Algebraic and Topological Indices of Molecular Pathway Networks in Human Cancers

Peter Hinow∗ Edward A. Rietman† Jack A. Tuszyński‡

May 8, 2014

Abstract
Protein-protein interaction networks associated with diseases have gained prominence as an area of research. We investigate algebraic and topological indices for protein-protein interaction networks of 11 human cancers derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. We find a strong correlation between relative automorphism group sizes and topological network complexities on the one hand and five year survival probabilities on the other hand. Moreover, we identify several protein families (e.g. PIK, ITG, AKT families) that are repeated motifs in many of the cancer pathways. Interestingly, these sources of symmetry are often central rather than peripheral. Our results can aide in identification of promising targets for anti-cancer drugs. Beyond that, we provide a unifying framework to study protein-protein interaction networks of families of related diseases (e.g. neurodegenerative diseases, viral diseases, substance abuse disorders).

Introduction
Biological networks have been an active area of research for some years, see e.g. [12, 16, 17] and the references therein. In earlier work [6] we reported that molecular signaling network complexity is correlated with cancer patient survival. In that work we reported a statistical mechanics measure of network complexity. Here we focus on the relative sizes of automorphism groups and the dimensions of the cycle spaces (cyclomatic numbers).

Complex real-world networks contain feedback loops to enable the network “communication” to continue in the face of node failure [2]. In the case of

∗Department of Mathematical Sciences, University of Wisconsin - Milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA; hinow@uwm.edu
†Newman-Lakka Institute, Tufts University School of Medicine, Boston, MA 02111, USA; erietman@gmail.com
‡Cross Cancer Institute and Department of Physics, University of Alberta, Edmonton, T6G 2E1, Canada; jack.tuszynski@gmail.com
protein-protein interaction (PPI) networks this means that inhibition of a spec-
cific node may or may not have any effect. It is well known that targeting hub
nodes in networks often causes the network to break up into multiple com-
ponents and this could be lethal, because many protein hubs in PPIs for cancer
are also important proteins in metabolic networks. As we argued in [6], tar-
geting nodes with high-betwenness has higher potential for improved cancer
treatment. Selective targeting of nodes in a PPI for cancer treatment is fraught
with difficulties.

In this letter we apply two more algebraic and topological indices to study
cancer PPI networks and show correlation with 5 year patient survival. We
identify several repeated motifs of proteins that are “interchangeable” in a sense
to be specified below. In the long run we anticipate that the methods described
here will aid identification of potential drug targets.

**Results and Discussion**

A network is an undirected graph $G = (V, E)$ with vertex set $V$ and edge set
$E$. The vertices are proteins and two vertices are connected by an edge if there
is a known interaction of the two partners, either by direct binding or by enzy-
matic catalysis. Beyond cancer pathways, the Kyoto Encyclopedia of Genes and
Genomes (KEGG) database also contains pathways related to immune diseases
(e.g. asthma), neurodegenerative diseases (Alzheimer’s disease, Parkinson’s dis-
ease), substance dependence, cardiovascular diseases, viral diseases and many
others [13]. The KEGG networks are assembled from the literature by searches
for experimental confirmation of the relevant interactions. Each interaction
is always confirmed by two or more different experimental techniques such as
pull-down mass spectrometry, yeast two-hybrid and various biochemical tests.
Naturally, networks constructed from experimental results are likely to contain
errors, which are however impossible to quantify.

An automorphism is a permutation $\phi : V \to V$ that preserves the adjacency
relation, that is,

$$(u, v) \in E \iff (\phi(u), \phi(v)) \in E.$$  

With the operation of composition, the automorphisms form a group $\text{Aut}(G)$. The relation on the set of vertices

$$u \sim v \iff \text{there exists a } \phi \in \text{Aut}(G) \text{ such that } v = \phi(u)$$

is an equivalence relation and its equivalence classes are called the (group) orbits.

MacArthur et al. [18] list 20 examples of real world networks and their rich
symmetry groups. This is in contrast to large random graphs, such as graphs
from the Erdős-Rényi model $\mathcal{G}(n, p)$. Here $n$ is the number of vertices. Edges are
independently present with probability $0 < p < 1$. Such graphs have only the
trivial automorphism, with probability approaching one, in the limit $n \to \infty$ [5,
Chapter IX]. The difference is not surprising if one realizes that real networks
display a modular structure, with vertices organized in communities tightly
connected internally and loosely connected to each other \[^{10}\]. This results in the presence of symmetric subgraphs such as trees and complete cliques.

