UCLA
UCLA Previously Published Works

Title

"Radiotranscriptomics": A synergy of imaging and transcriptomics in clinical assessment

Permalink

Journal

ISSN

Authors

Katrib, Amal
Hsu, William
Bui, Alex
et al.

Publication Date

DOI

"RADIOTRANSCRIPTOMICS": A synergy of imaging and transcriptomics in clinical assessment

Amal Katrib[1,*], **William Hsu**[2], **Alex Bui**[2], and **Yi Xing**[1,*]

[1]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA

[2]Department of Radiological Sciences, University of California, Los Angeles, Los Angeles, CA 90095, USA

Abstract

Recent advances in quantitative imaging and "omics" technology have generated a wealth of mineable biological "big data". With the push towards a P4 "predictive, preventive, personalized, and participatory" approach to medicine, researchers began integrating complementary tools to further tune existing diagnostic and therapeutic models. The field of radiogenomics has long pioneered such multidisciplinary investigations in neuroscience and oncology, correlating genotypic and phenotypic signatures to study structural and functional changes in relation to altered molecular behavior. Given the innate dynamic nature of complex disorders and the role of environmental and epigenetic factors in pathogenesis, the transcriptome can further elucidate serial modifications undetected at the genome level. We therefore propose "radiotranscriptomics" as a new member of the P4 medicine initiative, combining transcriptome information, including gene expression and isoform variation, and quantitative image annotations.

Keywords

quantitative imaging; transcriptomics; RNA-seq; genomics; image genomics; radiogenomics; systems biology; precision medicine

SYSTEMS BIOLOGY IN MEDICINE

Diagnostic paradigms of the past century witnessed the prevalence of an "Oslerian formalism for human disease" methodology, a term coined by network theorist and complex systems pioneer AL Barabasi [1]. Scientific investigators often relied upon hypothesis-driven correlations between clinical manifestations and pathological findings, bringing rise to a reductionism philosophy in medicine. Reinforced by early successes in diagnosis of acute diseases, the relationship between the conventional medical practice and the reductionist mechanistic approach led to an oversimplification of the complex etiology of

diseases. It was not until the evaluation of cases of chronic disorders and complex diseases that clinicians became aware of the shortcomings of a discrete assessment of symptoms. Given our current understanding of the body as a complex system of interconnected biological networks, and of disease as a culmination of a multi-step modification of those networks, new paradigms are needed.

With the contemporary transition of biology from a qualitative, descriptive discipline into a quantitative, multi-parameter field, the power of computational methods has begun to unravel novel biology. The emergence of high-throughput technologies; greater computational power; and advances in imaging modalities, processing, and analysis have rendered a wealth of publically available biological "big data". This newfound capacity has promoted a paradigm shift in the early 21st century in which the biological science community began to embrace a more integrative attitude to medical research, using data to define and drive hypotheses. Propelled by the conviction that "every object that biology studies is a system of systems" [2], scientists began using tools from seemingly unrelated disciplines to study the impact of pathogenic factors on the genome, transcriptome, proteome, and metabolome. These efforts have been supported by numerous National Institutes of Health (NIH)-sponsored initiatives in the United States, giving rise to a collaborative community in which data and resources are easily shared and field-specific expertise continuously exchanged to drive advances in biomedical research [3].

GENOMICS and GWAS

Initially driven by the preeminent "common disease–common variant" hypothesis at the time, genome-wide association studies (GWAS) overtook the role of inheritance genetic linkage studies in assessing genotypic-phenotypic relationships [4,5]. Scientists meticulously examined the genome for recurrent variation patterns believed to be associated with the development of diseases and complex traits [6]. Single-nucleotide polymorphisms (SNPs), point mutations involving base substitutions in DNA sequences occurring in more than one percent of the general population, are the most common type of genetic variation marker in GWAS studies [7,8]. The identification of commonly occurring SNPs, along with the analysis of altered allelic frequency and distribution range of common variants across subpopulations, have been extensively used to assess the hereditary susceptibility to Mendelian disorders [7]. Given its comprehensiveness in surveying the whole genome, GWAS has proven promising in uncovering novel pathogenic candidate genes and potential therapeutic molecular targets. Nonetheless, the predictive power of commonly identified genomic variants remains poor as a result of the inability to determine truly causal relationships between such variants and disease susceptibility [9]. This is mostly attributed to a lower statistical significance when applying single-locus approaches to measure monogenic, single haplotype marginal effects on complex and aggressive disorders or rare traits [10]. The majority of large-scale GWAS approaches is also incapable of accounting for allelic heterogeneity, epigenetic and environmental bearing on gene expression, and gene-gene interactions or epistasis [10–13]. GWAS's contribution to elucidating the genetic basis of disease susceptibility is notable; however, the incorporation of functional annotations, interaction information, and multi-level molecular analyses is necessary for a more holistic genomic approach to disease stratification.

QUANTITATIVE MEDICAL IMAGING ANALYSIS

Medical imaging has long been a standard in diagnostic and therapeutic assessment and clinical outcome inference. Physicians often monitor morphologic, functional, molecular, metabolic, and micro-environmental changes *in vivo* in patients over time. Structural medical imaging modalities, such as X-ray, magnetic resonance imaging (MRI), and computed tomography (CT), are used to observe anatomical abnormalities whereas functional imaging modalities, such as functional-MRI (fMRI), positron emission tomography (PET), and single-photon emission computed tomography (SPECT), are reserved for assessment of physiological activity [14]. Melding complementing imaging modalities, modern multimodal devices and techniques now offer anatomic correlates to functional and molecular variations associated with pathogenesis [15,16].

Advances in imaging protocols, biomarker identification, and analysis tools have considerably improved prognosis of disease manifestation and progression. Standardization of imaging protocols, along with automated image registration, alignment, and segmentation and availability of reference atlases, has facilitated the transition of imaging from a largely qualitative tool to a robust quantitative measurement [17,18]. This transition process has been supported by the abundance of open source, user-friendly image analysis software such as the Insight Segmentation and Registration Toolkit (ITK) [19], FSL [20], 3D Slicer [21], Osirix [22], and Statistical Parametric Mapping [23]. Using those tools, quality control pre- and post-processing protocols can be applied to alleviate acquisition scheme- and operator-induced errors [24], and attempt to normalize imaging data across normal variations. For MR imaging, this includes distortions due to multi-center protocol differences, magnetic field inhomogeneity, motion and geometric artifacts, image intensity dropouts, misalignment and incorrect image registration, and eddy currents [24–26]. As for PET/CT multi-center cohort studies, errors can arise as a result of discrepancies in acquisition mode, scanner parameters, reconstruction filter kernel, and attenuation correction method [27,28]. Investigators then have a variety of manual, semi-automatic, and automatic methods to select from for quantitative analysis of the images [29]. A relatively simple approach, region-of-interest (ROI) analysis is used to extract signal intensity information within pre-defined anatomical regions [30]. This approach can be limited in its precision, namely when evaluating smaller areas, coverage, and time-consumption [31]. When conducting larger-scale exploratory studies, radiologists often employ automated voxel-based methods that carry statistical tests across image voxels to identify correlates to preselect covariates of interest [32]. Voxel-based morphometry (VBM) is widely used; however, its results are prone to misinterpretation due to misalignment, imperfect registration to standard space, and arbitrary spatial smoothing [32]. Tract-based spatial statistics (TBSS) has been proposed to address such concerns while equally maintaining the strengths of voxel- and tractography-based analyses [32]. This is achieved by carefully tuning registration and projecting onto an alignment-invariant tract representation [32]. The shift towards quantization in imaging has propelled the role of radiology in transforming patient care. Uniting clinicians, radiologists, and medical researchers, the Radiological Society of North America (RSNA) established the Quantitative Imaging Biomarkers Alliance (QIBA) to address aspects of image acquisition protocols, process standardization, quantitative data analysis, and biomarker development

[33]. A synergy of art and science, imaging analysis has now paved the way for both qualitative visualization and quantitative measurement of phenotypic features associated with many biological disorders.

