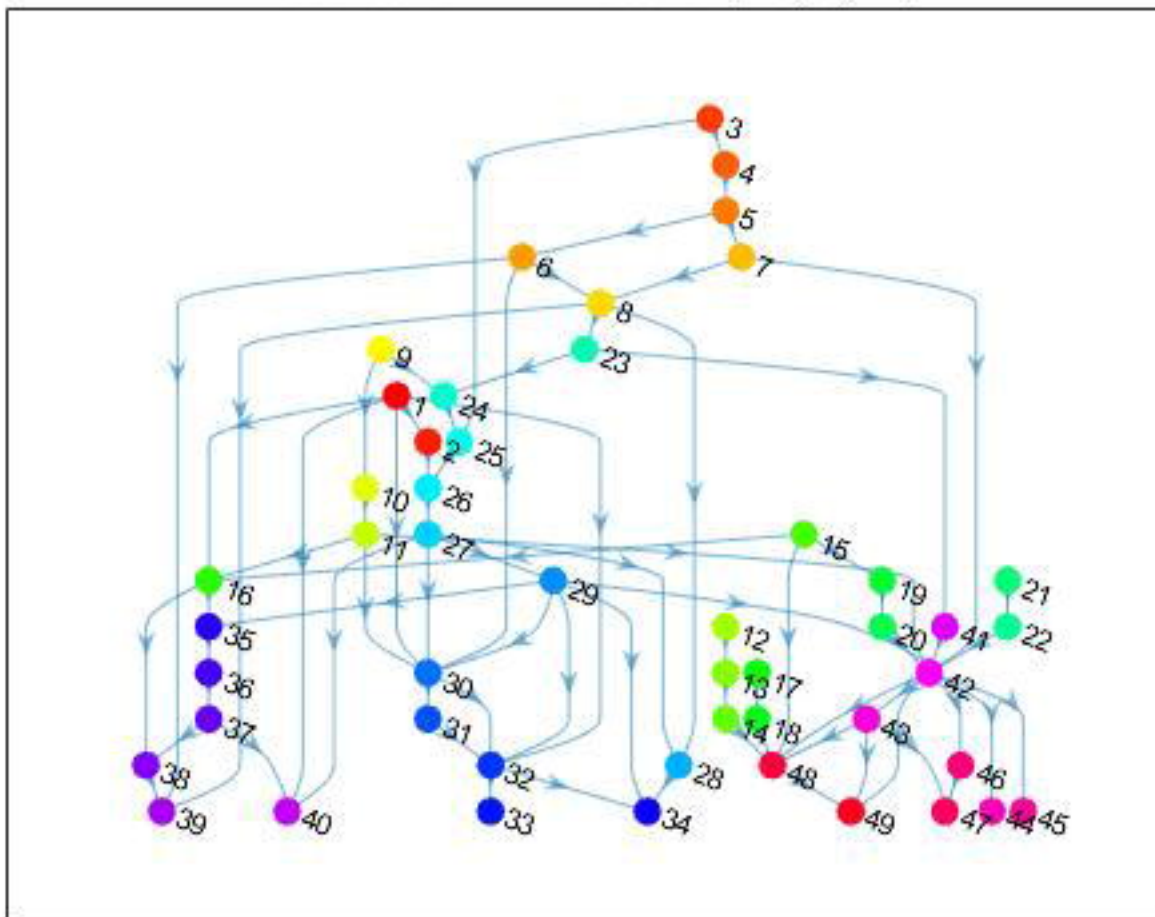


Modules for T-C-S-K Mutation [22,40,44]