Figure 1 shows as an example the protein-protein interaction network of pancreatic cancer as retrieved from the KEGG database. We find that the automorphism group of this network is the direct product of symmetric groups 

\[ \text{Aut}(G) = S_9^2 \times S_6^3 \times S_5^2 \times S_8, \]

see Table 1 for a complete list of automorphism groups. Remarkably, symmetries do not only arise due to tree subgraphs at the “ends” of the network, but also due to central nodes of high degree (highlighted in yellow in Figure 1). Thus any flow of information that passes through one node in such an orbit equivalence class may pass through any other node in the same equivalence class. The presence of such modular patterns indicates a high level of redundancy which confers robustness to the associated biological system (tumor cells). We suggest that to interrupt the flow through such a network most efficiently, the nodes adjacent to large central orbits are the best to be targeted for example by pharmacological agents that inhibit a specific protein-protein interaction pair. Similar suggestions have been made in\[^{7, 8, 22}\]. The use of automorphism groups has, to the best of our knowledge, not yet been proposed.

Automorphism groups are often used to measure the complexity of a network \[^{23}\]. In order to make automorphism group sizes of graphs with \( n = |V| \) vertices comparable, we follow the suggestion in \[^{23}\] and compute the ratio

\[ \beta_G = \left( \frac{|\text{Aut}(G)|}{n!} \right)^{\frac{1}{n}}. \]

This relates the size of \( \text{Aut}(G) \) to the size of the automorphism group \( S_n \) of the complete graph on \( n \) vertices.

A second graph invariant is of a more topological nature. A cycle is a sequence of adjacent vertices that starts and ends at the same vertex. The set of all cycles \( C(G) \) can be made a vector space over the field \( \mathbb{Z}_2 \) by taking the symmetric difference

\[ C_1 \Delta C_2 = (C_1 \cup C_2) \setminus (C_1 \cap C_2) \]

as addition, the identity as negation, and the empty cycle as zero. The dimension of this vector space is called the cyclomatic number \( \mu_G \), or the circuit rank. Loosely speaking, it is a count of the “independent” loops, see Figure 2. It is shown in \[^{4, 15}\] that for a graph with \( n \) vertices, \( m \) edges and \( c \) connected components, \( \mu_G \) is given by

\[ \mu_G = m - n + c. \]

We plot these two indices against the five year survival probability \( p \), obtained from the Surveillance, Epidemiology and End Results (SEER) database \[^{20}\] for 11 types of cancer in Figure 3. Interaction networks with larger values of
\( \beta_G \) or equivalently greater symmetry are associated with better chances of survival. A large value of \( \mu_G \) indicates high topological complexity and correlates with decreased chance of survival. We find that both coefficients of determination are \( R^2 = 0.52 \) with corresponding \( p \)-value \( p = 0.011 \) (the equality is coincidental). There are widespread differences in detection stage, metastasis status, treatment and general health of the patient which are unfortunately not accessible from the SEER database. Nevertheless, given this large amount of natural uncertainty in the data, this indicates a strong correlation of averages. It would be invaluable for future research to classify database entries according to some of the parameters mentioned above. Since both the automorphism group size \( \beta_G \) and the cyclomatic number \( \mu_G \) are correlated to the five year survival probability \( p \), it is to be expected that these two quantities are correlated to each other, for the the protein-protein interaction networks of cancers that are the object of our study, see Figure 4 left panel. However, it is easy to construct examples of graphs that show no correlation between \( \beta_G \) and \( \mu_G \), see Figure 4 right panel.

