**Generation of genes having similar node degrees as SMGs and their associated perturbations**

To comprehend whether SMGs were pivotal in edgetic perturbations, we compared the proportions of perturbations involving SMGs and those from randomly generated genes with a similar degree of interacting proteins. First, we downloaded lists of pan-cancer and cancer specific significantly mutated cancer genes from the COSMIC Cancer Gene Census (https://cancer.sanger.ac.uk/census)1 and from the TCGA consortium (https://cancergenome.nih.gov/publications)2. The genes amounted to 719 and 299 cancer genes from COSMIC and Bailey *et. al*, respectively, and were classified according to their significance as cancer specific or as pan-cancer. We considered a gene to be significantly mutated in a certain cancer type if it was characterized as either cancer specific or pan-cancer, but affected that cancer type. Finally, in each cancer type, a union of the significantly mutated genes from both the above sources were considered as cancer specific significantly mutated genes (Supplementary Table 3). To find the number of perturbations involving the SMGs, we searched for any perturbed interactions having an SMG as an interacting partner. Then, in each cancer type, we merged all the interactions observed in both cancer and healthy in all the patients to generate all possible interactions within a cancer type. Next, for each cancer type, we determined the degree of each of the proteins within all the possible interactions of a cancer type. We used the degree of each SMG involved in any perturbation to randomly query for other proteins having a similar number of interacting partners to them (Supplementary File 3b). Finally, we determined the number of perturbations associated with the genes having a similar degree to the SMGs.

For each cancer type, the Z-test of proportions3 was used to estimate the statistical significance of the extent of edgetic perturbations associated with cancer-specific SMGs compared to the extent of edgetic perturbations associated with genes having a similar network topology to the SMGs. To do this, we first determined if there were significant differences in the proportion of perturbations associated with SMGs and the proportion of perturbations associated with randomly generated genes. Then, for each significant difference, we sought to find only the cancer types where the proportion of perturbations associated with SMGs were significantly larger than those associated with the randomly generated genes.

**Significantly mutated genes together with proteins having high degrees of connectivity in the PPIN are crucial players in edgetic perturbations of cancer PPINs**

Elevated mutation rate is a hallmark of cancer driver genes4–6. We analysed the involvement of SMGs as well as their first and second network neighbours in edgetic perturbations. Leiserson *et al.* previously suggested that somatic mutations affect subnetworks within PPINs via a heat diffusion model where “hot” nodes/SMGs propagate their heat to neighbouring nodes7. First, we found that not all SMGs are involved in edgetic perturbations, but only a specific number in each cancer type (Supplementary Table 3 and Supplementary File 3a). Also, there were significant differences in the proportion of perturbations associated with SMGs and those associated with the randomly generated genes having similar node degrees in the PPINs (Supplementary Table 4a). A majority of the perturbations across the cancer types had more instances where the portion of the perturbations associated with random genes was more substantial than the proportion of perturbations associated with SMGs. This observation was prominent in BRCA, PRAD and STES where the portion of the perturbations associated with random genes at both the first and second neighbours was significant, while HNSC had no significant differences in the two proportions. However, a look into the proteins involved in the majority of the perturbations associated with the random genes (e.g., *SKIP, HIST1H3J* and *EZH2*) revealed that the proteins function in gene expression deregulation in cancer and are potential molecules for therapeutic intervention in cancer8–14 (Supplementary Table 4b, Supplementary File 3b). With a rise in the interest of therapeutic targeting of cancer enabling proteins at the PPIN level, our findings suggest that therapeutic targeting of only SMGs involved in edgetic perturbations particularly in BRCA, PRAD, STES and HNSC may not yet be a sound idea. However, additional incorporation of epigenetic markers engaged in tumourigenesis of these cancer types may be additionally beneficial as previously suggested 15.

Nevertheless, we found that in some cancer types such as BLCA, LIHC and LUSC, edgetic perturbations were significantly associated with the SMGs (p < 0.05, Table 4c) as compared to edgetic perturbations resulting from randomly generated genes with similar degrees. Our findings correspond to those of 16–17 who pointed out that somatic mutations occurring at protein interaction interfaces may alter protein-protein interaction networks for example by resulting in loss of interactions or gain of new interactions. Besides, Cui *et al.* while analysing the effects of somatic mutations on the PPIN of liver cancer patients found that SMGs significantly rewire liver cancer PPINs when compared to random non mutated genes18. In the cancer types showing significant effects of the SMGs in disrupting the PPIN, our work provides opportunities for therapeutic targeting at the PPIN level as it is in the case with BH3 like proteins19 .