More recently, researchers have placed a greater emphasis on generating extensive, mineable databases of radiological images and phenotypic features extracted and quantitatively analyzed in a high-throughput fashion [34,35]. The field was spurred by the observed heterogeneity in tumor microenvironment, metastatic nature of various cancers, and treatment responses. It employs automatic and semi-automatic segmentation and image trait extraction algorithms to generate individualized and targeted lists of imaging biomarkers [36]. Those features then undergo a conservative selection process in which only those that exhibit high specificity, stability, reproducibility, and information-to-count ratio are retained as potential phenotypic predictors.

GENOMICS and IMAGING META-ANALYSIS

The accumulation of a wealth of histopathologic, omic, and imaging data, in addition to electronic medical records, has fostered investigations in the assimilation of a variety of biological "big data" to better understand pathogenesis. A relatively new field, "radiogenomics" or "imaging genomics" emerged as part of the hybrid, multidisciplinary initiative to leverage and correlate phenotypic and genotypic traits of biological disorders [3,37,38].

Initially explored by neuroscientists, radiogenomic frameworks were utilized to study brain endophenotypes and decode neurodevelopment and mechanisms that mediate the complex traits of cognition, brain plasticity, and neurodegeneration. Amongst some of the collaborative neuroimaging genomics efforts in the field are the Enhancing NeuroImaging Genetics through Meta Analysis (ENIGMA) consortium [39], the Alzheimer's Disease Neuroimaging Initiative (ADNI) [40], and the Human Connectome Project [41]. By conducting GWAS, they are attempting to link SNP biomarkers and genetic variants with structural and functional neuroimaging phenotypes [39,42]. Integrating tools from genetics, "omics", and imaging, they aim to decipher the human brain and elucidate the intricate molecular underpinnings of neurodegeneration [39,43,44]. On a morphological level, many research groups have observed genetic influences on total intracranial and hippocampal volume [45], cortical thickness [46], subcortical structures [47], and white matter microstructure integrity [48]. To a lesser extent, genetic associations with physiological alterations have also been linked to human behavior and neurodegenerative disorders [49]. A few genetic studies have explored similar correlations with neuronal activation patterns [50], fractional anisotropy (FA) [51], and cerebral amyloid precursor protein accumulation [52].

Given the extent of processing information attainable through functional imaging, radiogenomics studies using PET and certain MR protocols offer the potential for a distinctive look into changes in the physiology and molecular activity during disease onset and progression [49]. Diffusion Tensor Imaging (DTI) MR and ^{18}F-FDG PET imaging annotations have been speculated to exhibit higher sensitivity to effects of genetic variants and as a result yield stronger association signals [53,54]. A recent study by the ENIGMA

consortium conducted a heritability analysis of tract-derived FA measures obtained from multi-site DTI MR brain images of healthy subjects [55]. FA scalar maps, which account for the directional preference of water diffusion and fiber coherence, are extensively used to infer the state of white matter (WM) microstructure integrity in neurodegenerative diseases [55,56]. They generated a pipeline composed of a standardized high-resolution FA-based registration space and WM skeleton to address inter-site variability in protocol administration, operation, and measurement. Using voxel-wise heritability analysis, they extracted FA phenotypes with high levels of heritability. Through imaging protocol harmonization, they proved successful at laying a strong foundation for future multi-site GWAS endeavors and rendering FA measures a more robust candidate endophenotype.

Scalable in its nature, radiogenomics is now being pursued in oncology, as the field applies a personalized genomic- and environmentally-centered approach to identify prognostic and predictive tumorigenesis biomarkers [3]. The network-scale rewiring of signal transduction and metabolic pathways during carcinogenesis has led to the hypothesis that genomic alterations foster adaptation to varying selection pressures, enabling cancer cells to evolve in a Darwinian fashion and as result prevail chemo and radiation treatment [57]. Mapping altered molecular behavior onto cancer imaging features enables noninvasive assessment of the underlying molecular machinery, providing exhaustive spatial and temporal coverage of tumor development, metastasis, and heterogeneity [58]. Research groups have already investigated genotypic associations with functional measures such as 18[F]-2-fluorodeoxyglucose (FDG) uptake [54] and morphologic traits such as non-uniform enhancement patterns and architectural distortion [59], cerebral edema and cellular invasion [60], and intratumoral vasculature and tumor margin definition [61].

TRANSCRIPTOMICS

Transcriptome sequencing and expression profiling have laid the foundations for studies of the transcriptional structure of genes. By cataloguing and mapping the transcriptome, researchers have been able to quantitatively evaluate altered expression levels during development and pathogenesis [62]. Analyzing total RNA derived from protein-coding and non-coding genes, which actively reflects time-resolved gene expression, offers a deeper understanding of the inner workings of a cell. This has rendered transcriptomics crucial in disease investigation, readily competing with other "omic" tools (Figure 1). In addition to providing a thorough snapshot of total gene activity, transcriptomics is impermeable to limitations of proteomics, including a large domain size, difficulty amplifying and detecting low-abundance proteins, and proteins' dynamic nature [63,64]. It also outperforms metabolomics in simplicity, maturity, and depth, despite being unable to relay cellular biochemical activity [65].

Various hybridization- and sequencing-based methods have been developed to profile gene expression and quantify the transcriptome. Notwithstanding the popularity of microarray technology, high-throughput next generation sequencing (NGS) is surpassing Moore's law predictions, rapidly becoming the platform of choice for transcriptional profiling [66]. Hybridization-based methods are limited by the introduction of individual bias due to reliance upon prior knowledge of the transcriptome. Restricted to closed systems, only

mRNA with corresponding homologous printed probes can be measured and only known transcripts are accounted for [67,68]. As a result of high noise and saturation background levels and cross-hybridization, microarrays are also constrained in detection dynamic range and suffer from quantitative inaccuracies due to nonlinear dye response [69]. Furthermore, intra- and inter-tumor diversity raises the concern of sampling errors when using microarray methods in oncology studies [70]. RNA-seq is therefore poised to address limitations of microarray-based investigations. Albeit with higher costs and further analytical complexity, RNA-seq proves superior for gene network construction [66]. By aligning reads against splice junctions, it can identify isoforms, gene fusions, alternative splicing sites, and post-transcriptional RNA editing events with greater sensitivity and higher spatio-temporal resolution [71]. Able to conduct a robust, limitless, and unbiased genome-wide survey of the transcriptome, RNA-seq can quantify transcripts with over six orders of magnitude of dynamic range [66,72]. With the exponential decrease in sequencing costs and advances in sequence detection methods, library preparation protocols, and multiplexing capabilities, NGS-based RNA-seq has revolutionized gene expression profiling studies, translating into countless novel discoveries of pathogenic biomarkers.