Further study of the automorphism groups reveals repeated motifs in several interaction networks. The eight proteins from the PIK3C\{A,B,D,G\} and PIK3R\{1,2,3,5\} family form a single orbit equivalence class in seven of the networks (AML, CML, colorectal, endometrial, pancreatic, renal and SCL cancers) and are split in two orbit equivalence class in two more networks (glioma and NSCL). The three proteins AKT\{1,2,3\} are orbit equivalent in eight networks (CML, colorectal, endometrial, glioma, NSCL, pancreatic, renal and SCL cancers), that is, whenever they appear in the network to begin with. These players have been known for a long time to be of crucial importance to the initiation and progression of cancer, mainly due to the various biological and biochemical assays performed on cancer cells. However, our conclusions stem directly from a group-theoretic analysis of the PPI networks and they are network specific. Since its initial discovery as a proto-oncogene, the serine/threonine kinase AKT has become a major focus of attention because of its critical regulatory role in diverse cellular processes, including cancer progression and insulin metabolism. The AKT cascade is activated by receptor tyrosine kinases, integrins, B and T cell receptors, cytokine receptors, G-protein-coupled receptors and other stimuli that induce the production of phosphatidylinositol (3,4,5)-triphosphates (PtdIns(3,4,5)P3) by phosphoinositide 3-kinase (PI3K). These lipids serve as plasma membrane docking sites for proteins that harbor pleckstrin-homology (PH) domains, including AKT and its upstream activator PDK1. The tumor suppressor PTEN is recognized as a major inhibitor of AKT and is frequently lost in human tumors. There are three highly related isoforms of AKT (AKT1, AKT2, and AKT3), which represent the major signaling arm of PI3K. For example, germline mutations of AKT have been identified in pathological conditions of cancer and insulin metabolism. AKT regulates cell growth through its effects on the TSC1/TSC2 complex and mTOR pathways, as well as cell cycle and cell proliferation through its direct action on the CDK inhibitors p21 and p27, and its indirect effect on the levels of cyclin D1 and p53. AKT is a major mediator of cell survival through direct inhibition of pro-apoptotic signals such as the
pro-apoptotic regulator BAD and the FOXO and Myc family of transcription factors. AKT has been demonstrated to interact with Smad molecules to regulate TGF-β signaling. These findings make AKT an important therapeutic target for the treatment of cancer.

Interestingly, the network of small cell lung cancer contains an enormous orbit of 18 equivalent nodes of degree six. This orbit consists of laminines, collagens and a fibronectin that are major proteins in the basal lamina. All nodes are connected to six members of the integrin family of transmembrane receptors.

**Conclusion**

We have shown that the relative size of the automorphism groups and the cyclomatic numbers for cancer pathway networks from the KEGG database are both correlated with five-year survival of cancer patients. Determination of the specific reasons for these great discrepancies in survival rates remains a topic for future research. Interestingly, cancers with more symmetric interaction networks are associated with better survival rates. This may be due to a greater robustness to failure, which, somewhat counterintuitive, is a positive feature in this context.

We suggest that selective removal of nodes from the network (clinically equivalent to protein inhibition) and reinterpolation on the linear curves helps to identify potential drug targets. This indicates that complexity of a biochemical network involved in a deregulated cell cycle as exemplified by cancer cells is of crucial importance to its robustness. This is manifested by various redundancies in the PPI network that make the search for a therapeutic “silver bullet” an impossible task. We suggest that selective removal of nodes from the network (clinically equivalent to protein inhibition) and reinterpolation on the linear curves helps to identify potential drug targets. We have shown that PI3K and AKT families of proteins appear to be the most suitable targets for pharmacological inhibition in the most number of cancer types studied. It is encouraging that there are several AKT pathway inhibitors in clinical development, e.g. perifosine (KRX-0401, Aeterna Zentaris/Keryx), MK-2206 (Merck), and GSK-2141795 (Glaxo-SmithKline) [3]. Similarly, Bayer, GlaxoSmithKline (GSK), Novartis, Merck & Co., Roche and Sanofi are just a few of the companies that have placed great importance on the development of a spectrum of agents targeting the PI3K pathway. Drug candidates including pan-PI3K inhibitors, PI3K isoform-specific inhibitors, AKT inhibitors and mTOR inhibitors are currently tested alone and in combinations in an array of cancer indications [11].

While the motivation for this focus has been stated as: The pathway is almost invariably on in cancer, our methodology identifies this pathway as the most crucial using mathematical analysis of the network. Moreover, we are able to identify those types of cancer where the pathway should be the main target and those types where targeting it may not produce the expected clinical outcomes.
Methods

The cancer pathways were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG) [13] with the help of the open source software packages KEGGgraph [24] and cytoscape [21]. The automorphism groups of the networks were found with saucy [14] and gap [9] (see Tables 2-12 for the complete group lists). Bases of the cycle spaces were found using python networkX.

Acknowledgements

JAT acknowledges funding support for this research from NSERC (Canada), Canadian Breast Cancer Foundation and the Allard Foundation. We thank Dr. Giannoula Klement (Newman-Lakka Institute, Tufts University School of Medicine) for valuable comments.

