PRECISION TOOLS

The Computational Network Imaging Frontier: Relevance for Digital Biomarkers in Precision Oncology.

by Marco Dominietto Ph.D and Enrico Capobianco Ph.D

Cancer treatment is no doubt the greatest of the big challenges in the newborn field of Precision Medicine. As an effect of advances in imaging technologies and methods, the assessment of therapeutic response in cancer patients now involves a mix of qualitative and quantitative aspects, thus calling for integrative approaches linking together various types of evidences obtained from molecular profiling, cell signaling, experimental omics and clinical records. Such multiplexing gives origin to a multitude of data, presenting an unprecedented opportunity for building multilevel inference algorithms targeted to cancer therapy. We describe a network-driven methodology and the rationale that leverages the plasticity and adaptability of possible configurations and its representative power.

From Images to Networks

Modern imaging techniques, such as Magnetic Resonance Imaging (MRI), Computer Tomography (CT), Positron Emission Tomography (PET), when used in combination with contrast agents and radiotracers, allow measurements of anatomical details with high resolution and also maps of the dynamics generated by several physiological parameters. The overall result is a large number of 3D datasets, each describing an anatomical or physiological feature of the tumor and of the surrounded healthy tissue. Roughly speaking, a parallel can be made with taking several pictures of the same object from different angles; while each image alone shows a detail of the target object, the aggregate of all images defines the picture of it. Similarly, in imaging only the integration of all the information arising from different techniques can eventually represent the status of a tumor.

Such integration is known to be unfeasible in the image space, due to the redundancy of effects from many variables that need to be accounted. It is therefore a necessity to transform the image in another domain, making thus calculations possible. Networks build an ideal framework for dealing with imaging by integrating many variables in a relatively simple structure consisting of nodes and edges, the latter indicating the interactions between the former.

All 3D images consist of a 3D array of voxel, which is defined as a unit of graphical information that defines a point in three-dimensional space (x,y,z) The graphics community introduced volumetric representations for geometric objects more than three decades ago.[1] Voxelization enables an approximation, i.e. a 3D scan conversion process of some continuous geometric shapes into an array of voxels in the 3D discrete space.[2,3]

Graph partitioning approaches have been proposed on the basis of a simple couple of steps: a) Clustering of similar voxels to form regularly spaced supervoxels of a more uniform size and useful to compute robust statistics; b) Connectivity of supervoxels with their neighbors by edges, thus forming the graph. In order to account for shape and boundaries information, ad hoc features were considered for the supervoxels.[4] The results of such segmentation algorithm is a reduction of the computational complexity, and the incorporation of qualitative features.

Notably, the content of each voxel corresponds to a set of features F(f1, f2, … fn) determined by the different image techniques. Such features are quite heterogeneous, and can be represented beyond physical parameters (density, diffusivity, etc.) and shape and composition parameters[5] by measures, qualitative characteristics, scores, functions collected as quantifiable characteristics destined to be algorithmically processed to learn and generalize.[6] In the image domain *(I-space)*, each voxel can be therefore identified by the following equation:

$$v(x,y,z) = F(f_1, f_2, \ldots f_n)x,y,z$$

For instance, macro categories of features may refer to:

- f_1 = physiological properties, such as perfusion, oxygenation, ph, hypoxia…
- f_2 = anatomical structure
- f_3 = tissue characterization, such as hypoxic, necrotic, viable, …. etc.

