INTERSECTION BETWEEN GENES CONTROLLING VASCULARIZATION AND ANGIOGENESIS IN RENAL CELL CARCINOMAS

S. Souchelnytskyi

College of Medicine, Qatar University, Doha 2713, Qatar

Aim: To show that application of the systemic analysis may significantly improve comparison of different datasets. Different genes and proteins may converge on the same functional outputs. A comparison of 2 datasets by only identification names of affected molecules may miss that, leading to a conclusion that there is nothing in common for these datasets. Systemic analysis may overcome this limitation, by focusing on functions represented by the identification names. ***Materials and Methods***: Datasets were retrieved from open sources. Systemic analysis of vascularization features and angiogenesis signature was performed by using Cytoscape and its plugs-in. ***Results***: In contrary to the initial statement of the lack of overlap between the vascularization features and the angiogenesis genes-signature in renal carcinomas, we observed an intersection on the functional level. Analysis of the networks built with identification names of vascularization and angiogenesis datasets showed an intersection, which included potent regulators of vessel formation and growth. ***Conclusion***: Analysis of networks may expose functional links, which may be missed by a direct identification names comparison.
Key Words: cancer, vascularization, angiogenesis, systemic analysis.

Systemic analysis of experimental data may prevent erroneous conclusions when comparing different datasets. Systemic analysis may unveil overlapping patterns of genes, which can be missed by comparison of genes name by name. Vascularization and angiogenesis are crucial cancer hallmarks [1, 2]. Recently there has been proposed a 14-genes vascularization signature to separate good and poor prognosis survivors with renal cell carcinomas [3]. The authors concluded that there was no overlap with the previously reported 48-genes angiogenesis signature [3, 4], despite the fact that both signatures refer to angiogenesis. This conclusion may be the result of an identification name for identification name comparison, which could miss functional connections. In a living system, different genes, transcripts as well as proteins work in a cooperative way to ensure functional responses [5, 6]. Therefore, different genes and proteins may have a similar functional output. Such functional connections would be unveiled by a systemic analysis [5–8]. Systemic analysis focuses on affected functions rather than on identification names, and may lead to better clinically relevant conclusions.

MATERIALS AND METHODS

The lists of the 14-genes and the 48-genes signatures were used to build networks using the Cytoscape tool (version 3.6.0; www.cytoscape.org) [9]. Cytoscape allows building, visualization and analysis of networks in different formats. Flexibility of Cytoscape in managing of different datasets, options for retrieval of information from open source databases and for comparison of different networks prompted use of this tool in this study. UniProt database was used to import a network of nodes and edges built with the genes of the signatures. The 14-genes network contained 178 nodes and 303 edges, and the 48-genes network contained 836 nodes and 1763 edges. These two networks were analyzed in Cytoscape for an intersection. Intersection detected 62 nodes. Retrieved shared nodes were further searched in Cytoscape for regulators of angiogenesis, migration, vessel and endothelial cells.

RESULTS AND DISCUSSION

Re-examination of the published 14-genes vascularization [3] and the 48-genes angiogenesis [4] signatures showed that there is an overlap between these signatures on the networks level (Figure). An intersection of the corresponding networks retrieved 62 shared nodes, with the use of the UniProt database. Moreover, 17 of these nodes are involved in regulation of angiogenesis, migration and endothelial cell functions (the Figure and the Table). The network built with the list of 14 genes generated a network with 178 nodes and 303 edges. The network built with the list of the 48 genes generated a network with 846 nodes and 1763 edges. The shared 62 nodes and the 17 genes of relevance to angiogenesis, migration and endothelial cells functions are listed in the Table.