References

[1] Ahuja RK, Magnanti TL, Orlin JB, Network Flows. Theory, Algorithms, and Applications, Prentice Hall, Englewood Cliffs, NJ, 1993
[2] Albert R, Jeong H, Barabási AL, Error and attack tolerance of complex networks, Nature 406, 378-382, (2000)
[3] Alexander W, Inhibiting the Akt pathway in cancer treatment. Three leading candidates, Pharmacy and Therapeutics 36, 225-227, (2011)
[4] Berger F, Gritzmann P, de Vries S, Minimum cycle bases for network graphs, Algorithmica 40, 51-62, (2004)
[5] Bollobás B, Random Graphs, Cambridge University Press, Cambridge, 2001
[6] Breitkreutz D, Hlatky L, Rietman E, Tuszynski JA, Molecular signaling network complexity is correlated with cancer patient survivability, Proc. Natl. Acad. Sci. USA 109, 9209-9212, (2012)
[7] Chandra N, Padiadpu J, Network approaches to drug discovery, Expert Opin. Drug Discov. 8, 7-20, (2013)
[8] Csermely P, Kocsmáros T, Kiss HJM, London G, Nussinov R, Structure and dynamics of molecular networks: A novel paradigm of drug discovery. A comprehensive review, Pharmacol. Therapeut. 138, 333-408, (2013), arXiv:1210.0330
[9] The GAP Group, GAP – Groups, Algorithms, and Programming, Version 4.6.4, 2013, http://www.gap-system.org
[10] Garlaschelli D, Ruzzentini F, Basosi R, Complex networks and symmetry I: A review, *Symmetry* 2, 1683-1709, (2010), [arXiv:1006.3923](http://arxiv.org/abs/1006.3923)

[11] Holmes D, PI3K Pathway inhibitors approach junction, *Nat. Rev. Drug Discov.* 10, 563-564, (2011)

[12] Junker BH, Schreiber F (eds.), *Analysis of Biological Networks*, John Wiley & Sons, Hoboken, NJ, 2008

[13] Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M, The KEGG resource for deciphering the genome, *Nucleic Acids Res.* 32, D277-D280, (2004), [http://www.genome.jp/kegg/](http://www.genome.jp/kegg/)

[14] Katebi H, Sakallah KA, Markov IL, Graph Symmetry Detection and Canonical Labeling: Differences and Synergies, In: Turing-100. The Alan Turing Centenary, A Voronkov (ed.), p. 181-195, Manchester, UK, (2012), [http://vlsicad.eecs.umich.edu/BK/SAUCY/](http://vlsicad.eecs.umich.edu/BK/SAUCY/)

[15] Kavitha T, Liebchen C, Mehlhorn K, Michail D, Rizzie R, Ueckerdt T, Zweig KA, Cycle bases in graphs characterization, algorithms, complexity, and applications, *Computer Science Review* 3, 199-243, (2009)

[16] Koonin EV, Wolf YI, Karev GP (eds.), *Power Laws, Scale-Free Networks and Genome Biology*, Landes Bioscience, Austin, TX, 2006

[17] Ma’ayan A, MacArthur BD (eds.), *New Frontiers of Network Analysis in Systems Biology*, Springer Verlag, Dordrecht, Heidelberg, New York, London, 2012

[18] MacArthur BD, Sánchez-García RJ, Anderson JW, Symmetry in complex networks, *Discr. Appl. Math.* 156, 3525-3531, (2008), [arXiv:0705.3215](http://arxiv.org/abs/0705.3215)

[19] Newman MEJ, *Networks: An Introduction*, Oxford University Press, Oxford, New York, 2010

[20] National Cancer Institute, Surveillance, Epidemiology and End Results (SEER) Program, 2013, [http://seer.cancer.gov/](http://seer.cancer.gov/)

[21] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res.* 13, 2498-2504, (2003), [http://cytoscape.org](http://cytoscape.org)

[22] Winterbach W, van Mieghem P, Reinders M, Wang H, de Ridder D, Topology of molecular interaction networks, *BMC Syst. Biol.* 7, 90, (2013)

[23] Xiao Y, MacArthur BD, Wang H, Xiong M, Wang W, Network quotients: Structural skeletons of complex systems, *Phys. Rev. E* 78, 046102, (2008), [arXiv:0802.4318](http://arxiv.org/abs/0802.4318)
[24] Zhang JD, Wiemann S, KEGGgraph: a graph approach to KEGG PATHWAY in R and Bioconductor, *Bioinformatics* 25, 1470-1471, (2009), http://bioconductor.org

Figures

Figure 1: *The protein-protein interaction network of pancreatic cancer.* The network was retrieved from the KEGG database [13] and its automorphism group determined with saucy [14], namely $\text{Aut}(G) = S_3^2 \times S_2^4 \times S_2^2 \times S_8$. Highlighted in yellow are three central orbits of nodes of degrees 3, 4 and 8, respectively. Two of these are the PI3K and the AKT families, respectively.