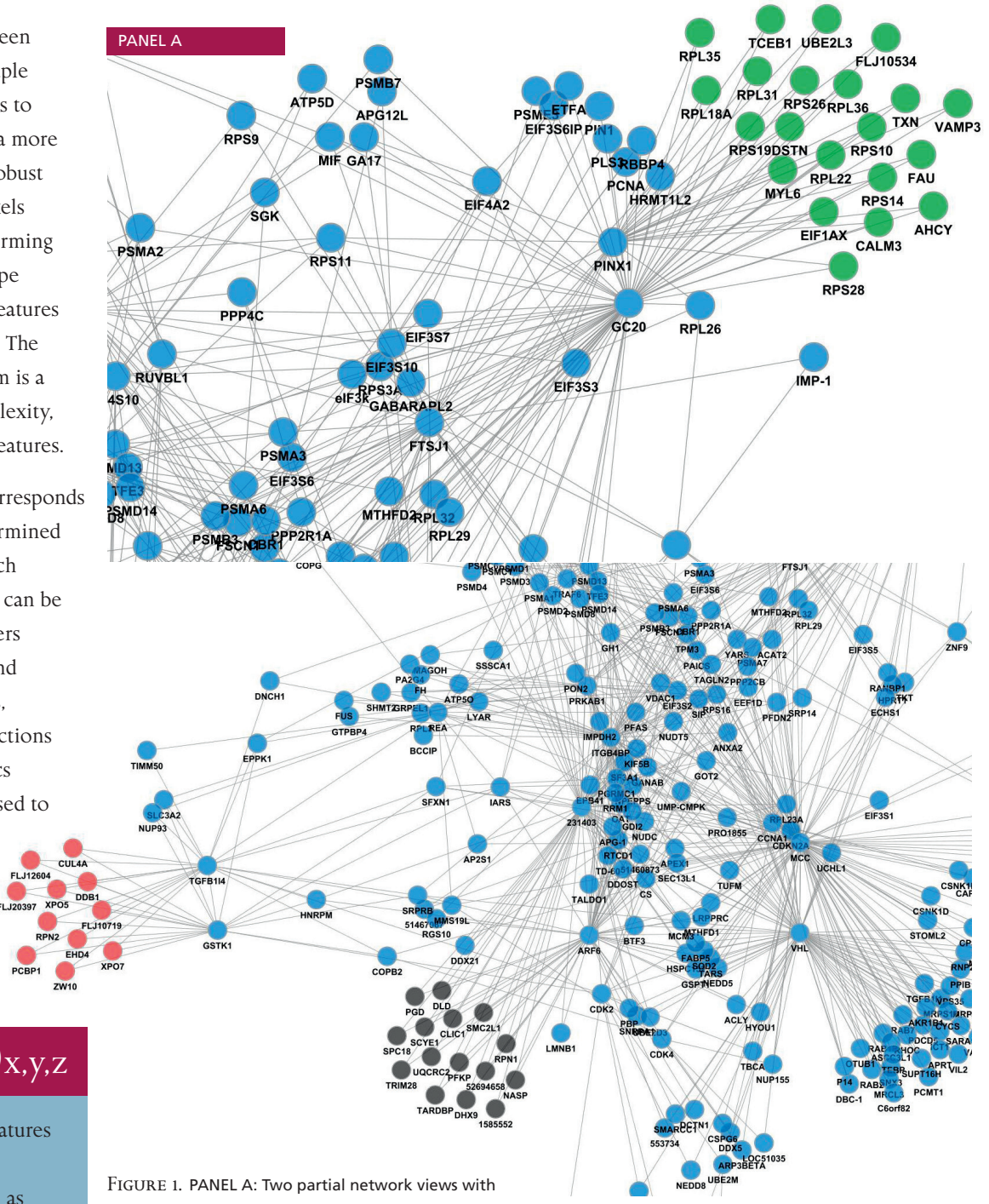


FIGURE 1. PANEL A: Two partial network views with hubs (nodes rich in connectivity) and modules (black, red, green). Interactome networks from integration of various sources and small experiments identifying 23462 interactions from 7385 proteins. PANEL B: Interactome network from large scale Mass Spectrometry experiments identifying 6463 interactions between 2235 human cell proteins (Ewing et al, 2007).[7]

In network domain (N-space), nodes are corresponding to image voxels. Edges linking nodes (or communicating voxels) are defined on the basis of similarity/dissimilarity measures of the features F. By means of this simple transformation, we keep both spatial (x,y,z) and feature F information, this latter delivering a key context for conducting inferential analyses. Networks are well-known computational frameworks, elucidating relationships between nodes through edges that can create dense or modular configurations (Fig 1).