Quantitative analysis of the transcriptome by RNA-seq encompasses many layers that extend beyond the mere evaluation of gene expression levels. Investigators often aim to detect novel splice junctions, transcripts, and fusion genes and identify translocation, differential alternative splicing, and post-transcriptional events as they relate to the evolution of complex traits and pathogenesis [73]. Gene expression levels are estimated by mapping RNA-seq reads against a reference genome or transcriptome. By comparing expression levels across various conditions, researchers can then extract key differential expression genes and embed results in a network framework for a systems evaluation of gene-pathway interactions. Assessment of alternative splicing events and mRNA isoform expression levels has also become an emergent field of interest when analyzing RNA-seq data. The role of pre-mRNA alternative splicing, as well as other types of transcript isoform variation in modulating gene expression and consequently rendering diverse transcriptome and proteome populations, has elicited a belief in a direct causal relationship between isoform-level dysregulation and disease susceptibility and initiation [74]. The basic modes of alternative splicing include (Figure 2): (i) exon skipping, (ii) mutually exclusive exon usage, (iii) alternative 5′ donor or 3′ acceptor splice sites, and (iv) intron retention [75]. Additionally, alternative initiation and alternative polyadenylation, sometimes coupled with alternative splicing, are two other common sources of transcript isoform diversity. Through varying combinations of those modes, alternative splicing and transcript isoform processing can dynamically regulate gene function, spatially and temporally accounting for environmental impacts on phenotypic development [75].

RNA-seq technology has opened up diagnostic applications beyond the purview of microarrays, providing information on disease-specific gene expression, RNA variants, and fusions at a fraction of the cost of whole-genome sequencing [76]. However, its transition to the clinical practice is hindered by the prerequisite for standardization and establishment of benchmark datasets to assess analytical sensitivity, specificity, accuracy, and reproducibility [77]. Researchers are in the process of investigating potential sources of bias across sites, including library preparation methods, sample collection, and sequencing platform selection,

to develop standardized metrics that can minimize artifacts and false discoveries. The development of rigorous protocols that can account and correct for such biases and perform cross-validation of results is critical in a clinical sequencing workflow. Additionally, clinical protocols will also have to be optimized to account for low-quality and low-abundance specimens [76]. A rapid turnaround in the delivery of analysis results, along with the maintenance of a reasonable cost of operation and support from insurance companies, will also have to be met prior to the wide adoption of RNA-seq in clinical settings [76].

"RADIOTRANSCRIPTOMICS"

Currently, the term "radiogenomics" has been used broadly to refer to studies that correlate imaging data with diverse types of omics data. However, in its strictest sense "radiogenomics" should address specifically variations at the DNA level. Given the extent of available "omic" data and the variety of information provided through each type, transcriptomic data can be combined with imaging to offer further insight into the molecular intricacies of diseases. Studying the transcriptome allows us to venture into the intermediate stage from gene to protein, providing functional context to key genes, along with regulatory mechanisms through which they confer selective expression variations in pathogenesis. We therefore propose a more specific term "radiotranscriptomics" as the newest member of expanding efforts in omics-phenotype data integration (Figure 3).

A few currently-existing efforts have begun exploring the integration of microarray expression profiling data with imaging features to non-invasively study the molecular characteristics of various tumor types and predict clinical outcome. A study conducted by a group in Stanford investigated oncogenomic correlates of non-small cell lung cancer (NSCLC) by linking FDG uptake PET imaging features to genome-wide expression signatures [54]. A prognostic multivariate FDG uptake model was generated by associating single gene and co-expressed gene clusters with various standard uptake values (SUVs) and survival information. The heterogeneity in tissue-specific uptake of FDG, an established surrogate for glycolysis, was observed to be a comprehensive transcriptomic-level marker for dysregulated cellular bioenergetics in tumorigenesis. Another study explored the association of MR volumetrics with mRNA and microRNA (miRNA) expression levels in glioblastoma multiforme (GBM) [60]. They measured fluid attenuation inversion recovery (FLAIR) signal abnormalities in relation to edema and cellular migration and correlated select radiophenotypes with gene expression. By means of Ingenuity Pathway Analysis (IPA), they identified top concordant genes and molecular pathways that can potentially serve as diagnostic determinants of cancer invasion and metastasis. Another recent study by Aerts *et al.* examined the association of a pre-defined list of quantitative imaging annotations with gene expression profiles of lung cancer patients [35]. Through gene-set enrichment analysis (GSEA), they unveiled a coupling between several oncogenic cell cycle and proliferation pathways and four such features. Given the abundant concern of false positive gene hits as a result of the probe-dependent nature of microarrays, we believe RNA-seq technology to be extremely promising in its ability to improve upon existing results in the field of "radiotranscriptomics".

A working "radiotranscriptomics" pipeline will involve the (i) selection of an adequate mRNA profiling dataset and transcript analyses method, (ii) extraction of quantitative image annotations, (iii) data sharing and harmonization, (iv) application of appropriate statistical and analysis algorithms for association studies, (v) visual representation and predictive modeling, and (vi) potential integration of additional molecular and clinical variables. The Cancer Genome Atlas (TCGA), in connection with The Cancer Imaging Archive (TCIA), creates an effective dataset repository for exploratory studies [78]. It offers a wide range of publically accessible microarray and RNA-seq mRNA and miRNA gene expression data from surgically resected tumors, along with corresponding patient radiological information [78,79]. Analysis of total gene expression, isoform variation, and alternative splicing events can be interchangeably pursued to account for RNA's complex nature. A subset within the associated imaging dataset should then be selected, varying the choice of modality per interest in either morphologic or physiologic phenotype extraction. Preference should be given to protocols that render high reproducibility, reliability, and accuracy, as well as scans of ample sample size and sufficient pre- and post-treatment/resectioning time incremental measures. Following that, a controlled and comprehensive set of imaging features should be defined based on prominence, reproducibility, and independence from other traits [36]. Validation by domain experts will help moderate inhomogeneity and obfuscation in quantitative analysis [80]. Collaborations across multiple cohorts are encouraged to ensure sufficient power, statistical significance, and credibility for the proposed association studies. As a result, it is crucial to harmonize the process by establishing a regimented protocol that addresses multi-site acquisition and patient population variability. This involves image-space registration along with the use of consistent sequence alignment methods to map transcriptomic reads and data analysis tools to limit spurious positives. Afterwards, a quantitative statistical analysis scheme (e.g., weighted gene co-expression network analysis [81]) should be employed to elucidate a predictive relationship between gene expression/ isoform variation and phenotype. Investigators can refer to their research goals to guide their selection of correlation linkage, association, or regression estimation methods to obtain strong signals for single gene, gene set, and network analyses. In the process, additional mathematical methods will be needed to mitigate limitations of data dimensionality variations and insufficient "good data" sample sizes. Researchers will have the opportunity to consolidate supplementary clinical and molecular signature levels within patient records to improve upon their predictive "radiotranscriptomic" model. Finally, validation of results will have to be performed to differentiate true associations from false discoveries. This can be done orthogonally through basic experimental science bench work. Replication studies using independent datasets can also be used to validate findings.