Figure 2: *The cycle space of a simple graph with four vertices.* The graph $G$ is shown in black and three cycles $C_1$, $C_2$ and $C_3$ are marked in red, blue and green, respectively. Any of the sets $\{C_1, C_2\}$, $\{C_1, C_3\}$ and $\{C_2, C_3\}$ is a basis of the cycle space $\mathcal{C}(G) = \{C_1, C_2, C_3\}$. 
Figure 3: *Automorphism group sizes and cyclomatic numbers.* A plot of relative automorphism group size $\beta_G$ (left) and cyclomatic number $\mu_G$ (right) against five year survival probability $p$ for 11 types of cancer. In both cases we have the coefficients of determination $R^2 = 0.52$ and $p = 0.011$. The cancer types are *AML:* acute myelogeneous leukemia, *Bld:* bladder cancer, *CML:* chronic myelogeneous leukemia, *Col:* colorectal cancer, *End:* endometrial cancer, *Gli:* Glioma, *NSCL:* non-small cell lung cancer, *Pc:* pancreatic cancer, *Ren:* renal cancer, *SCL:* small cell lung cancer, *Thyr:* thyroid cancer. The width of the horizontal error bars is 0.1, the width of the vertical error bars is 0.02 in the left panel and 20 in the right panel. These errors are estimated as the actual error is unknown.

Figure 4: *Mutual relation between $\beta_G$ and $\mu_G$. (Left)* A plot of cyclomatic number $\mu_G$ against relative automorphism group size $\beta_G$ for the 11 protein-protein interaction networks of cancers. (Right) In general, these quantities are not related. Shown are a highly symmetric graph with trivial cycle space and an asymmetric graph with large cycle space.

Tables
Table 1: Automorphism groups of all cancers. Column $n$ contains the number of vertices, column $c$ contains the number of connected components of the protein-protein interaction network. The abbreviations are the same as in Figure 3.

| Cancer | $n$ | $Aut(G)$ | $|Aut(G)|$ | $c$ | $\beta_G$ | $\mu_G$ |
|--------|-----|----------|-----------|-----|-----------|---------|
| AML    | 62  | $S_2^6 \times S_3^2 \times S_4 \times S_7^2$ | $1.011 \cdot 10^{14}$ | 1   | $7.03 \cdot 10^{-2}$ | 108     |
| Bld    | 29  | $S_2^4 \times S_3^2 \times S_4$ | $1.658 \cdot 10^9$ | 3   | $1.29 \cdot 10^{-1}$ | 20      |
| CML    | 73  | $S_2^8 \times S_3^6 \times S_4 \times S_7^2 \times S_8$ | $9.98 \cdot 10^{17}$ | 3   | $6.3 \cdot 10^{-2}$ | 115     |
| Col    | 49  | $S_2^6 \times S_3^2 \times S_4^3 \times S_8$ | $9.63 \cdot 10^{14}$ | 3   | $9.2 \cdot 10^{-2}$ | 58      |
| End    | 44  | $S_2^4 \times S_3^2 \times S_4 \times S_5 \times S_8$ | $4.82 \cdot 10^{12}$ | 2   | $1.1 \cdot 10^{-1}$ | 45      |
| Gli    | 64  | $S_2^8 \times S_3^6 \times S_4^3 \times S_6$ | $2.85 \cdot 10^{19}$ | 2   | $7.0 \cdot 10^{-2}$ | 128     |
| NSCL   | 48  | $S_2^4 \times S_3^2 \times S_4$ | $1.37 \cdot 10^{14}$ | 1   | $8.6 \cdot 10^{-2}$ | 75      |
| Pc     | 68  | $S_2^8 \times S_3^6 \times S_4^2 \times S_8$ | $1.39 \cdot 10^{16}$ | 4   | $6.4 \cdot 10^{-2}$ | 72      |
| Ren    | 60  | $S_2^8 \times S_3^6 \times S_6 \times S_8 \times S_{18}$ | $1.7 \cdot 10^{14}$ | 3   | $1.2 \cdot 10^{-1}$ | 52      |
| SCL    | 77  | $S_2^8 \times S_3^6 \times S_4 \times S_6 \times S_8 \times S_{12} \times S_{18}$ | $4.4 \cdot 10^{14}$ | 2   | $1.0 \cdot 10^{-1}$ | 150     |
| Thyr   | 28  | $S_2^4 \times S_3^2 \times S_4 \times S_7$ | $7.0 \cdot 10^2$ | 3   | $1.7 \cdot 10^{-1}$ | 24      |