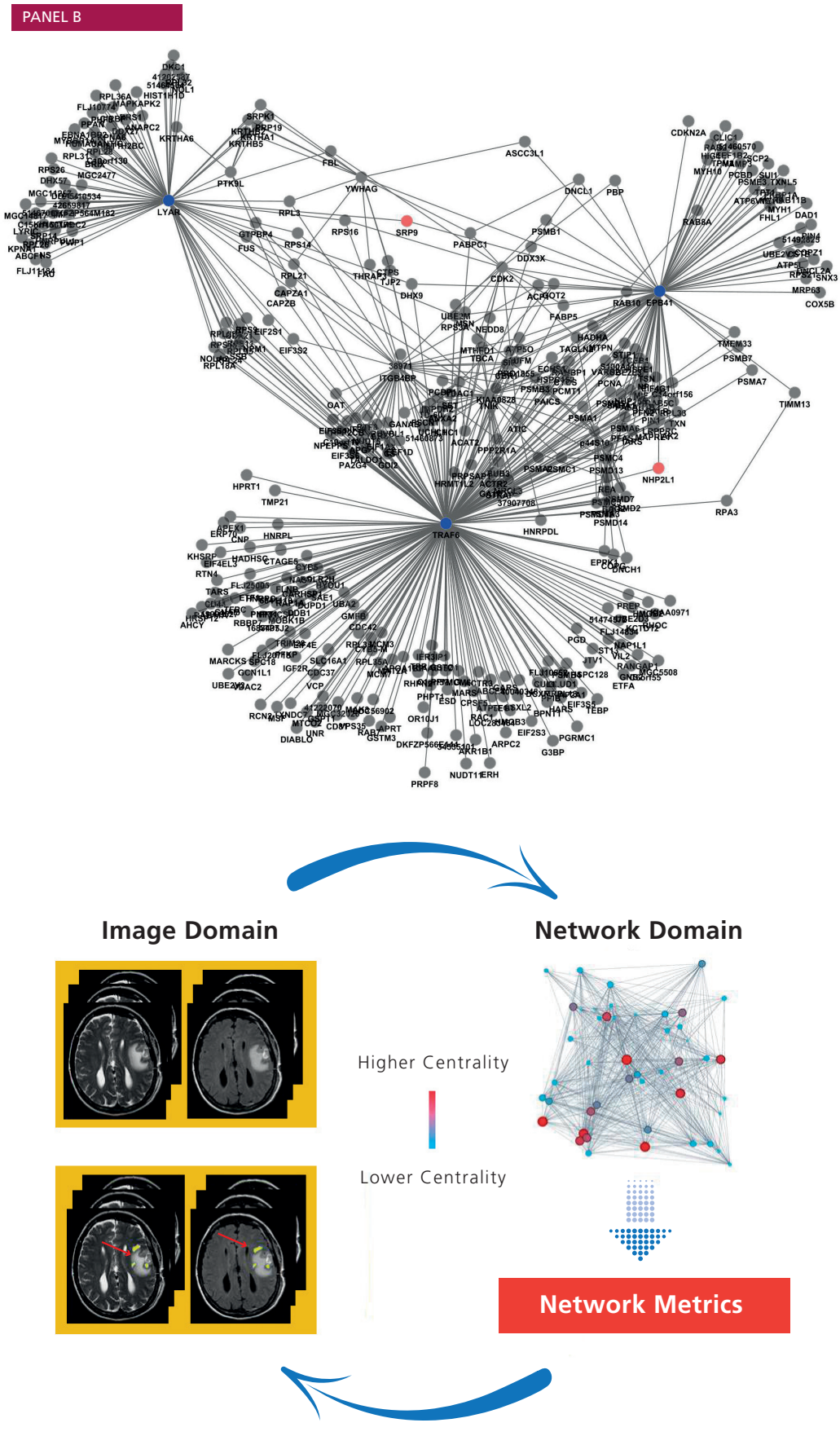


FIGURE 2. From image to network via voxel-node associations. Central nodes allow traffic between many nodes, thus they have high participation rates to many biological processes.

Voxel-node associations (Fig 2) imply the possibility of mapping from image (I) to network (N) spaces aiming to gain a few important advantages:

- A parsimonious representation of the informative contents, such that I → N allows to reduce the overall complexity (dimensionality, redundancy etc.) and improve the inference process
- The network metric can guide through interdependent relationships between nodes which appear in the form of connectivity patterns, and establish confidence and significance to such relationships
- Through the features that were identified in I-space, an ensemble use of them can be operated in N-space and this can lead to identify interesting modules

From Networks to Images

From N-space, once patterns (paths, motifs, etc.) and modules have been detected and assessed as significant, one can return to the I-space with improved discrimination power. If the network metrics have revealed the presence of significant nodes regulating other node dynamics or the overall network, these are the preferred candidate nodes to be back transformed and highlighted in I-space. This is the case displayed in Figure 2, in which the network topology allows the identification of nodes with high centrality and/or degree in the N-space. Assuming that these nodes are highly important in the network hierarchy, their back-transformation in the I-space may reveal important features.

This sort of back-transform might be particularly relevant under conditions such as tumor microenvironment in which tumor and stromal characterizations might be intertwined and convoluted to an extent that does not allow clear assessment of specific effects. In other terms, the kind of deconvolution that can be operated in the N-space can reveal relationships between nodes, beyond each specific single network entity therefore, and thus pinpointing ensemble dynamics which were much harder to detect in the I-space.