"RADIOTRANSCRIPTOMICS" CHALLENGES

"Radiotranscriptomics" has all the ingredients fitted to explore new biomedical questions and complements other methodologies in clinical assessment. Nonetheless, it is noteworthy to bring attention to limitations and workarounds potentially associated with such an approach. First and foremost, the field's greatest challenge is the limited number of existing, well-documented datasets that include concurrent RNA-seq and medical imaging data. We anticipate this will become less of a concern in the near future, in view of the increased

interest in biomedical Big Data and allocation of funds for cross-disciplinary and multi-omic research.

Moreover, the unsupervised nature of quantitative association studies makes it difficult to distinguish true biological signals from artifacts due to technical variability and confounding factors. Technical variability can occur throughout the various stages of a "radiotranscriptomics" workflow: study design and protocol, time points of data collection, RNA-seq library preparation, image acquisition and processing, to name a few. Confounding factors can include changes in gene expression due to physical environment, genetic or demographic variables such as population stratification, and inherent variations between subjects arising from age, gender, or other features. Careful study design and data analysis strategies are needed to minimize the effect of technical artifacts and confounding factors. We should also note that the association studies between transcriptome and imaging data will reveal correlation but not causality. Causal inference in "radiotranscriptomics" will nonetheless continue to be challenging, as in other areas of transcriptome-phenotype mapping efforts.

In addition, the majority of collected RNA-seq data are from analyses at the cell population level. This can fail to account for intratumoral heterogeneity, in which apparently identical cells can still harbor varying degrees of mutation and consequently promote an inconsistent behavioral pattern of disease progression and recurrence and treatment response [82]. Single-cell RNA-seq can help mitigate this shortcoming, identifying cell-type specific characteristics, albeit with higher sampling errors and technical noise.

SUMMARY

Given the temporal and functional nature of transcriptome-level analysis, "radiotranscriptomic" models have an edge over other data-driven frameworks in disease risk stratification and clinical assessment. By harnessing RNA-mediated regulation of gene expression, the field can appropriately capture the transcriptional state of a cell, offering a keener insight into the current underlying molecular and functional state. It also provides information on patient-specific epigenetic and environmental modifications believed to be major contributors to the acquisition of pathogenic traits. Accounting for the intrinsic dynamic nature of biological systems and the relationship between molecular, functional, and anatomical stages, "radiotranscriptomics" has naturally lent itself to investigations in complex, multifactorial disorders such as cancer, neurodegeneration, and autoimmune diseases (Figure 4).

Medical imaging is often used to grossly characterize underlying molecular features. Nonetheless, "radiotranscriptomics" can also pave the way for mining the reverse relationship, in which transcriptomic signatures serve as input to predicting imaging annotations. Imaging offers an aggregate representation of the underlying function. Studying the microenvironment of a transcriptome can elucidate the associated physiology and help guide clinicians in scouting for pathogenic phenotypes and generating all-inclusive lists of imaging features. Furthermore, medical imaging continues to lag behind in its ability to adequately capture biomarkers, namely in cases of early disease onset or multi-faceted

disorders that lack a standard trend. By studying thematic expression patterns in relation to observed regional physical properties, researchers are better equipped to design novel imaging protocols that can specifically target such signatures.

Moreover, the translational capabilities of "radiotranscriptomics" render it an equally informative tool for both researchers and clinicians. Research studies are typically meticulous in following controlled imaging protocols and quality control procedures when using data across multiple sites. However, image acquisition parameters are constrained by local settings and scanner capabilities, therefore making it difficult to standardize the parameters and subsequent image. The application of an endophenotypic approach to disease assessment and stratification has the potential to overcome imaging-only biomarkers that are prone to this issue. The integration of multi-level biological data, including prognostic quantitative imaging biomarkers, radiogenomic and radiotranscriptomic signatures, and clinical variables, can generate more robust, precise descriptors as input to personalized patient plans. Through "radiotranscriptomics", specifically, clinicians can further tune treatment regimens by building temporal models to match time-series transcriptomic information with serial imaging and patient record observations.

The potential applications of "radiotranscriptomics" extend far beyond the immediate rejoice of joining many systems biology efforts in the integration of "big data" in medicine. Through a Venn diagram synergistic coupling of molecular indexes from transcriptomics, phenotypic traits from imaging, and clinical data from medical records (Figure 5), the field offers the potential to radically transform the face of modern medicine. "Radiotranscriptomics", thus, poses as a promising contributor to the P4 "predictive, preventive, personalized, and participatory" and precision medicine initiatives, promoting both proactive and reactive measures through disease detection in pre-symptomatic phases and tailored targeting of therapeutic plans.

Acknowledgments

REFERENCES

1. Loscalzo J, Barabasi AL. Systems biology and the future of medicine. Wiley Interdiscip. Rev. Syst. Biol. Med. 2011; 3:619–627. [PubMed: 21928407]

2. Trewavas A. A Brief History of Systems Biology: "Every object that biology studies is a system of systems" Francois Jacob (1974). Plant Cell Online. 2006; 18:2420–2430.

3. Jaffe CC. Imaging and genomics: is there a synergy? Radiology. 2012; 264:329–331. [PubMed: 22821693]

4. Lander ES. The new genomics: global views of biology. Science. 1996; 274:536–539. [PubMed: 8928008]

5. Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant…or not? Hum. Mol. Genet. 2002; 11:2417–2423. [PubMed: 12351577]

6. Strachan, T., Read, AP., Strachan, T. Human Molecular Genetics. 4th. New York: Garland Science; 2011.

7. Bush WS, Moore JH. Chapter 11: Genome-Wide Association Studies. PLoS Comput. Biol. 2012; 8:e1002822. [PubMed: 23300413]

8. Moody, G. Digital code of life: how bioinformatics is revolutionizing science, medicine, and business. Hoboken: Wiley; 2004.

9. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am. J. Hum. Genet. 2012; 90:7–24. [PubMed: 22243964]

10. Li J, Horstman B, Chen Y. Detecting epistatic effects in association studies at a genomic level based on an ensemble approach. Bioinformatics. 2011; 27:i222–i229. [PubMed: 21685074]

11. Iles MM. What can genome-wide association studies tell us about the genetics of common disease? PLoS Genet. 2008; 4:e33. [PubMed: 18454206]

12. George AL Jr. Appraising the value of genomic association studies. J. Am. Soc. Nephrol. 2008; 19:1840–1842. [PubMed: 18632845]

13. Gibson G. Rare and common variants: twenty arguments. Nat. Rev. Genet. 2012; 13:135–145. [PubMed: 22251874]

14. Bushberg, JT. The essential physics of medical imaging. 3rd. Philadelphia: Wolters Kluwer Health/ Lippincott Williams & Wilkins; 2012.

15. Martí-Bonmatí L, Sopena R, Bartumeus P, Sopena P. Multimodality imaging techniques. Contrast Media Mol. Imaging. 2010; 5:180–189. [PubMed: 20812286]

16. Padhani AR, Miles KA. Multiparametric imaging of tumor response to therapy. Radiology. 2010; 256:348–364. [PubMed: 20656830]

17. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 2002; 17:825–841. [PubMed: 12377157]

18. Woods RP, Mazziotta JC, Cherry SR. MRI-PET registration with automated algorithm. J. Comput. Assist. Tomogr. 1993; 17:536–546. [PubMed: 8331222]

19. Martin K, Ibáñez L, Avila L, Barré S, Kaspersen JH. Integrating segmentation methods from the Insight Toolkit into a visualization application. Med. Image Anal. 2005; 9:579–593. [PubMed: 16185910]

20. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. Fsl. Neuroimage. 2012; 62:782–790. [PubMed: 21979382]

21. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn. Reson. Imaging. 2012; 30:1323–1341. [PubMed: 22770690]

22. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. J. Digit. Imaging. 2004; 17:205–216. [PubMed: 15534753]

23. Friston KJ. Commentary and opinion: II. Statistical parametric mapping: ontology and current issues. J. Cereb. Blood Flow Metab. 1995; 15:361–370. [PubMed: 7713993]

24. Filippi M, Horsfield MA, Adèr HJ, Barkhof F, Bruzzi P, Evans A, Frank JA, Grossman RI, McFarland HF, Molyneux P, et al. Guidelines for using quantitative measures of brain magnetic resonance imaging abnormalities in monitoring the treatment of multiple sclerosis. Ann. Neurol. 1998; 43:499–506. [PubMed: 9546332]

25. Netsch T, van Muiswinkel A. Quantitative evaluation of image-based distortion correction in diffusion tensor imaging. IEEE Trans. Med. Imaging. 2004; 23:789–798. [PubMed: 15250631]

26. Hutton C, Bork A, Josephs O, Deichmann R, Ashburner J, Turner R. Image distortion correction in fMRI: A quantitative evaluation. Neuroimage. 2002; 16:217–240. [PubMed: 11969330]

27. Gallardo-Estrella L, Lynch DA, Prokop M, Stinson D, Zach J, Judy PF, van Ginneken B, van Rikxoort EM. Normalizing computed tomography data reconstructed with different filter kernels: effect on emphysema quantification. Eur. Radiol. 2015; 26:478–486. [PubMed: 26002132]

28. Fahey FH, Kinahan PE, Doot RK, Kocak M, Thurston H, Poussaint TY. Variability in PET quantitation within a multicenter consortium. Med. Phys. 2010; 37:3660–3666. [PubMed: 20831073]

29. Snook L, Plewes C, Beaulieu C. Voxel based versus region of interest analysis in diffusion tensor imaging of neurodevelopment. Neuroimage. 2007; 34:243–252. [PubMed: 17070704]

30. Poldrack RA. Region of interest analysis for fMRI. Soc. Cogn. Affect. Neurosci. 2007; 2:67–70. [PubMed: 18985121]

31. Park HJ, Kubicki M, Shenton ME, Guimond A, McCarley RW, Maier SE, Kikinis R, Jolesz FA, Westin C-F. Spatial normalization of diffusion tensor MRI using multiple channels. Neuroimage. 2003; 20:1995–2009. [PubMed: 14683705]

32. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage. 2006; 31:1487–1505. [PubMed: 16624579]

33. Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC. A collaborative enterprise for multi-stakeholder participation in the advancement of quantitative imaging. Radiology. 2011; 258:906–914. [PubMed: 21339352]

34. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJWL, Dekker A, Fenstermacher D, et al. Radiomics: the process and the challenges. Magn. Reson. Imaging. 2012; 30:1234–1248. [PubMed: 22898692]

35. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat. Commun. 2014; 5:4006. [PubMed: 24892406]

36. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, Zegers CML, Gillies R, Boellard R, Dekker A, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur. J. Cancer. 2012; 48:441–446. [PubMed: 22257792]

37. Medland SE, Jahanshad N, Neale BM, Thompson PM. Whole-genome analyses of whole-brain data: working within an expanded search space. Nat. Neurosci. 2014; 17:791–800. [PubMed: 24866045]

38. Colen R, Foster I, Gatenby R, Giger ME, Gillies R, Gutman D, Heller M, Jain R, Madabhushi A, Madhavan S, et al. NCI Workshop Report: Clinical and computational requirements for correlating imaging phenotypes with genomics signatures. Transl. Oncol. 2014; 7:556–569. [PubMed: 25389451]

39. Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME, Toro R, Jahanshad N, Schumann G, Franke B, et al. large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging Behav. 2014; 8:153–182. [PubMed: 24399358]

40. Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, L Whitwell J, Ward C, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J. Magn. Reson. Imaging. 2008; 27:685–691. [PubMed: 18302232]

41. Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, Chang A, Chen L, Corbetta M, Curtiss SW, et al. The Human Connectome Project: a data acquisition perspective. Neuroimage. 2012; 62:2222–2231. [PubMed: 22366334]

42. van Erp, TGM., Cannon, TD., Tran, HL., Wobbekind, AD., Huttunen, M., Lonnqvist, J., Kaprio, J., Salonen, O., Valanne, L., Poutanen, V-P. Genetic influences on human brain morphology; Biomedical Imaging: Nano to Macro, IEEE International Symposium; 2004. p. 583-586.