Table 2: Orbit equivalence classes for the AML group.

| Nodes       | Degree |
|-------------|--------|
| PIK3R5      | 9      |
| PIK3CA      |        |
| PIK3CB      |        |
| PIK3CD      |        |
| PIK3R1      |        |
| PIK3R2      |        |
| PIK3R3      |        |
| CCND1       | 12     |
| MYC         |        |
| PPARD       |        |
| LEF1        |        |
| TCF7        |        |
| TCF7L2      |        |
| TCF7L1      |        |
| KRAS        | 5      |
| NRAS        |        |
| HRAS        |        |
| RPS6KB1     | 5      |
| RPS6KB2     |        |
| EIF4EBP1    |        |
| STAT3       | 5      |
| STAT5A      |        |
| STAT5B      |        |
| RARA        | 3      |
| PML         |        |
| LOC652346   |        |
| LOC652671   |        |
| ARAF        | 1      |
| RAF1        |        |
| BRAF        |        |
| KIT         | 12     |
| LOC652799   |        |
| IKBKB       | 5      |
| IKBKG       |        |
| MAP2K1      | 5      |
| MAP2K2      |        |
| SOS1        | 4      |
| SOS2        |        |
| PIM1        | 3      |
| PIM2        |        |
| NFKB1       | 3      |
| RELA        |        |
| MAPK1       | 2      |
| MAPK3       |        |
| Nodes  | Degree |
|--------|--------|
| DAPK1  | 2      |
| DAPK3  |        |
| DAPK2  |        |
| RPS6KA5| 5      |
| KRAS   | 5      |
| NRAS   |        |
| HRAS   |        |
| ARAF   | 5      |
| RAF1   |        |
| BRAF   |        |
| E2F1   | 1      |
| E2F2   |        |
| E2F3   |        |
| MAPK1  | 6      |
| MAPK3  |        |
| MAP2K1 | 5      |
| MAP2K2 |        |
| RASSF1 | 3      |
| FGFR3  |        |
| CCND1  | 3      |
| CDK4   |        |
| EGFR   | 1      |
| ERBB2  |        |

Table 3: Orbit equivalence classes for the bladder cancer group.

| Nodes  | Degree |
|--------|--------|
| PIK3R5 | 9      |
| PIK3CA |        |
| PIK3CB |        |
| PIK3CD |        |
| PIK3CG |        |
| PIK3R1 |        |
| PIK3R2 |        |
| PIK3R3 | 11     |
| CRK    | 9      |
| CRKL   |        |
| CBLC   |        |
| CBL    |        |
| CBLB   |        |
| SHC2   | 2      |
| SHC4   |        |
| SHC3   |        |
| SHC1   |        |
| MYC    |        |
| HDAC1  | 2      |
| HDAC2  |        |
| CTBP1  |        |
| CTBP2  |        |
| AKT1   | 14     |
| AKT2   |        |
| AKT3   |        |
| ARAF   | 5      |
| RAF1   |        |
| BRAF   |        |
| KRAS   | 5      |
| NRAS   |        |
| HRAS   |        |
| CHUK   | 4      |
| IKBKB  |        |
| IKBK   |        |
| CDK4   | 3      |
| CDK6   |        |
| CCND1  |        |
| E2F1   | 1      |
| E2F2   |        |
| E2F3   |        |
| ABL1   | 13     |
| BCR    |        |
| MAP2K1 | 5      |
| MAP2K2 |        |
| SOS1   | 4      |
| SOS2   |        |
| MECOM  | 4      |
| RUNX1  |        |
| STAT5A | 3      |
| STAT5B |        |
| MAPK1  | 2      |
| MAPK3  |        |
| TGFRB1 | 2      |
| TGFRB2 |        |
| NFKB1  | 1      |
| RELA   |        |

Table 4: Orbit equivalence classes for the CML group.
| Nodes     | Degree |
|-----------|--------|
| PIK3R5   | 4      |
| PIK3CA   | 4      |
| PIK3CB   | 4      |
| PIK3CD   | 4      |
| PIK3CG   | 4      |
| PIK3R1   | 4      |
| PIK3R2   | 4      |
| PIK3R3   | 4      |
| RHOA     | 4      |
| RAC1     | 4      |
| RAC2     | 4      |
| RAC3     | 4      |
| LEF1     | 4      |
| TCF7     | 4      |
| TCF7L1   | 4      |
| TCF7L2   | 4      |
| AKT1     | 8      |
| AKT2     | 8      |
| AKT3     | 8      |
| MAPK8    | 4      |
| MAPK9    | 4      |
| MAPK10   | 4      |
| ARAF     | 2      |
| RAF1     | 2      |
| BRAF     | 2      |
| SMAD4    | 2      |
| TGFBR1   | 2      |
| TGFBR2   | 2      |
| MAPK1    | 4      |
| MAPK3    | 4      |
| JUN      | 3      |
| FOS      | 3      |
| DCC      | 3      |
| SMAD2    | 3      |
| SMAD3    | 3      |