Tumor-Host tissue interaction

It is widely accepted that tumors show high level of heterogeneity when compared to the healthy tissue. Nevertheless, the tumor is not growing independently in the hosting organ, it is fully integrated in it. The interaction between tumor and hosting tissue is therefore crucial for two aspects: on one side, the tumor invades the surrounding tissue, on the other side the healthy tissue tries to prevent this from happening.

This competition involves the surface of the tumor, together with deep regions connected with the outer part of the organ by means of vascular or lymphatic vessels. As a result of short-distance, such complex interaction can be described only by taking into account both tumor and healthy tissue physiology. In other words, only by considering tumors as subnetworks of the hosting organ, each subject to different rules, can we adhere to real tumor-host tissue interaction dynamics.

Following the same approach, we can include the interaction of tumor with the outer organs. Such interaction, which is mediated by the moving cells flowing from the primary tumor via both arteries and vein or lymphatic vessels, is responsible for the colonization of other organs that originates the metastatic process. We thus assume that metastases formation is a sort of long-distance interaction effect observable inside an expanded network covering tumor, hosting organ and metastatic organ.

Controllability

Ideally, the most intuitive way of controlling complex networks is to rely on just a few of their structural elements, nodes and links. For instance, the identification may target a minimum set of nodes through which the control over the entire networks could be exerted. These would be called driver nodes, and would of course be critical, however non unique, meaning that multiple sets of drivers of the same size but with different nodes are expected to exist. Another aspect worth the consideration is that the functional characterization of complex networks is very hard to achieve, due to the dynamic convolution of structural and systematic features with more transient and even random ones. This is to say that we need to identify the object of our control, before finding the best possible way to implement it through selected drivers.

Thus controllability requires from one hand an ensemble view, in which the denser the network presents the fewer are the drivers needed to control it, and the sparser are the networks the more drivers would be needed and the harder the effective control[8]. From another hand, selectivity is required with reference to specific targets, whose prioritization depends on structural or functional network characteristics. Owing to the impossibility of achieving full control, some risks must be taken into consideration due to limited load and capacity of each specific node, indicating the possibility that receiving links is possible till a limit beyond which the propagation of effects cannot be prevented. This will lead to so-called cascading failures or catastrophic events.[9]

Having established a controlled setting in N-space may be very important also for the impacts in I-space, after back-mapping. In particular, monitoring could be focalized in a restricted and landmarked tissue area, but in case a more spread area needs surveillance, this can be performed through a number of well-selected anchor points. Recently, network inference work revealed the role of sentinel nodes designed to bridge between cancer phenotypes. A clearly modular organization came out to indicate cancer hallmark specific and topological coupled characterization, elucidating further functional controllability aspects.[10]

Stability

The inherent complexity of tumor systems implies the presence of instabilities affecting the states and its critical transitions. In turn, network stability can be investigated in correspondence with perturbations and the overall resilience. In physics this type of dynamics is called phase transitions from ordered to disordered states and allows the inspection of critical points on the basis of control parameters.[11]

In short, the state of the system is seen as a vector with length established by the number of nodes, thus given by $X = (x_1,\ldots,x_n)$, with Xi as the state of node i. We assume here that time t has been fixed, i.e. X is stationary. Stability can be locally assessed by checking:

$$\Delta Y = dY/dt = d(X\text{-}X^*)/dt = AY$$

The importance of the matrix A is known: it embeds the node interactions, and from its eigen-decomposition the system stability properties can be investigated. The off-diagonal terms in A involve of course interconnectivities which may be simply on and off, or be a function of some parameters which may be totally or partially known. Since the network dynamics depend on the interactions, naturally enough also stability is influenced, and the main question is to find indicators of such stability, in particular telling us when the system is approaching conditions of instability. Such indicators correspond to early warnings and are often referred as tipping points, and typically a slow return to equilibrium after small system perturbations is an indication of a so-called critical slowing down, typically monitored by inspecting measures such as variance and power spectrum of state variables.

In real world applications, stationarity can be replaced by quasi-stationarity owing to the presence of attractor states.[12] In the presence of perturbations, the system's resilience is enabled to maintain normal functionality, and can induce such states in correspondence to the existence or the establishment of particularly robust phenotypes and/or in response to particular biological processes. Most states are thus to be considered unstable, and even if it is hard to predict when such condition can start, the ultimate effect is that of eventually inducing a system's convergence to stable or low-energy states, the attractors. Network resilience was recently examined in its complexity, which is typically depending

on multi-dimensional manifolds involving bifurcation points or equivalently implying transitions to lowly resilient and less desirable states.[13]

An interesting aspect in cancerous systems refers to adaptive dynamics that include functional advantages from proximity to critical points.[14] Tumor microenvironment is very likely an example of a system working near a critical point, between order and disorder, and characterized by vast and differentiated phenotypic variability, which may be interpreted as a strategy for continuous adaptation and eventually survival. It's important to note that during tumor growth, while the inner part of the tumor tends to reach a sort of macroscopic stability, the border of the tumor continuously interacts with the healthy tissue modifying the microenvironment in order to promote tissue invasion. Of interest, the fact that with network ensembles the analytical and computational aspects of stability can be investigated, allowing the most efficient representation for heterogeneity in the origin domain, and likely observed close to criticality.