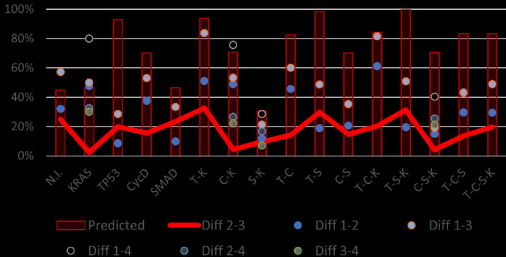




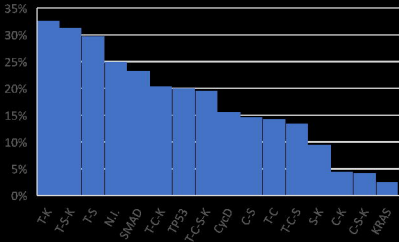
Modules for T-K Mutation [15,32,42]



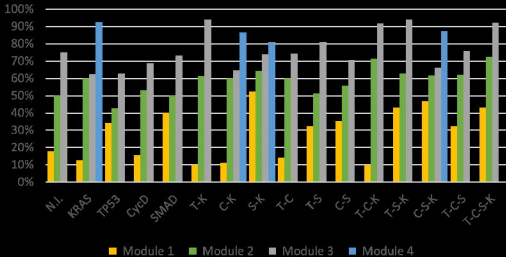
Toposort Gaps



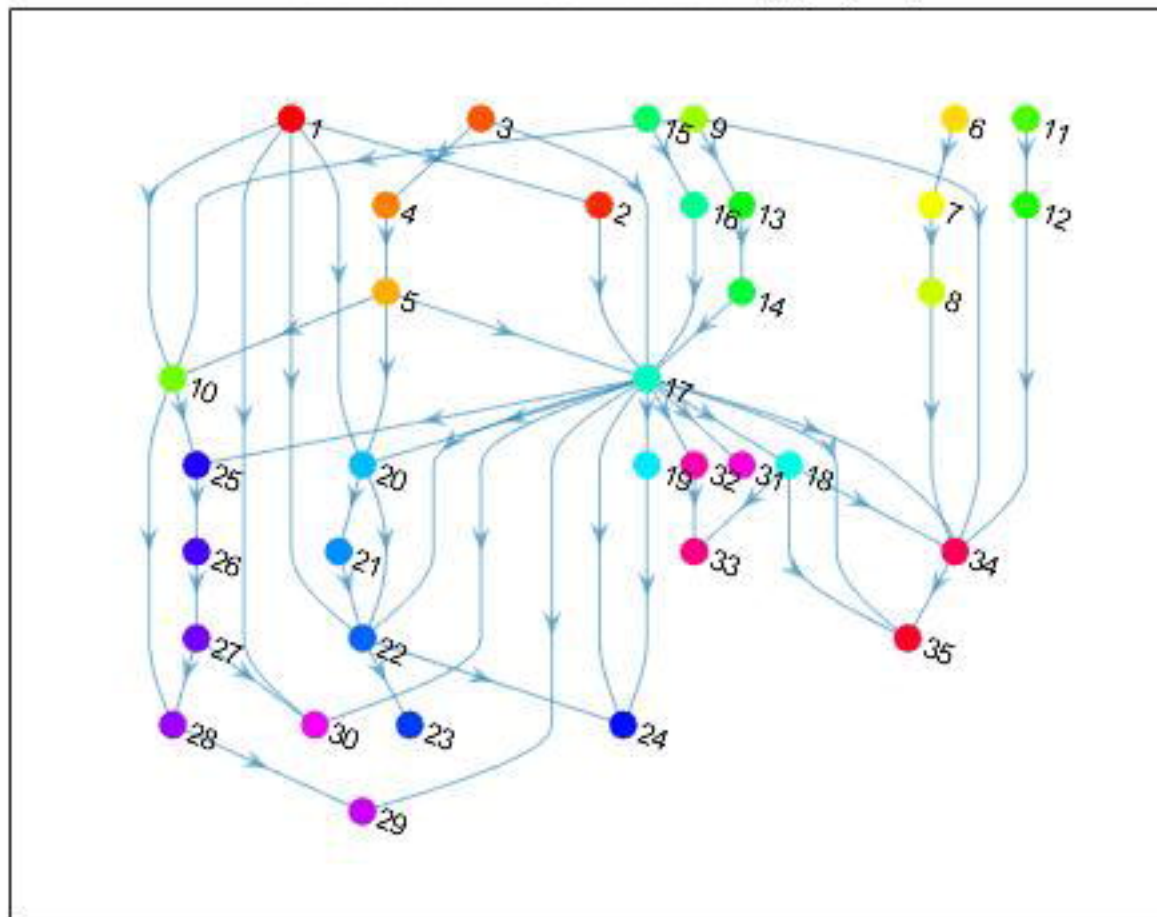
Toposort Margins Between Modules 2 and 3 Ranks