Supplementary Table 3: Specific cancer SMGs are involved in edgetic perturbations of cancer PPINs.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cancer type | Cancer type SMGs | Gained edges | Number of edges gained as 1st neighbours of SMGs | Number of edges gained as 1st or 2nd neighbours of SMGs | SMG protein products involved in edgetic gains | % of gains linked to SMGs | Lost edges | Number of edges lost as 1st neighbours of SMGs | Number of edges lost as 1st or 2nd neighbours of SMGs | SMG protein products involved in edgetic losses | % of losses linked to SMGs |
| THCA | 35(36) | 22831 | 254 | 7741 | 32 | 33.9 | 28065 | 334 | 9788 | 32 | 34.8 |
| BLCA | 52(54) | 20739 | 759 | 11046 | 47 | 53.26 | 19030 | 626 | 9695 | 42 | 50.94 |
| BRCA | 49(52) | 22195 | 638 | 11627 | 47 | 52.4 | 25516 | 553 | 12385 | 46 | 48.5 |
| COAD | 82(87) | 10065 | 449 | 5966 | 54 | 59.27 | 21024 | 679 | 10533 | 60 | 50.1 |
| KIRC | 24(26) | 18258 | 215 | 7001 | 24 | 38.34 | 33005 | 380 | 14418 | 25 | 43.7 |
| KIRP | 18(21) | 17174 | 154 | 5402 | 19 | 31.5 | 27141 | 233 | 9048 | 20 | 33.33 |
| KICH | 16(18) | 12423 | 102 | 3409 | 17 | 27.4 | 27490 | 246 | 9337 | 17 | 34 |
| HNSC | 50(52) | 21913 | 626 | 11439 | 46 | 52.2 | 27485 | 817 | 13947 | 44 | 50.7 |
| LUAD | 52(58) | 16622 | 480 | 9109 | 39 | 54.8 | 21907 | 608 | 11819 | 46 | 54 |
| PRAD | 39(40) | 17529 | 261 | 8173 | 28 | 46.63 | 22468 | 320 | 9060 | 30 | 40.32 |
| LUSC | 34(35) | 13108 | 426 | 5295 | 26 | 40.39 | 23242 | 430 | 10307 | 28 | 44.35 |
| STES | 24(27) | 24326 | 160 | 4685 | 20 | 19.26 | 20800 | 182 | 3827 | 20 | 18.4 |
| LIHC | 41(49) | 36458 | 1293 | 22998 | 44 | 63.1 | 51445 | 2161 | 32554 | 46 | 63.27 |

The numbers in the brackets next to the Cancer type SMGs indicate the protein products of the SMGs.