Big Data and Digital Biomarkers

Single-molecule centered research and developments aimed to digital detection of biomolecules for clinical applications are expected to generate a new wave of biomarkers.[15] An interesting direction was undertaken with the endo-phenotype characterization induced by the joint modeling of imaging and genetic variants associated with disease, supported by the assignment of probabilistic measures of relevance to both.[16] A more general quantification of heterogeneity classified in multiple categories to assess imaging biomarkers was also recently proposed.[17]

Especially digital detection may ensure measurement resolution and marker sensitivity on the basis of discrete counts which are not available from system's approaches. These evidences will in part contribute to the novel Big Data, coupled with generated clinical evidence, such as phenotypic signatures. Digital measurements will thus represent in perspective the type of data liquidity widely considered the most valuable resource for precision medicine. The reason relies in the translational power of digital biomarkers, a mix of physiological, pathological, behavioral and technological contents characterizing both normal and altered biological processes.

Among digital biomarkers, imaging-related ones are expected to be relevant for establishing health vs disease conditions, thus monitoring health status as well as assessing therapeutic interventions and drug responses. A multifaceted fingerprint of individuals is thus obtained by leveraging on a synergy of new data types (behavioral and contextual beyond subjective and observational clinical evidences).

Overall, the most expected impact is the possibility to be able to measure more objectively because of the synthesis of three newly acquired data properties: collective (aggregated group- or community-wise), comprehensive (multi-profiling) and integrative (multi-source evidences, including bias information). Consequently, more complete signatures of phenotypes explaining significant variation at individual and population scales could deliver timely and accurate personalized clinical decisions and patient stratification. Important sources of digital biomarkers are appearing worldwide, including all the Electronic Health Record (EHR) repositories. More recently, through The Cancer Imaging Archive (TCIA), the National Cancer Institute has supported the creation of both a Big Data resource and of a Quantitative Imaging Network to burst field developments and expand critical tumor response biomarkers in clinical trials.[18] Another initiative refers to the Multimodal Brain Tumor Image Segmentation Benchmark (BRATS), in which high-quality image data are analyzed by state-of-the-art algorithms and manual annotations are provided.[19]

Quantitative Imaging to assess cancer-targeted therapies

Heterogeneity implies that in each patient and in each tumor a variety of features can be present (Fig 3). Imaging techniques allow to monitor intra-tumoral evolution by extracting, mining and analyzing quantitative data. Variation is present in imaging features identifying cell clusters, reflected into differentiated molecular characteristics, contrast enhancement or necrosis identification (Fig 4).

Establishing proper quantitative metrics is thus crucial to measure temporal and spatial heterogeneity.[20] While statistical inference could be conducted from the extracted data, the most important step is to build suitable representation systems in which the mapping to tumor regions with spatially distinct environments can be operated.