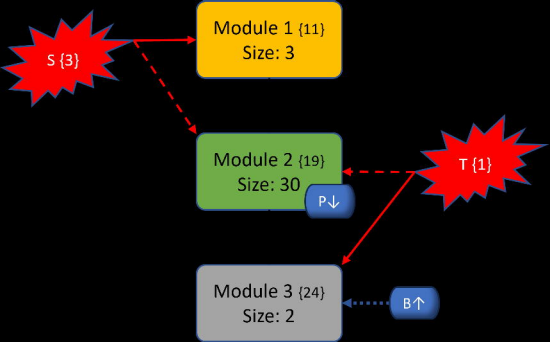


Topological Module Rankings

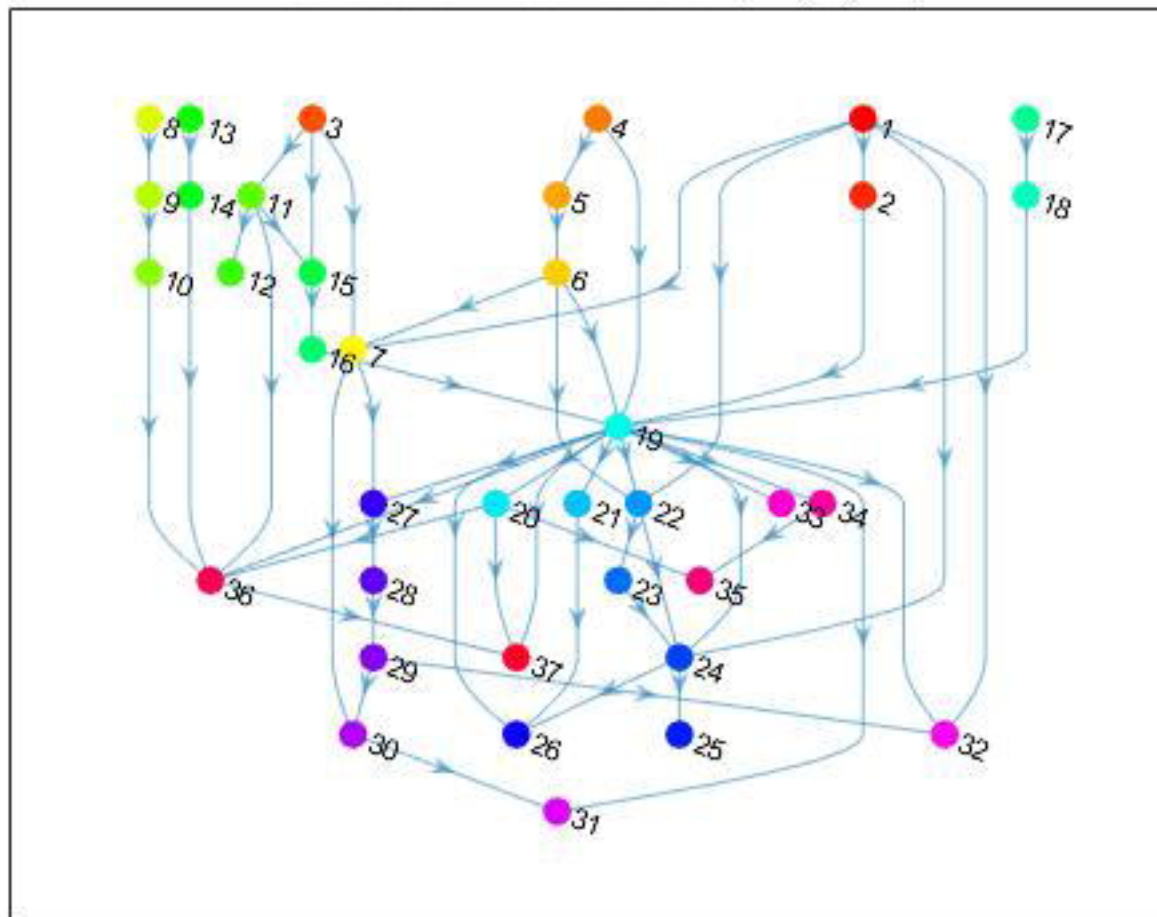


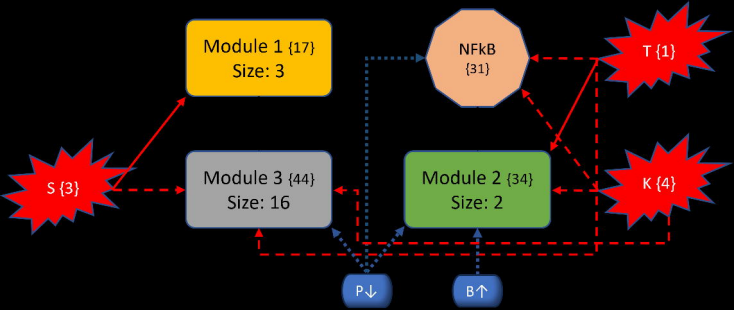
Modules for TP53 Mutation [9,17,22]