Table 5: Orbit equivalence classes for the colorectal cancer group.

| Nodes     | Degree |
|-----------|--------|
| PIK3R5   | 3      |
| PIK3CA   | 3      |
| PIK3CB   | 3      |
| PIK3CD   | 3      |
| PIK3CG   | 3      |
| PIK3R1   | 3      |
| PIK3R2   | 3      |
| PIK3R3   | 3      |
| APC2     | 1      |
| APC      | 1      |
| AXIN1    | 1      |
| AXIN2    | 1      |
| GSK3B    | 1      |
| LEF1     | 3      |
| TCF7     | 3      |
| TCF7L1   | 3      |
| TCF7L2   | 3      |
| AKT1     | 5      |
| AKT2     | 5      |
| AKT3     | 5      |
| KRAS     | 5      |
| NRAS     | 5      |
| HRAS     | 5      |
| ARAF     | 5      |
| RAF1     | 5      |
| BRAF     | 5      |
| BAD      | 3      |
| CASP9    | 3      |
| FOXO3    | 3      |
| PDPK1    | 10     |
| ILK      | 10     |
| MAP2K1   | 5      |
| MAP2K2   | 5      |
| MAPK1    | 3      |
| MAPK3    | 3      |
| SOS1     | 4      |
| SOS2     | 4      |
| MYC      | 4      |
| CCND1    | 4      |
| DCC      | 3      |
| SMAD3    | 3      |

Table 6: Orbit equivalence classes for the endometrial cancer group.
Table 7: Orbit equivalence classes for the glioma group. The nodes MDM2 and TP53 can only be permuted simultaneously with the nodes CDKN1A and CDKN2A.

| Nodes   | Degree |
|---------|--------|
| PIK3CA  | 9      |
| PIK3R5  | 7      |
| SHC1    | 5      |
| KRAS    | 20     |
| ARAF    | 5      |
| PRKCA   | 3      |
| CCND1   | 3      |
| AKT1    | 1      |
| AKT2    | 1      |
| AKT3    | 1      |
| E2F1    | 1      |
| E2F2    | 1      |
| E2F3    | 1      |
| PDGFRA  | 16     |
| MAP2K1  | 9      |
| PLCG1   | 4      |
| SOS1    | 4      |
| CDKN2A  | -      |
| CDKN1A  | (4)    |

Table 8: Orbit equivalence classes for the NSCL group.

| Nodes   | Degree |
|---------|--------|
| PIK3CA  | 7      |
| PIK3R5  | 5      |
| PLCG1   | 2      |
| AKT1    | 12     |
| ARAF    | 8      |
| PRKCA   | 3      |
| CCND1   | 3      |
| BAD     | 3      |
| CASP9   | 3      |
| CDK4    | 3      |
| CDK6    | 3      |
| SOS1    | 4      |
| SOS2    | 4      |
| MAPK1   | 4      |
| MAPK3   | 4      |
| MAP2K1  | 4      |
| MAP2K2  | 4      |
| STK4    | 2      |
| RASSF5  | 2      |
| CDK4    | 2      |
| CDK6    | 2      |
| Nodes       | Degree |
|------------|--------|
| PIK3R5     | 3      |
| PIK3CA     |        |
| PIK3CB     |        |
| PIK3CD     |        |
| PIK3CG     |        |
| PIK3R1     |        |
| PIK3R2     |        |
| PIK3R3     |        |
| FIGF       | 4      |
| PGF        |        |
| VEGFA      |        |
| VEGFB      |        |
| VEGFC      |        |
| MAPK1      | 1      |
| MAPK3      |        |
| MAPK8      |        |
| MAPK9      |        |
| MAPK10     |        |
| AKT1       | 8      |
| AKT2       |        |
| AKT3       |        |
| CHUK       | 5      |
| IKBKB      |        |
| IKBKG      |        |
| RAC1       | 4      |
| RAC2       |        |
| RAC3       |        |
| ARAF       | 2      |
| RAF1       |        |
| BRAF       |        |
| TGFB1      | 2      |
| TGFB2      |        |
| TGFB3      |        |
| E2F1       | 1      |
| E2F2       |        |
| E2F3       |        |
| RELA       | 14     |
| NKB1       |        |
| STAT1      | 6      |
| STAT3      |        |
| TGFBR1     | 6      |
| TGFBR2     |        |
| BAD        | 4      |
| BCL2L1     |        |
| RALA       | 3      |
| RALB       |        |
| CDK4       | 3      |
| CDK6       |        |
| SMAD2      | 3      |
| SMAD3      |        |
| BRCA2      | 1      |
| RAD51      |        |
| TGFA       | 1      |
| EGF        |        |