Supplementary Table 4a: The proportions of edgetic perturbations associated with SMGs and those associated with random genes with a similar degree significantly differ in size.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cancer type | Number of edges gained as 1st neighbours of SMGs | Number of edges gained as 1st neighbours of random genes | A | Number of edges gained as 1st or 2nd neighbours of SMGs | Number of edges gained as 1st or 2nd neighbours of random genes | B | Number of edges lost as 1st neighbours of SMGs | Number of edges lost as 1st neighbours of random genes | C | Number of edges lost as 1st or 2nd neighbours of SMGs | Number of edges lost as 1st or 2nd neighbours of random SMGs | D |
| THCA | 254 | 461 | 6.06e-15 | 7741 | 7282 | 5.84e-07 | 334 | 609 | 1.69e-19 | 9788 | 11057 | 1.47e-28 |
| BLCA | 759 | 662 | 0.009 | 11046 | 10818 | 0.02 | 626 | 851 | 2.35e-09 | 9695 | 10191 | 3.58e-07 |
| BRCA | 638 | 935 | 2.44e-14 | 11627 | 12997 | 4.09e-39 | 553 | 927 | 5.86e-23 | 12385 | 13776 | 7.16e-35 |
| COAD | 449 | 406 | 0.1 | 5966 | 5345 | 1.13e-18 | 679 | 805 | 0.0008 | 10533 | 11817 | 4.06e-36 |
| KIRC | 215 | 477 | 8.69e-24 | 7001 | 8001 | 2.01e-26 | 380 | 452 | 0.012 | 14418 | 12681 | 5.58e-43 |
| KIRP | 154 | 341 | 2.53e-17 | 5402 | 4567 | 3.19e-23 | 233 | 354 | 5.12e-07 | 9048 | 9168 | 0.27 |
| KICH | 102 | 89 | 0.34 | 3409 | 2799 | 3.93e-19 | 246 | 187 | 0.004 | 9337 | 7824 | 4.43e-44 |
| HNSC | 626 | 693 | 0.06 | 11439 | 11589 | 0.15 | 817 | 814 | 0.9 | 13947 | 13941 | 0.9 |
| LUAD | 480 | 431 | 0.09 | 9109 | 8375 | 7.49e-16 | 608 | 450 | 1.18e-06 | 11819 | 11082 | 1.12e-06 |
| PRAD | 261 | 247 | 0.53 | 8173 | 7109 | 2.1e-30 | 320 | 467 | 1.25e-05 | 9060 | 9265 | 0.05 |
| LUSC | 426 | 261 | 1.78e-10 | 5295 | 5127 | 0.03 | 430 | 470 | 0.17 | 10307 | 10132 | 0.1 |
| STES | 160 | 166 | 0.74 | 4685 | 5067 | 1.51e-05 | 182 | 200 | 0.35 | 3827 | 4868 | 3.85e-36 |
| LIHC | 1293 | 1508 | 3.43e-05 | 22998 | 23235 | 0.06 | 2161 | 1705 | 7.68e-14 | 32554 | 30974 | 3.86e-24 |

A: P-value showing if the difference between the proportion of edgetic gain perturbations associated with SMGs significantly differs from the proportion of edgetic perturbations associated with random genes at the first degree neighbours.

B: P-value showing if the difference between the proportion of edgetic gain perturbations associated with SMGs significantly differs from the proportion of edgetic perturbations associated with random genes at both the first and second degree neighbours.

C: P-value showing if the difference between the proportion of edgetic loss perturbations associated with SMGs significantly differs from the proportion of edgetic perturbations associated with random genes at the first degree neighbours.

D: P-value showing if the difference between the proportion of edgetic gain perturbations associated with SMGs significantly differs from the proportion of edgetic perturbations associated with random genes at both the both first and second degree neighbours.

Supplementary Table 4b: 5 cancer types show a significantly larger proportion of edgetic perturbations (1st neighbours of SMGs) associated with SMGs when compared to the proportion of edgetic perturbations (1st neighbours of random genes) associated with random genes with similar degrees.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cancer type | genes  in gains | Number of edges gained as 1st neighbours of random genes | E | Number of edges gained as 1st or 2nd neighbours of random genes | F | genes in losses | Number of edges lost as 1st neighbours of random genes | G | Number of edges lost as 1st or 2nd neighbours of random genes | H |
| THCA | 33 | 461 | - | 7282 | 2.42e-06 | 33 | 609 | - | 11057 | - |
| BLCA | 42 | 662 | 0.004 | 10818 | 0.01 | 46 | 851 | - | 10191 | - |
| COAD | 45 | 406 | - | 5345 | 5.64e-19 | 65 | 805 | - | 11817 | - |
| KIRC | 24 | 477 | - | 8001 | - | 23 | 452 | - | 12681 | 2.79e-43 |
| KIRP | 13 | 341 | - | 4567 | - | 19 | 354 | - | 9168 | - |
| KICH | 16 | 89 | - | 2799 | 1.97e-19 | 18 | 187 | 0.002 | 7824 | 2.21e-44 |
| LUAD | 34 | 431 | - | 8375 | 3.75e-16 | 46 | 450 | 4.39e-07 | 11082 | 8.99e-13 |
| PRAD | 28 | 247 | - | 7109 | 1.05e-30 | 30 | 467 | - | 9265 | - |
| LUSC | 20 | 261 | 8.9e-11 | 5127 | 0.01 | 25 | 470 | - | 10132 | - |
| LIHC | 44 | 1508 | - | 23235 | - | 45 | 1705 | 3.84e-14 | 30974 | 1.93e-24 |

E: P-value showing how significantly large the proportion of edgetic gains associated with SMGs is when compared to edgetic gain perturbations associated with randomly generated genes on the first degree neighbours.