It has to be noted that there are many systems biology tools for building and analyzing of networks. Here is reported an analysis with the use of Cytoscape, as this tool has a well-developed plugs-in (Apps) and allows easy retrieval of data from different databases. Other tools, such as FunCoup or String, may also be used, as long as they retrieve complete datasets. To evaluate a completeness of a dataset, the author did cross-checks with published reports. This cross-check shows whether in the network would be retrieved interactions described earlier. It would indicate comprehensiveness of the coverage of interactions and whether the network represents available knowledge. If a tool does not retrieve such interactions, it would
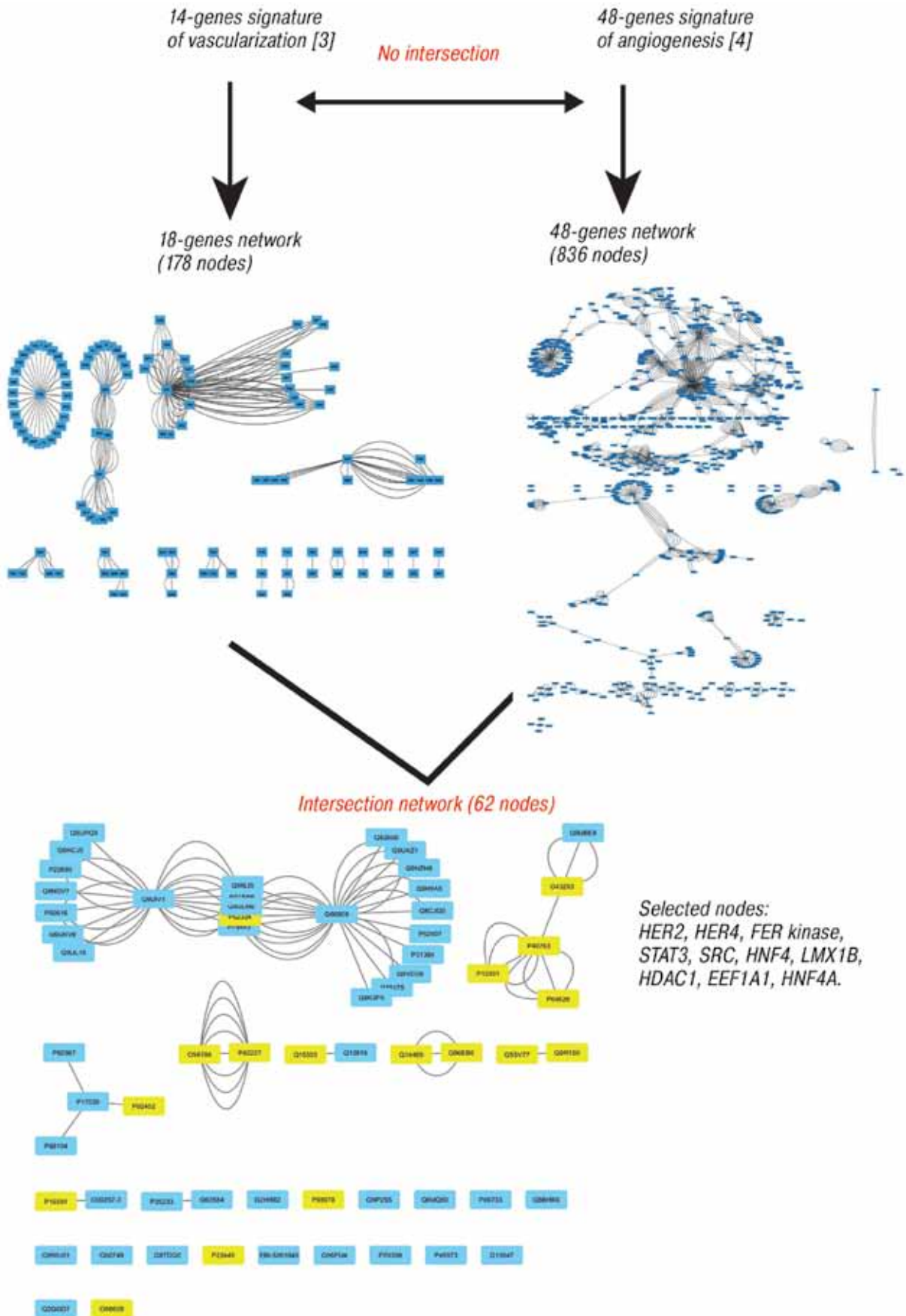
Figure. Systemic analysis identifies links of the vascularization signature with the angiogenesis genes-signature, which were not detectable by a direct comparison. A direct comparison of the 14-genes and 48-genes signatures of vascularization and angiogenesis did not show an intersection [3, 4]. The analysis of the corresponding networks detects 62 nodes common for the networks. There are nodes (in yellow) with relevance to regulation of angiogenesis, migration, vessel formation and endothelial cells. The nodes are annotated in the Table

Table. Annotations of the 14-genes signature of vascularization [3], the 48-genes signature of angiogenesis [4], 62 shared nodes between the networks and 17 nodes which were retrieved from the shared 62-nodes by a search for "angiogenesis AND vessel AND migration AND endothelial". Fourteen and 48 genes are annotated by their GO terms. Sixty-two and 17 nodes are annotated by the accession numbers and names in parentheses