Table 9: Orbit equivalence classes for the pancreatic cancer group.

| Nodes       | Degree |
|------------|--------|
| ARNT       | 2      |
| FIGF       |        |
| PGF        |        |
| TDGFBR2    |        |
| PIK3R5     | 4      |
| PIK3CA     |        |
| PIK3CB     |        |
| PIK3CD     |        |
| PIK3CG     |        |
| PIK3R1     |        |
| PIK3R2     |        |
| PIK3R3     |        |
| PAK1       | 2      |
| PAK2       |        |
| PAK3       |        |
| PAK4       |        |
| PAK6       |        |
| PAK7       |        |
| AKT1       | 8      |
| AKT2       |        |
| AKT3       |        |
| ARAF       | 2      |
| RAF1       |        |
| BRAF       |        |
| EPAS1      | 18     |
| HIF1A      |        |
| RAC1       | 6      |
| CDC42      |        |
| MAPK1      | 5      |
| MAPK3      |        |
| MAP2K1     | 5      |
| MAP2K2     |        |
| ETS1       | 2      |
| JUN        |        |
| CRK        | 2      |
| CRKL       |        |
| SOS1       | 1      |
| SOS2       |        |
| RAP1A      | 1      |
| RAP1B      |        |

Table 10: Orbit equivalence classes for the renal group.
| Nodes | Degree |
|-------|--------|
| LAMA1 | 6      |
| LAMB4 |       |
| COL4A6|       |
| LAMA2 |       |
| LAMC1 |       |
| FN1   |       |
| LAMA3 |       |
| LAMC2 |       |
| LAMA4 |       |
| LAMC3 |       |
| LAMA5 |       |
| COL4A1|       |
| LAMB1 |       |
| COL4A2|       |
| LAMB2 |       |
| COL4A4|       |
| LAMB3 |       |
| COL4A5|       |
| TRAF1 | 2      |
| PTGS2 |       |
| NOS2  |       |
| TRAF2 |       |
| BIRC2 |       |
| TRAF3 |       |
| BIRC3 |       |
| TRAF4 |       |
| XIAP  |       |
| TRAF5 |       |
| PIK3R5| 4      |
| PIK3CA|       |
| PIK3CB|       |
| PIK3CD|       |
| PIK3CG|       |
| PIK3R1|       |
| PIK3R2|       |
| PIK3R3|       |
| PIAS1 | 3      |
| PIAS2 |       |
| PIAS3 |       |
| PIAS4 |       |
| AKT1  | 11     |
| AKT2  |       |
| AKT3  |       |
| CHUK  | 4      |
| IKBKB |       |
| IKBKG |       |
| CCNE1 | 2      |
| CCNE2 |       |
| CDK2  |       |
| E2F1  | 13     |
| E2F2  |       |
| E2F3  |       |
| RELA  | 8      |
| NKFB1 |       |
| MAX   | 8      |
| MYC   |       |
| CDK4  | 4      |
| CDK6  |       |
| APAF1 | 1      |
| CYCS  |       |

Table 11: Orbit equivalence classes for the SCL group.

| Nodes | Degree |
|-------|--------|
| RET   | 3      |
| CCDC6 |       |
| NCOA4 |       |
| TFG   |       |
| NTRK1 |       |
| TPM3  |       |
| TPR   | 3      |
| LEF1  | 3      |
| TCF7  |       |
| TCF7L1|       |
| TCF7L2|       |
| KRAS  | 8      |
| NRAS  |       |
| HRAS  |       |
| RXRA  | 2      |
| RXRB  |       |
| RXRG  |       |
| MYC   | 4      |
| CCND1 |       |
| MAP2K1| 3      |
| MAP2K2|       |
| PPARG | 3      |
| PAX8  |       |
| MAPK1 | 2      |
| MAPK3 |       |

Table 12: Orbit equivalence classes for the thyroid cancer group.