43. McIntosh A, Deary I, Porteous DJ. Two-back makes step forward in brain imaging genomics. Neuron. 2014; 81:959–961. [PubMed: 24607220]

44. Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S, Risacher SL, Nho K, Huentelman MJ, Craig DW, et al. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. Alzheimers Dement. 2010; 6:265–273. [PubMed: 20451875]

45. Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM, Toro R, Appel K, Bartecek R, Bergmann Ø, et al. Identification of common variants associated with human hippocampal and intracranial volumes. Nat. Genet. 2012; 44:552–561. [PubMed: 22504417]

46. Kochunov P, Glahn DC, Nichols TE, Winkler AM, Hong EL, Holcomb HH, Stein JL, Thompson PM, Curran JE, Carless MA, et al. Genetic analysis of cortical thickness and fractional anisotropy of water diffusion in the brain. Front. Neurosci. 2011; 5:120. [PubMed: 22028680]

47. Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivières S, Jahanshad N, Toro R, Wittfeld K, Abramovic L, Andersson M, et al. Common genetic variants influence human subcortical brain structures. Nature. 2015; 520:224–229. [PubMed: 25607358]

48. Sprooten E, Fleming KM, Thomson PA, Bastin ME, Whalley HC, Hall J, Sussmann JE, McKirdy J, Blackwood D, Lawrie SM, et al. White matter integrity as an intermediate phenotype: Exploratory genome-wide association analysis in individuals at high risk of bipolar disorder. Psychiatry Res. 2013; 206:223–231. [PubMed: 23218918]

49. Hariri AR, Weinberger DR. Imaging genomics. Br. Med. Bull. 2003; 65:259–270. [PubMed: 12697630]

50. van den Heuvel MP, van Soelen ILC, Stam CJ, Kahn RS, Boomsma DI, Hulshoff Pol HE. Genetic control of functional brain network efficiency in children. Eur. Neuropsychopharmacol. 2013; 23:19–23. [PubMed: 22819192]

51. Kochunov P, Jahanshad N, Marcus D, Winkler A, Sprooten E, Nichols TE, Wright SN, Hong LE, Patel B, Behrens T, et al. Heritability of fractional anisotropy in human white matter: a comparison of Human Connectome Project and ENIGMA-DTI data. Neuroimage. 2015; 111:300–311. [PubMed: 25747917]

52. Ramanan VK, Risacher SL, Nho K, Kim S, Shen L, McDonald BC, Yoder KK, Hutchins GD, West JD, Tallman EF, et al. GWAS of longitudinal amyloid accumulation on ^{18}F-florbetapir PET in Alzheimer's disease implicates microglial activation gene *IL1RAP*. Brain. 2015; 138:3076–3088. [PubMed: 26268530]

53. Jahanshad, N., Kohannim, O., Toga, AW., McMahon, KL., de Zubicaray, GI., Hansell, NK., Montgomery, GW., Martin, NG., Wright, MJ., Thompson, PM. Diffusion Imaging Protocol Effects on Genetic Associations; Proc. IEEE Int. Symp. Biomed Imaging; 2012. p. 944-947.

54. Nair VS, Gevaert O, Davidzon G, Napel S, Graves EE, Hoang CD, Shrager JB, Quon A, Rubin DL, Plevritis SK. Prognostic PET 18F-FDG uptake imaging features are associated with major oncogenomic alterations in patients with resected non-small cell lung cancer. Cancer Res. 2012; 72:3725–3734. [PubMed: 22710433]

55. Jahanshad N, Kochunov PV, Sprooten E, Mandl RC, Nichols TE, Almasy L, Blangero J, Brouwer RM, Curran JE, de Zubicaray GI, et al. Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: a pilot project of the ENIGMA-DTI working group. Neuroimage. 2013; 81:455–469. [PubMed: 23629049]

56. Thomason ME, Thompson PM. Diffusion imaging, white matter, and psychopathology. Annu. Rev. Clin. Psychol. 2011; 7:63–85. [PubMed: 21219189]

57. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012; 481:306–313. [PubMed: 22258609]

58. Chowdhury R, Ganeshan B, Irshad S, Lawler K, Eisenblätter M, Milewicz H, Rodriguez-Justo M, Miles K, Ellis P, Groves A, et al. The use of molecular imaging combined with genomic techniques to understand the heterogeneity in cancer metastasis. Br. J. Radiol. 2014; 87:20140065. [PubMed: 24597512]

59. Yamamoto S, Maki DD, Korn RL, Kuo MD. Radiogenomic analysis of breast cancer using MRI: a preliminary study to define the landscape. AJR Am. J. Roentgenol. 2012; 199:654–663. [PubMed: 22915408]

60. Zinn PO, Majadan B, Sathyan P, Singh SK, Majumder S, Jolesz FA, Colen RR. Radiogenomic mapping of edema/ cellular invasion MRI-phenotypes in glioblastoma multiforme. PLoS One. 2011; 6:e25451. [PubMed: 21998659]

61. Karlo CA, Di Paolo PL, Chaim J, Hakimi AA, Ostrovnaya I, Russo P, Hricak H, Motzer R, Hsieh JJ, Akin O. Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations. Radiology. 2014; 270:464–471. [PubMed: 24029645]

62. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 2009; 10:57–63. [PubMed: 19015660]

63. Dove A. Proteomics: translating genomics into products? Nat. Biotechnol. 1999; 17:233–236. [PubMed: 10096288]

64. Smaczniak C, Li N, Boeren S, America T, van Dongen W, Goerdayal SS, de Vries S, Angenent GC, Kaufmann K. Proteomics-based identification of low-abundance signaling and regulatory protein complexes in native plant tissues. Nat. Protoc. 2012; 7:2144–2158. [PubMed: 23196971]

65. Fiehn O. Metabolomics—the link between genotypes and phenotypes. Plant Mol. Biol. 2002; 48:155–171. [PubMed: 11860207]

66. Hitzemann R, Bottomly D, Darakjian P, Walter N, Iancu O, Searles R, Wilmot B, McWeeney S. Genes, behavior and next-generation RNA sequencing. Genes Brain Behav. 2013; 12:1–12. [PubMed: 23194347]

67. Steger D, Berry D, Haider S, Horn M, Wagner M, Stocker R, Loy A. Systematic spatial bias in DNA microarray hybridization is caused by probe spot position-dependent variability in lateral diffusion. PLoS One. 2011; 6:e23727. [PubMed: 21858215]

68. Hartl, DL., Ruvolo, M. Genetics : analysis of genes and genomes. 8th. Burlington: Jones & Bartlett Learning; 2012.

69. Koltai H, Weingarten-Baror C. Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. Nucleic Acids Res. 2008; 36:2395–2405. [PubMed: 18299281]

70. Barry WT, Kernagis DN, Dressman HK, Griffis RJ, Hunter JD, Olson JA, Marks JR, Ginsburg GS, Marcom PK, Nevins JR, et al. Intratumor heterogeneity and precision of microarray-based predictors of breast cancer biology and clinical outcome. J. Clin. Oncol. 2010; 28:2198–2206. [PubMed: 20368555]

71. Sugano S. Introduction: next-generation DNA sequencing and bioinformatics. Tanpakushitsu kakusan koso. Protein, nucleic acid, enzyme. 2009; 54:1233–1237.

72. Coppola, G. The OMICs : applications in neuroscience. New York: Oxford University Press; 2014.

73. Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, Mason CE, Socci ND, Betel D. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol. 2013; 14:R95. [PubMed: 24020486]

74. Lu ZX, Jiang P, Xing Y. Genetic variation of pre-mRNA alternative splicing in human populations. Wiley Interdiscip. Rev. RNA. 2012; 3:581–592. [PubMed: 22095823]

75. Liu SL, Cheng CH. Alternative RNA splicing and cancer. Wiley Interdiscip. Rev. RNA. 2013; 4:547–566. [PubMed: 23765697]

76. Damodaran S, Berger MF, Roychowdhury S. Clinical tumor sequencing: opportunities and challenges for precision cancer medicine. Am. Soc. Clin. Oncol. Educ. Book. 2015; 35:e175–e182.