14-genes signature of vascularization [3]	48-genes signature of angiogenesis [4]	62 shared nodes between networks of the 14- and 48-genes	17 nodes retrieved with search terms: angiogenesis, vessels, migration and endothelial cells
ADH5, NLRC4, RPL36A, RPLP2, SLC16A4, TNFSF8, ZNF16, SOCB, GOSR2, IFNA13, CMYA5, STAT3, KCNJ12, MED10	*ACVRL1, AGGF1, AMOT, ANG, ANGPTL3, ANGPTL4, ATPIF1, BTG1, C1GALT1, CANX, CDH13, CHRNA7, COL4A2, COL4A3, EGF, EMCN, EPGN, ERAP1, FOXO4, HTATIP2, IL17F, IL18, IL8, MYH9, NCL, NF1, NOTCH4, NPPB, NPR1, PF4, PLG, PML, PROK2, RHOB, RNH1, ROBO4, RUNX1, SCG2, SERPINF1, SHH, SPHK1, SPINK5, STAB1, TGFB2, THY1, TNFSF12, TNNI3, VEGFA*	P16591 (p94-Fer kinase), O54784 (Dapk3), O43293 (DAP-like kinase), Q9BUH7 (WD repeat-containing protein antisense to TP73 gene), Q63584 (Tmp21), P11520 (Tis11a), Q86UE5 (KIAA1582), Q96EB6 (Sirt1), Q9WU01 (Slm1), Q2XNF6 (SIR2L1), Q5VYR3 (NIRF), Q9Y2L0 (CNOT1), Q9UF92 (CNOT6L), Q9UEG1 (UBA52), Q9R1S0 (Eppb9), Q9NU89 (MAGE-D3), Q9H9A5 (CNOT10), Q9H5A8 (Proto-oncogene c-Src), Q9H3E0 (HSPC131), Q96IX4 (RQCD1), Q8VEG6 (Cnot6l), Q8TDQ0 (TIMD3), Q8TCR0 (CAGH26), Q8CJ65 (Slamf7), Q7L3L1 (ZXDC), Q6P4S0 (AGO1), Q6P2M9 (C2orf29), Q60735 (Src-associated in mitosis 68 kDa protein), Q2PNI9 (LAK1), Q15303 (HER4), Q14469 (HES1), P35233 (Ptpn2), Q5SV78 (Dioxin-inducible factor 3), P06733 (Enolase 1), P02452 (Alpha-1 type I collagen), O00257-3 (CBX4), Q9NZN7 (LENG2), P78543 (BTG family member 2), O88609 (LIM/homeobox protein LMX1B), Q9UPQ9 (TNRC6B), Q2HR82 (K8), P45973 (CBX5), Q13547 (HDAC1), P70338 (Gfi1), O43846 (ErbB3-binding protein 1), P68104 (EEF1A1), P50616 (Transducer of erbB-2, 1), P31384 (Carbon catabolite repressor protein 4), Q8K3P5 (Cytoplasmic deadenylase, CCR4), P52907 (CAPZA1), Q9ULM6 (CCR4a), Q60809 (Caf1), Q9UIV1 (BTG1-binding factor 1), P31607 (B-cell translocation gene 1 protein), Q96FG0 (HZF1), P04626 (Proto-oncogene Neu), Q9UKV8 (AGO2), Q4VAB3 (Eif2c2), Q923E1 (Dab2), O14916 (APRF), P42227 (Stat3), P22449 (Hnf4a)	P12931 (pp60c-src), Q15303 (HER4), Q14469 (HES1), Q5SV78 (Dioxin-inducible factor 3), P02452 (Alpha-1 type I collagen), O88609 (LIM/ homeobox protein LMX1B), P62324 (B-cell translocation gene 1 protein), P04626 (HER2), P98078 (Differentially expressed in ovarian carcinoma 2), P40763 (STAT3), P42227 (Aprf), P22449 (Transcription factor HNF-4), P16591 (FER tyrosine kinase), O54784 (Dapk3), Q96EB6 (SIRT-1), Q9R1So (B9d1), O43293 (DAP-like kinase)

be advised to change settings of a search or even change a tool.

Accessibility and type of databases is another important point in selection of a tool for network analysis. Use of open source databases is advised, as proprietary databases may have restrictions on access to their data. Open source databases are often specialized and may focus on certain types of molecules (DNA, RNA, proteins, metabolites, small molecules), interactions (direct, functional, correlations) or species (*H. sapiens*, *M. musculus, D. melanogaster, C. elegance, D. rerio, etc*) [10]. Therefore, it is advised to ensure that the databases used for a network building represent types of data which are relevant to a study.

The detected with Cytoscape shared 62 nodes represent proteins with reported roles in tumorigenesis, angiogenesis and stroma formation. Examples are Her2, Her4, pp60c-src, stat3, p94-fer and Dapk3 kinases (see Table). Thus, detection of shared nodes confirms that the vascularization includes regulators of angiogenesis.