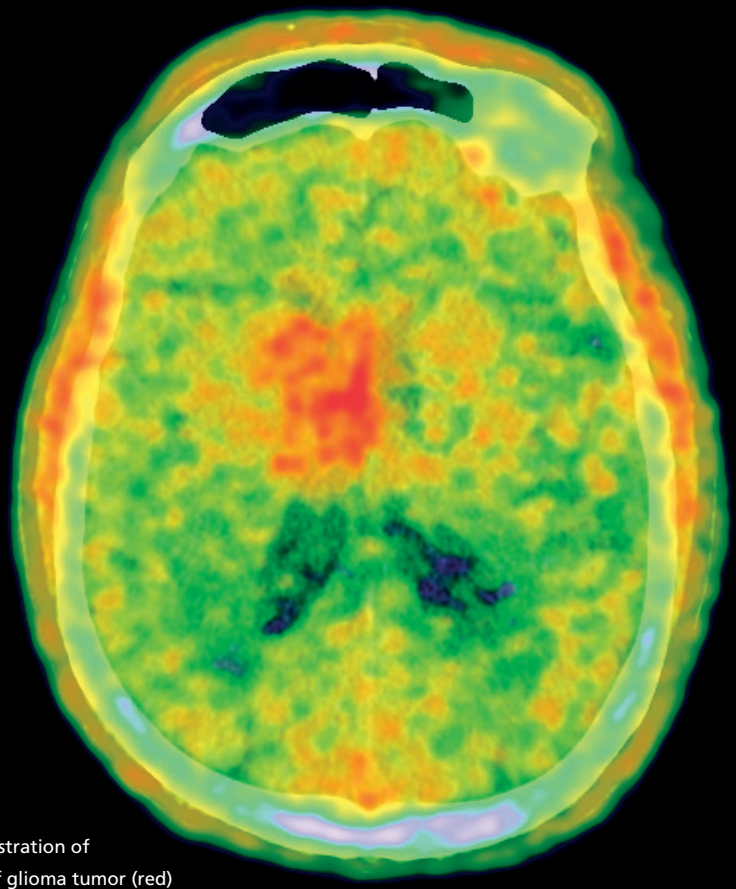


FIGURE 3. co-registration of FDG-PET and CT of glioma tumor (red) (http://www.osirix-viewer.com/datasets/).

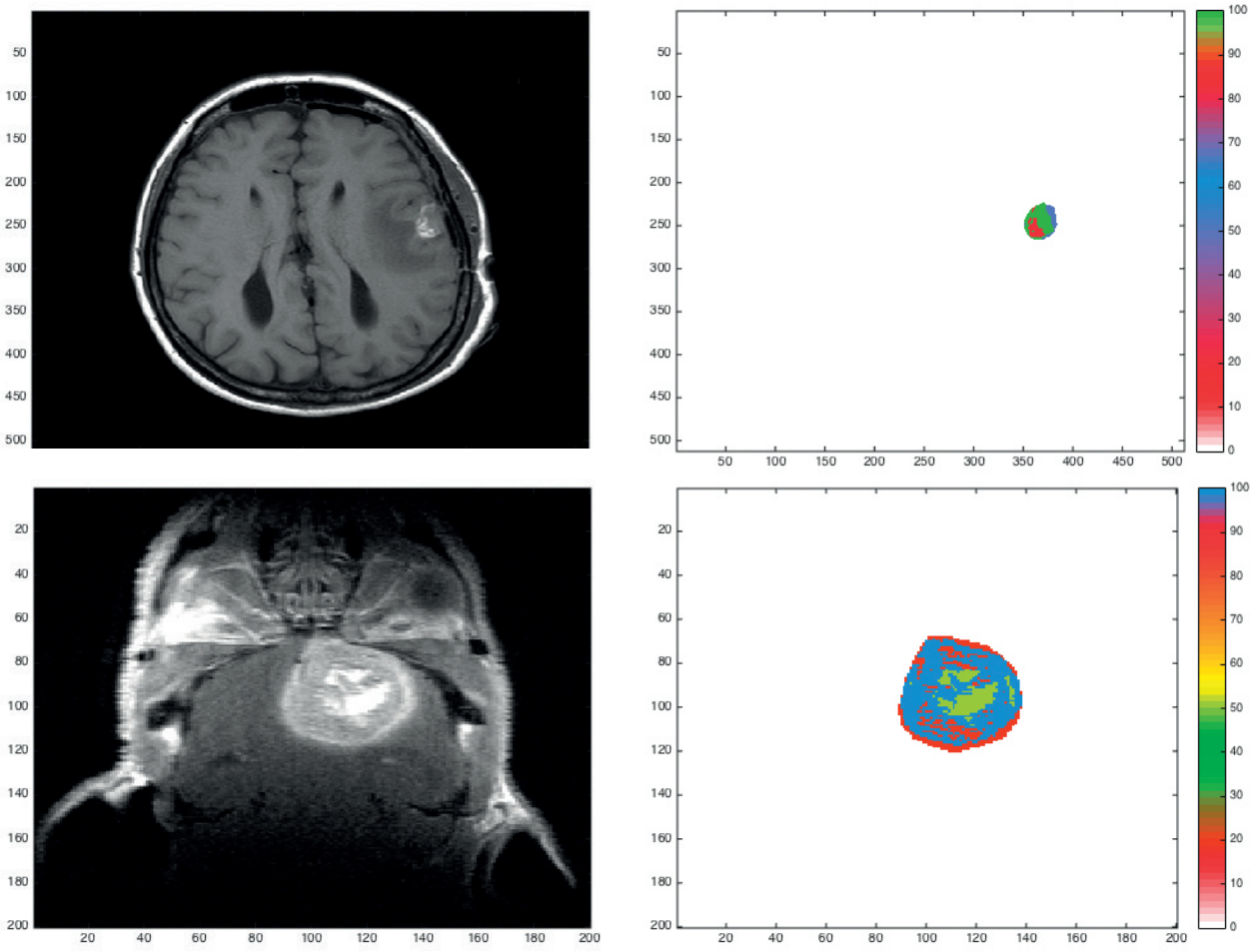


FIGURE 4. Tissue clustering (right side) of glioma tumor in human (top) and in mouse brain (bottom) visualized with MRI. Three types of tissues are identifiable: necrotic (green), proliferating (light blue), and highly-angiogenic tissue (red).