77. Van Keuren-Jensen K, Keats JJ, Craig DW. Bringing RNA-seq closer to the clinic. Nat. Biotechnol. 2014; 32:884–885. [PubMed: 25203037]

78. Gonzalez-Angulo AM, Hennessy BT, Mills GB. Future of personalized medicine in oncology: a systems biology approach. J. Clin. Oncol. 2010; 28:2777–2783. [PubMed: 20406928]

79. Prior FW, Clark K, Commean P, Freymann J, Jaffe C, Kirby J, Moore S, Smith K, Tarbox L, Vendt B, et al. TCIA: An information resource to enable open science. Conf. Proc. IEEE Eng. Med. Biol. Soc. 2013; 2013:1282–1285. [PubMed: 24109929]

80. Gutman DA, Cooper LAD, Hwang SN, Holder CA, Gao JJ, Aurora TD, Dunn WD Jr, Scarpace L, Mikkelsen T, Jain R, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set. Radiology. 2013; 267:560–569. [PubMed: 23392431]

81. Zhang B, Horvath S. Ageneral framework for weighted gene co-expression network analysis. Stat. Appl. Genet. Mol. Biol. 2005; 4 10.2202/1544-6115.1128.

82. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT, Martuza RL, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014; 344:1396–1401. [PubMed: 24925914]
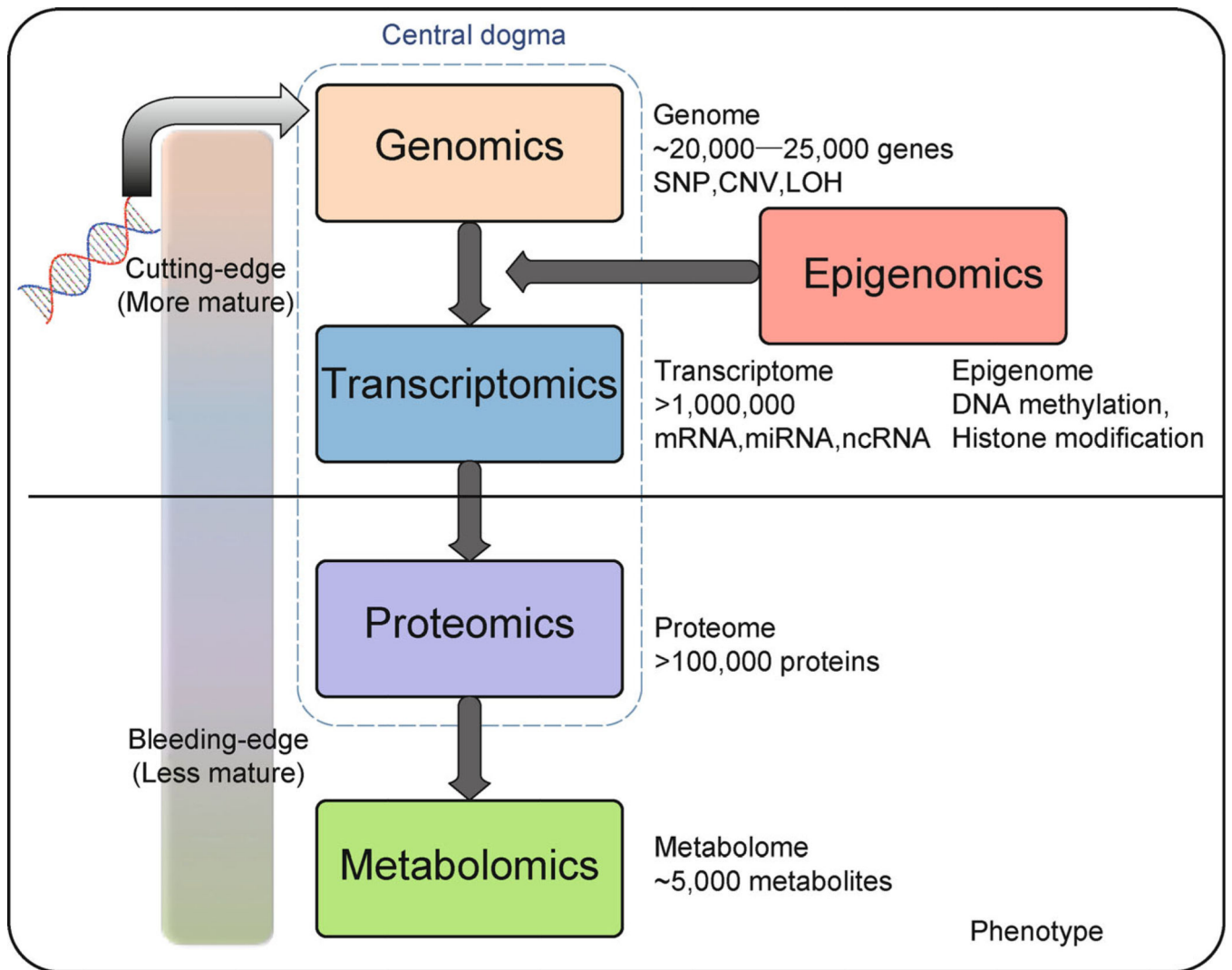
Figure 1. A systems biology approach through "omic" tools