Detection of the shared nodes emphasizes the power of systemic analysis in unveiling of hidden links. As the old saying goes "all roads lead to Rome", the same is valid for cancer — many mechanisms may lead to the same result, i.e. tumor vascularization [1, 2]. These cancer-roads may be traced with systems biology tools, which may identify common functional targets. The shared nodes represent well-known regulators of angiogenesis and vascularization of tumors. For example, Her2, Her4, STAT3, SIRT-1 and Src have been reported to regulate tumor angiogenesis [11–15]. It has to be noted that these

shared nodes show functional links to well-known genes of relevance for development of renal carcinoma, e.g. *VHL*, *VEGFR*, *HIF*, *mTOR* [11–15]. Moreover, the shared nodes represent regulators of functions targeted in treatment of renal carcinoma, such as anti-VEGFR treatments [16]. This confirms an importance of the systemic analysis for further clinical applications, as it allows linking of diagnostic and prognostic marker signatures with pathways targeted in treatments of patients. Thus, the signatures of angiogenesis and vascularization do have a functional overlap, but to detect it, a network analysis has to be applied.

CONCLUSIONS

Network-based analysis may detect functional similarities between seemingly not intersecting datasets. A network-based analysis is rooted in the systemic character of tumorigenesis, and it unveils how different genes and proteins may regulate the same functions. This study may be used as a template for an analysis of intersections for other datasets.

REFERENCES

1. Wang Z, Dabrosin C, Yin X, *et al*. Broad targeting of angiogenesis for cancer prevention and therapy. Semin Cancer Biol 2015; **35**: Suppl: S224–S43.

2. Johannessen TC, Wagner M, Straume O, *et al*. Tumor vasculature: the Achilles' heel of cancer? Expert Opin Ther Targets 2013; **17**: 7–20.

3. Ing N, Huang F, Conley A, *et al*. A novel machine learning approach reveals latent vascular phenotypes predictive of renal cancer outcome. Sci Rep 2017; **7**: 13190.

4. Liberzon A, Birger C, Thorvaldsdottir H, *et al*. The molecular signature database hallmark gene set collection. Cell Systems 2015; **1**: 417–25.

5. Califano A, Alvarez MJ. The recurrent architecture of tumour initiation, progression and drug sensitivity. Nat Rev Cancer 2017; **17**: 116–30.

6. Yan W, Xue W, Chen J, Hu G. Biological networks for cancer candidate biomarkers discovery. Cancer Inform 2016; **15**: 1–7.

7. Attarha S, Andersson S, Mints M, Souchelnytskyi S. Individualised proteome profiling of human endometrial tumours improves detection of new prognostic markers. Br J Cancer 2013; **109**: 704–13.

8. Zakharchenko O, Cojoc M, Dubrovska A, Souchelnytskyi S. A role of TGFβ1 dependent 14-3-3σ phosphorylation at Ser69 and Ser74 in the regulation of gene transcription, stemness and radioresistance. PLoS One 2013; **8**: e65163.

9. Shannon P, Markiel A, Ozier O, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2017; **13**: 2498–504.

10. Miryala SK, Anbarasu A, Ramaiah S. Discerning molecular interactions: A comprehensive review on biomolecular interaction databases and network analysis tools. Gene 2018; **642**: 84–94. doi: 10.1016/j.gene.2017.11.028.

11. Jomrich G, Schoppmann SF. Targeting HER 2 and angiogenesis in gastric cancer. Expert Rev Anticancer Ther 2016; **16**: 111–22.

12. Alameddine RS, Otrock ZK, Awada A, Shamseddine A. Crosstalk between HER2 signaling and angiogenesis in breast cancer: molecular basis, clinical applications and challenges. Curr Opin Oncol 2013; **25**: 313–24.

13. Gao P, Niu N, Wei T, *et al*. The roles of signal transducer and activator of transcription factor 3 in tumor angiogenesis. Oncotarget 2017; **8**: 69139–61.

14. Wang J, Zhuang S. Src family kinases in chronic kidney disease. Am J Physiol Renal Physiol 2017; **313**: F721–8.

15. Kunhiraman H, Edatt L, Thekkeveedu S, *et al*. 2-Deoxy glucose modulates expression and biological activity of VEGF in a SIRT-1 dependent mechanism. J Cell Biochem 2017; **118**: 252–62.

16. Zarrabi K, Fang C, Wu S. New treatment options for metastatic renal cell carcinoma with prior anti-angiogenesis therapy. J Hematol Oncol 2017; **10**: 38.