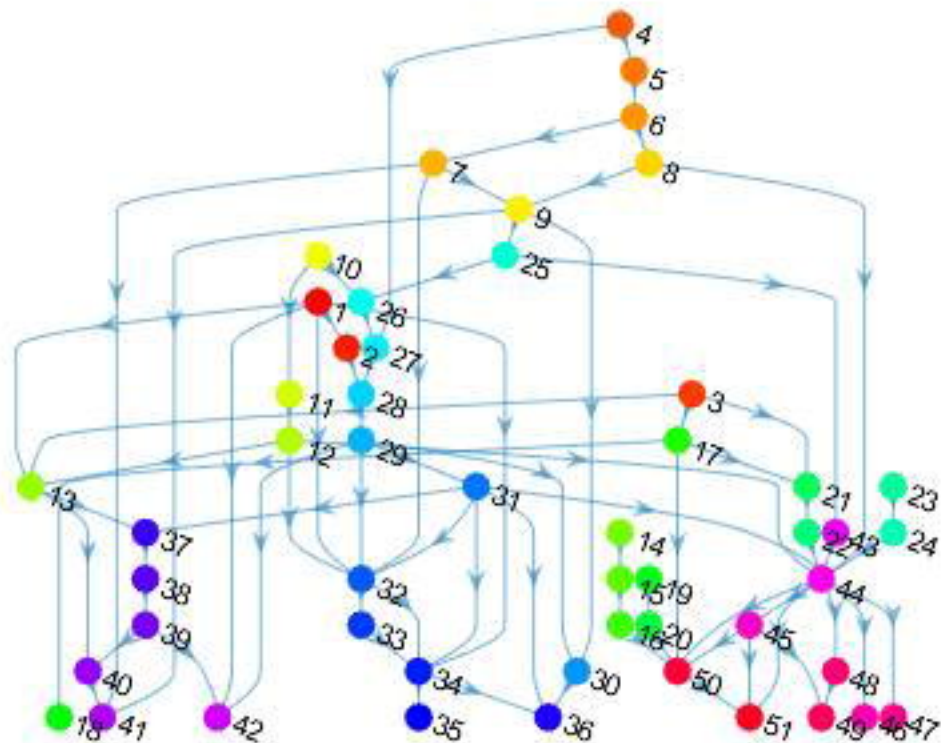


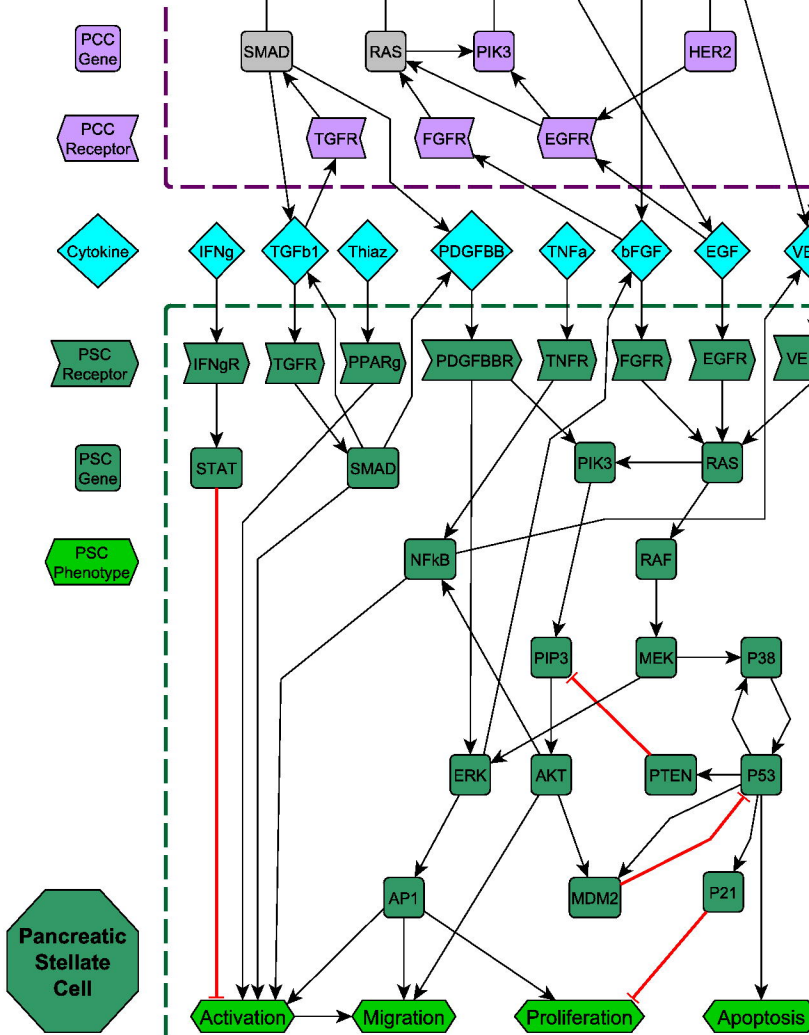
Modules for T-S Mutation [11,19,24]



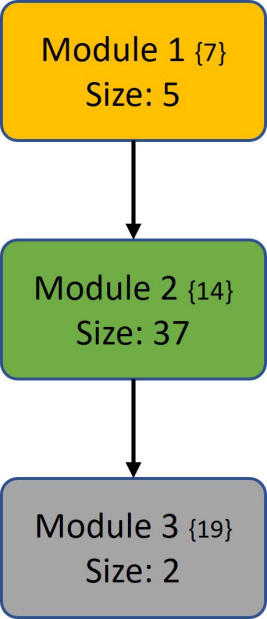


Modules for T-S-K Mutation [17,34,44]