BOX 1. Linking cancer imaging to networks

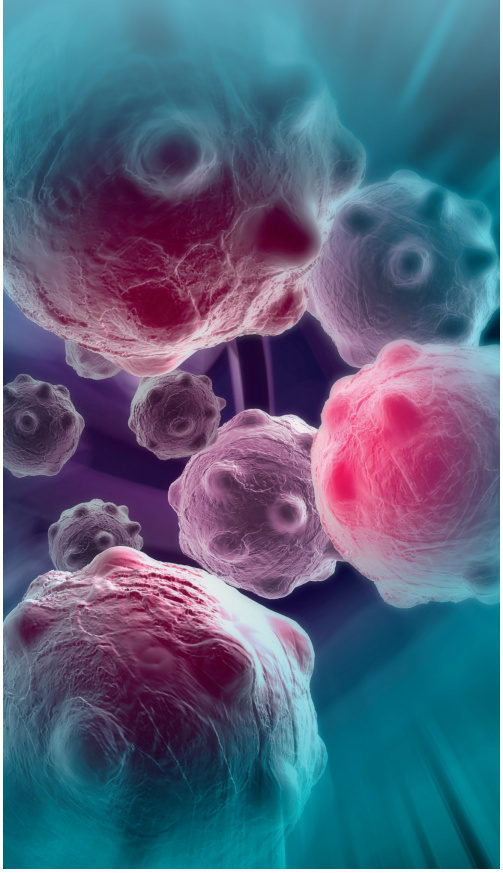
Tumor is a multifactorial disease characterized by substantial heterogeneity which multimodal image techniques can quantify.
Networks integrate images information encapsulated at a voxel level and transfer them from multiscale physical to multiscale computational settings.

Perturbations = Treatments, induce changes in network architecture, and identify:

- Treatment efficacy (tumor tissue)
- Side effects (healthy tissue)

Tumor Microenvironment:
Multiple tissues participate to form a network in which the metric identifies:

- **Short-distance interactions** involving immune-cells infiltration, vasculature, inflammation dynamics, etc.
- **Long-distance effects** involving metastases.



Cancer treatments can thus be seen as network perturbations, with the tumor-driven network as part of the hosting (organ-driven) network. The effects of any perturbation affect not only the tumor behavior but also the healthy tissue. Looking at a therapeutic perspective, we can use the modification caused in the network to monitor different types of effects, such as treatment efficacy in tumor regions and treatment toxicity in the healthy tissue (See Box 1).

Tumor habitats become particularly complex when microenvironment is considered, being the latter a context in which multiple tissue types interact by enabling processes such as inflammation, immune response and energy metabolism.[21] As previously mentioned, short and long distance interactions between cancer and stromal cells occur, depending on prevailing tumor and host environment programs (for instance, pro-invasion versus anti-metastasis, respectively) (see Box 1). These interactions can be naturally mapped onto networks and their spatiotemporal dynamics reconstructed due to the structural organization with built-in quantitative metrics.

Acknowledgement:
The authors thank their institutes for support.

Marco Dominietto, Ph.D Biomedical Engineering, is a medical physicist at the Biometerials Science Center of the University of Basel (CH). His research focuses on tumor development and image analysis both in mouse models and human beings. Among other sites, he worked at ETHZ (Zurich, Switzerland), CERN (Geneva, Switzerland), and University Hospital of Novara (Italy).

Enrico Capobianco, Ph.D Statistics, is Lead Scientist at the Center for Computational Science of the University of Miami (US), and associate with the Institute of Clinical Physiology of the National Research Council in Italy. PhD in Statistical Sciences, postdoc in computational science at Stanford Un. (US), CWI Amsterdam (NL), Boston Un. (US), among other sites. Recipient of CAS (China) and CAPES FIOCRUZ (Brazil) professorships, and visiting scientist of IHES (France). He has a vast international research experience in Cancer Genomics, Systems Medicine, Network Science, and Big Data.