F: P-value showing how significantly large the proportion of edgetic gains associated with SMGs is when compared to edgetic gain perturbations associated with randomly generated genes on both the first and second degree neighbours.

G: P-value showing how significantly large the proportion of edgetic losses associated with SMGs is when compared to edgetic loss perturbations associated with randomly generated genes on the first degree neighbours.

H: P-value showing how significantly large the proportion of edgetic losses associated with SMGs is when compared to edgetic loss perturbations associated with randomly generated genes on both the first and second degree neighbours.

-: Indicates no significant differences in the proportion of perturbations associated with SMGs and those associated with random genes, or the proportion of perturbations associated with SMGs is not larger than the proportion of perturbations associated with random genes.

Genes in gains: number of randomly generated genes with similar degrees to the SMGs and involved in edgetic gain perturbations.

Genes in losses: number of randomly generated genes with similar degrees to the SMGs and involved in edgetic loss perturbations.

**References**

1. Futreal, P. A. *et al.* A census of human cancer genes. *Nature Reviews Cancer* **4,** 177–183 (2004).

2. Bailey, M. H. *et al.* Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **173,** 371-385.e18 (2018).

3. Zou, K. H., Fielding, J. R., Silverman, S. G. & Tempany, C. M. C. Hypothesis Testing I: Proportions. *Radiology* **226,** 609–613 (2003).

4. Merid, S. K., Goranskaya, D. & Alexeyenko, A. Distinguishing between driver and passenger mutations in individual cancer genomes by network enrichment analysis. *BMC Bioinformatics* **15,** 308 (2014).

5. Pon, J. R. & Marra, M. A. Driver and passenger mutations in cancer. *Annu Rev Pathol* **10,** 25–50 (2015).

6. Tokheim, C. J., Papadopoulos, N., Kinzler, K. W., Vogelstein, B. & Karchin, R. Evaluating the evaluation of cancer driver genes. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 14330–14335 (2016).

7. Leiserson, M. D. M. *et al.* Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat. Genet.* **47,** 106–114 (2015).

8. Chen, Y., Zhang, L. & Jones, K. A. SKIP counteracts p53-mediated apoptosis via selective regulation of p21Cip1 mRNA splicing. *Genes Dev.* **25,** 701–716 (2011).

9. Chan, K.-M. *et al.* The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev.* **27,** 985–990 (2013).

10. Liu, G. *et al.* High SKIP expression is correlated with poor prognosis and cell proliferation of hepatocellular carcinoma. *Med. Oncol.* **30,** 537 (2013).

11. Ramakrishnan, S., Ellis, L. & Pili, R. Histone modifications: implications in renal cell carcinoma. *Epigenomics* **5,** 453–462 (2013).

12. Wang, X. *et al.* Clinical and prognostic relevance of EZH2 in breast cancer: A meta-analysis. *Biomed. Pharmacother.* **75,** 218–225 (2015).

13. Wang, Y. *et al.* Overexpression of YB1 and EZH2 are associated with cancer metastasis and poor prognosis in renal cell carcinomas. *Tumour Biol.* **36,** 7159–7166 (2015).

14. Yan, K.-S. *et al.* EZH2 in Cancer Progression and Potential Application in Cancer Therapy: A Friend or Foe? *Int J Mol Sci* **18,** (2017).

15. Poornima, P., Kumar, J. D., Zhao, Q., Blunder, M. & Efferth, T. Network pharmacology of cancer: From understanding of complex interactomes to the design of multi-target specific therapeutics from nature. *Pharmacological Research* **111,** 290–302 (2016).

16. Engin, H. B., Kreisberg, J. F. & Carter, H. Structure-Based Analysis Reveals Cancer Missense Mutations Target Protein Interaction Interfaces. *PLoS ONE* **11,** e0152929 (2016).

17. Jubb, H. C. *et al.* Mutations at protein-protein interfaces: Small changes over big surfaces have large impacts on human health. *Prog. Biophys. Mol. Biol.* **128,** 3–13 (2017).

18. Cui, H., Zhao, N. & Korkin, D. Multilayer View of Pathogenic SNVs in Human Interactome through in-silico Edgetic Profiling. *Journal of Molecular Biology* (2018). doi:10.1016/j.jmb.2018.07.012

19. Green, D. R. A BH3 Mimetic for Killing Cancer Cells. *Cell* **165,** 1560 (2016).