Module 1 {7}
Size: 5

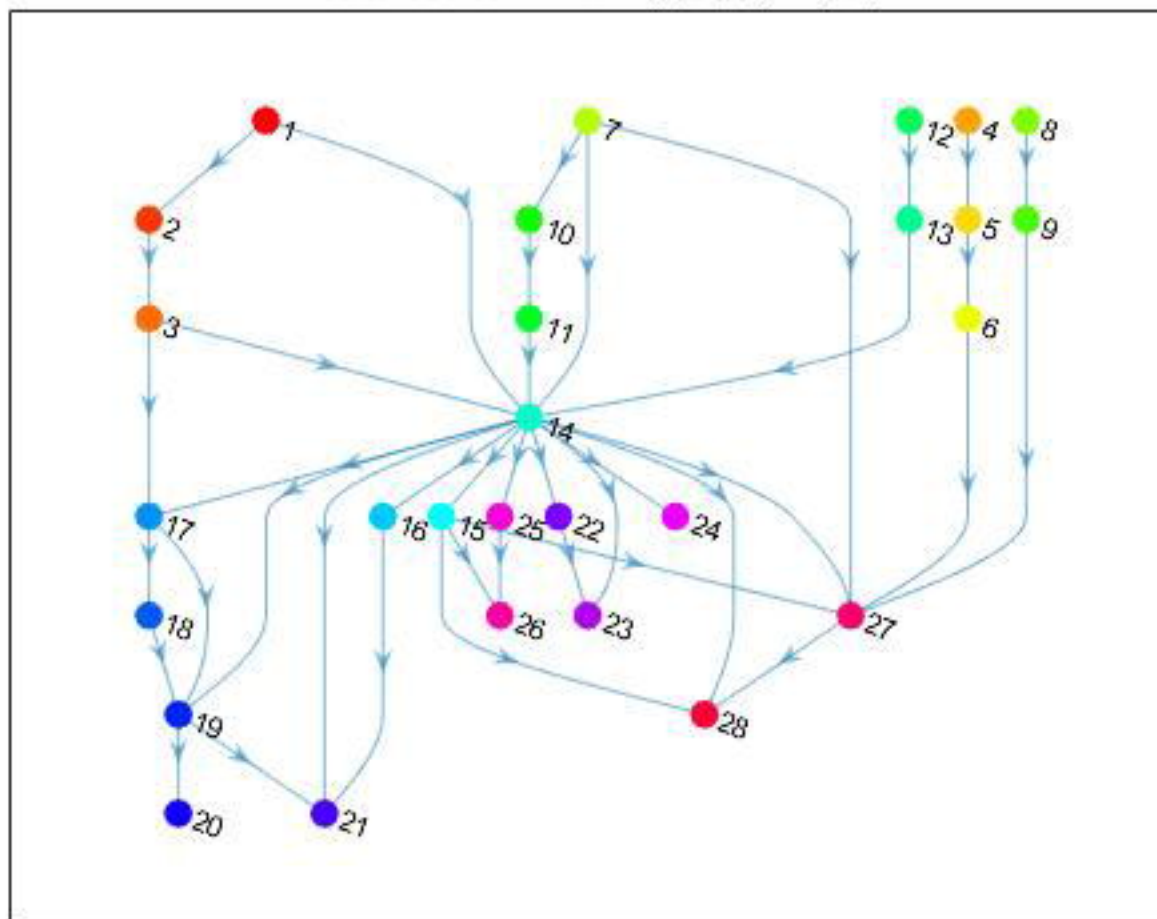


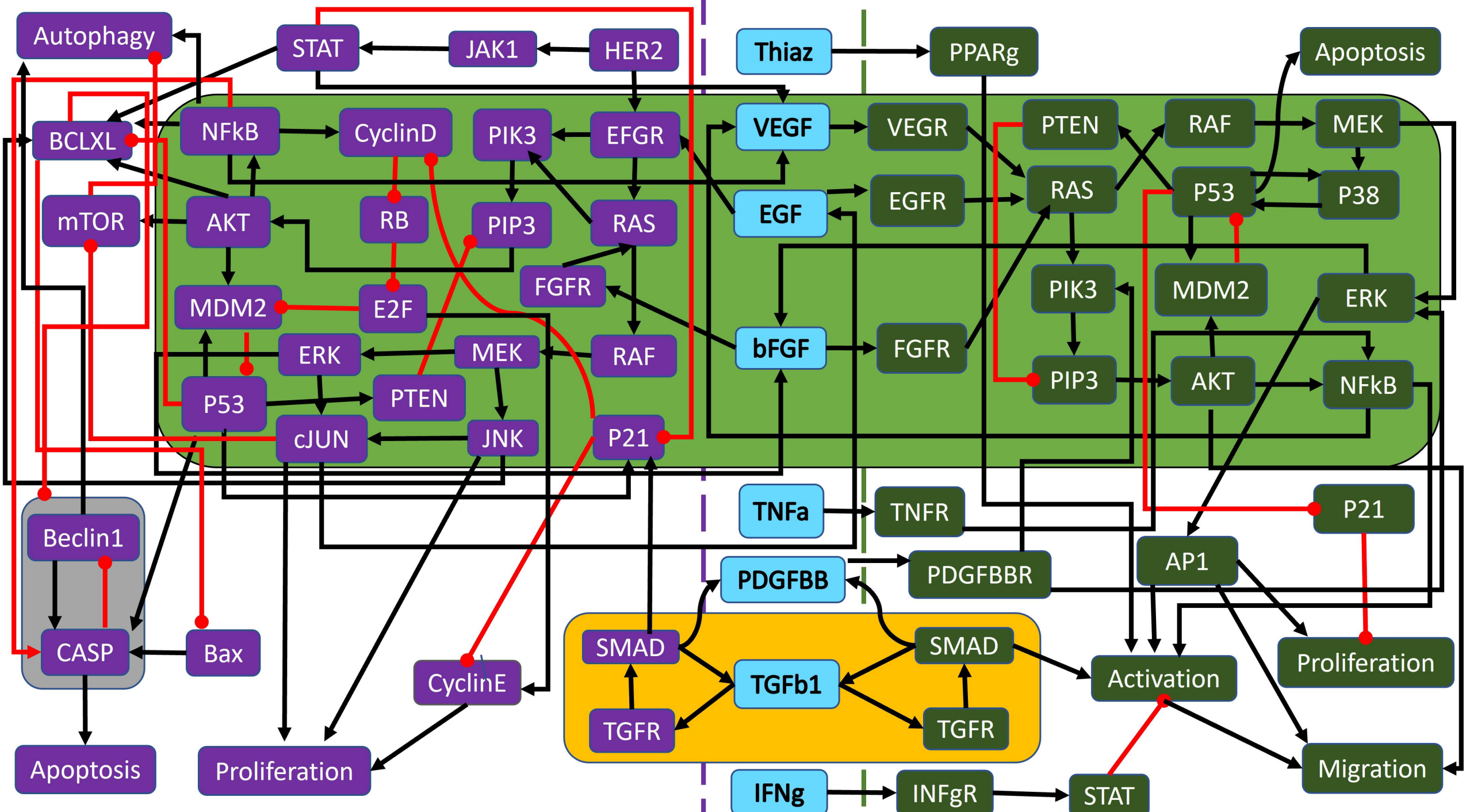
```
graph TD; M1[Module 1 {7} Size: 5] --> M2[Module 2 {14} Size: 37]; M2 --> M3[Module 3 {19} Size: 2];
```

Module 2 {14}
Size: 37

Module 3 {19}
Size: 2

Modules for Wild-Type [7,14,19]





Module 1 {7}
Size: 5

Module 2 {14}
Size: 37

P↓

Module 3 {19}
Size: 2

B↑