References

1. Kaufman A, Cohen D, Yagel R. Volume graphics. IEEE Computer, 26(7):51–64:1993.
2. Kaufman, A. Efficient algorithms for 3d scan-conversion of parametric curves, surfaces, and volumes. In Proc ACM SIGGRAPH 1987, pg 171–179, USA, July 1987. ACM Press.
3. Kaufman A, Shimony E. 3d scan-conversion algorithms for voxel-based graphics. In Proc ACM Workshop on Interactive 3D Graphics, pg 45–76, Chapel Hill, NC, USA, October 1986. ACM Press.
4. Lucchi A, Smith K, Achanta R, Knott G, Fua P. Supervoxel-based segmentation of mitochondria in EM image stacks with learned shape features. IEEE Trans Med Imaging, 31(2): 474-86: 2012.
5. Meyer CR, Armato SG, Fenimore CP, McLennan G, Bidaut LM, Barboriak DP, Gavrielides MA, Jackson EF, McNitt-Gray MF, Kinahan PE, Petrick N, Zhao B. Quantitative imaging to assess tumor response to therapy: common themes of measurement, truth data and error sources. Translat Oncology 2(4): 198-210: 2009.
6. Dominietto M, Tsinoremas N, Capobianco E. Integrative analysis of cancer imaging readouts by networks. Molec Oncology, 9(1); 1-16, 2015.
7. Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson, MD, O'Connor L, Li M, Taylor R, Dharsee M, Ho Y, Heilbut A, Moore L, Zhang S,Ornatsky O, Bukhman YV, Ethier M, Sheng Y, Vasilescu J, Abu-Farha M, Lambert JP, Duewel HS, Stewart II, Kuehl B, Hogue K, Colwill K, Gladwish K, Muskat B, Kinach R, Adams SL, Moran MF, Morin GB, Topaloglou T, Figeys D. Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol. Syst. Biol. 3(89); 1-17: 2007.
8. Liu YY, Slotine JJ, Barabasi AL. Controllability of complex networks. Nature 473(7346); 167-73; 2011.
9. Serra-Musach J, Aguilar H, Iorio F, Comellas F, Berenguer A, Brunet J, Saez-Rodriguez J, Pujana MA. Cancer develops, progresses and responds to therapies through restricted perturbation of the protein-protein interaction network. Integr Biol 4(9): 1038-48: 2012.
10.Mora A, Taranta M, Zaki N, Cinti C, Capobianco E. Epigenetically driven network cooperativity: meta-analysis in multi-drug resistant osteosarcoma. J Compl Networks, 2015. doi: 10.1093/comnet/cnv017.
11. Davies PCW, Demetrius L., , Tuszynski JA. Cancer as a dynamical phase transition. Theor Biol Med Model. 2011; 8: 30.
12. Davila-Velderrain J, Martinez-Garcia JC, Alvarez-Buylla ER. Modeling the epigenetic attractors landscape: toward a post-genomic mechanistic understanding of development. Front Genet. 2015; 6: 160.
13. Gao J, Barzel B, Barabasi AL. Universal resilience patterns in complex networks. Nature 530; 307-12: 2016.
14. Hidalgo J, Grilli J, Suweis S, Munoz MA, Banavar JR, Maritan A. Information-based fitness and the emergence of criticality in living systems. PNAS 111(28); 10095-10100: 2014.
15. Cretich M, Daaboul GG, Sola L, Selim Unlu M, Chiari M. Digital detection of biomarkers assisted by nanoparticles: application to diagnostics. Cell 33(6); 343-51:2015.
16. Batmanghelich NK, Dalca AV, Sabuncu MR, Golland P, ADNI. Joint modeling of imaging and genetics.. Inf Process Med Imaging 23; 766-77:2013.
17.Alic L, Niessen, WJ, Veenland JF. Quantification of Heterogeneity as a Biomarker in Tumor Imaging: a systematic review. PLOS One 9(10), e110300: 2014.
18. Kalpathy-Cramer J, Freymann JB, Kirby JS, Kinahan PE, Prior FW. Quantitative Imaging Network: data sharing and competitive algorithm validation leveraging. The Cancer Imaging Archive. Translat Oncology 7(1); 147-152: 2014.
19. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K. et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Trans Med Imag 34(10); 1993-2024; 2015.
20. Gatenby RA, Grove O, Gillies RJ. Quantitative Imaging in Cancer Evolution and Ecology Radiology, 269(1); 2013.
21. Hainaut P, Plymoth A. Targeting the hallmarks of cancer: towards a rational approach to next-generation cancer therapy. Curr Opin Oncol 25(1):50-1: 2013.