The directional relationship of "omic" tools enables researchers to study the phenotypic progression from genome to metabolome. Multi-level data exists at the (i) genomic-level: single-nucleotide polymorphism (SNP), copy number variation (CNV), and loss of heterozygosity (LOH); (ii) transcriptomic-level: messenger RNA (mRNA), micro-RNA (miRNA), and non-coding RNA (ncRNA); (iii) proteomic-level: proteins; (iv) metabolomic-level: metabolites, and (v) epigenomic-level: DNA methylation and histone modification.
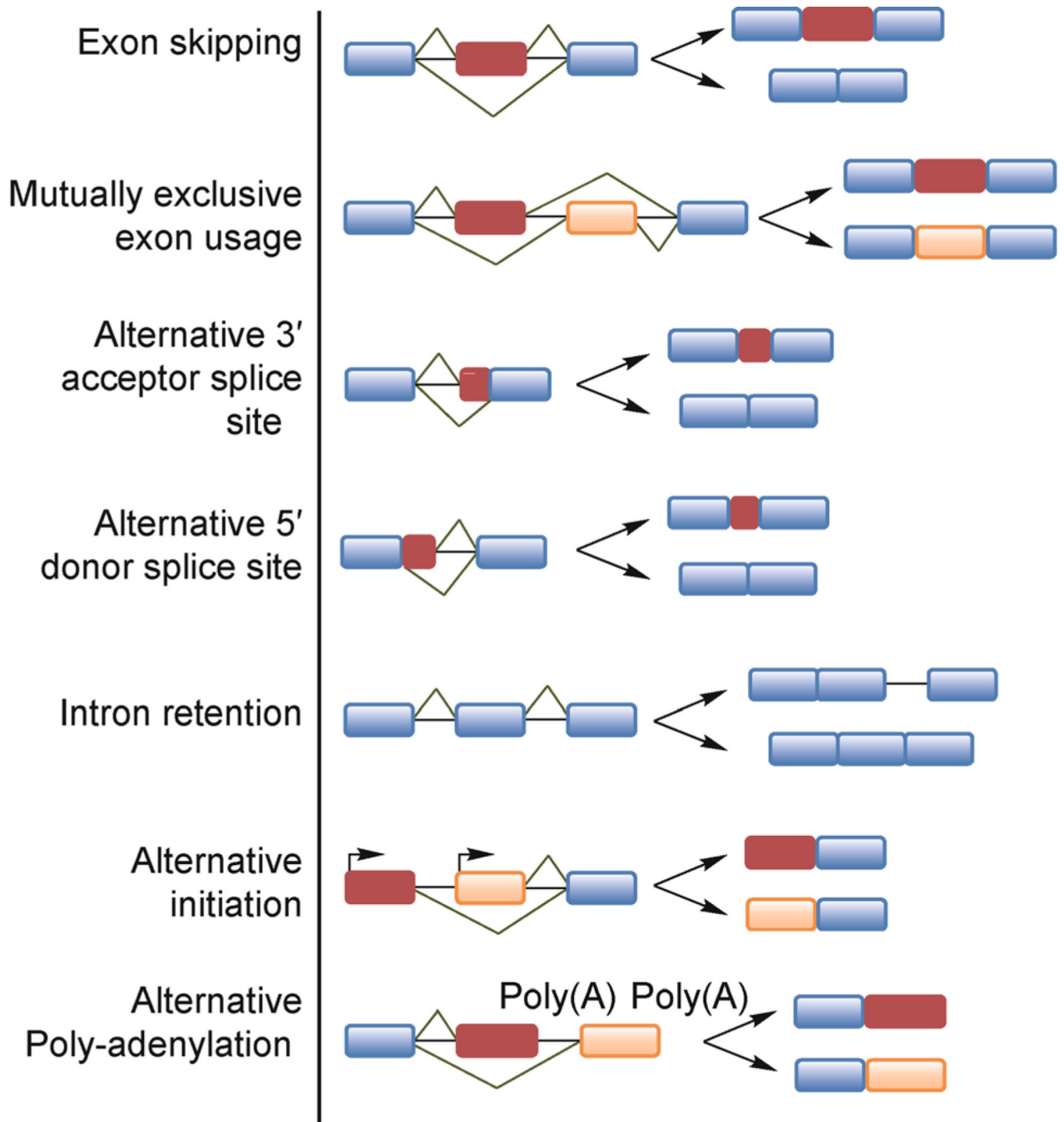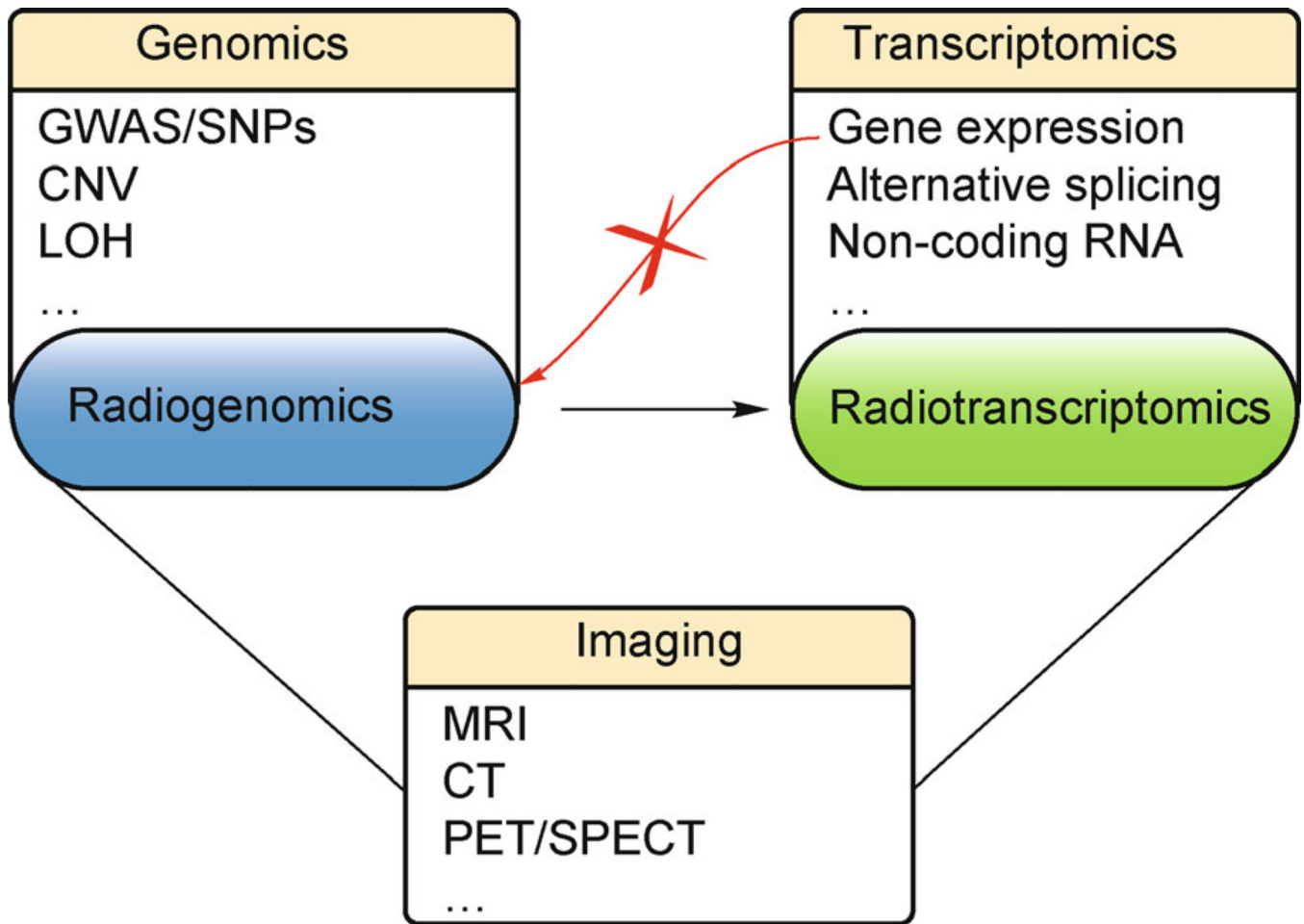
Figure 2. Alternative splicing modes of isoform variation

Figure 3. From Radiogenomics to Radiotranscriptomics

We propose a transition from the integration of imaging with genomics to that with transcriptomics, to unravel the impact of epigenetic and environmental factors on the static DNA code in the etiology of complex diseases. Note that studies involving gene expression and medical imaging have already been explored under the umbrella of radiogenomics. However, radiotranscriptomics is a more specific term, given that gene expression analysis is done at the transcriptome level.
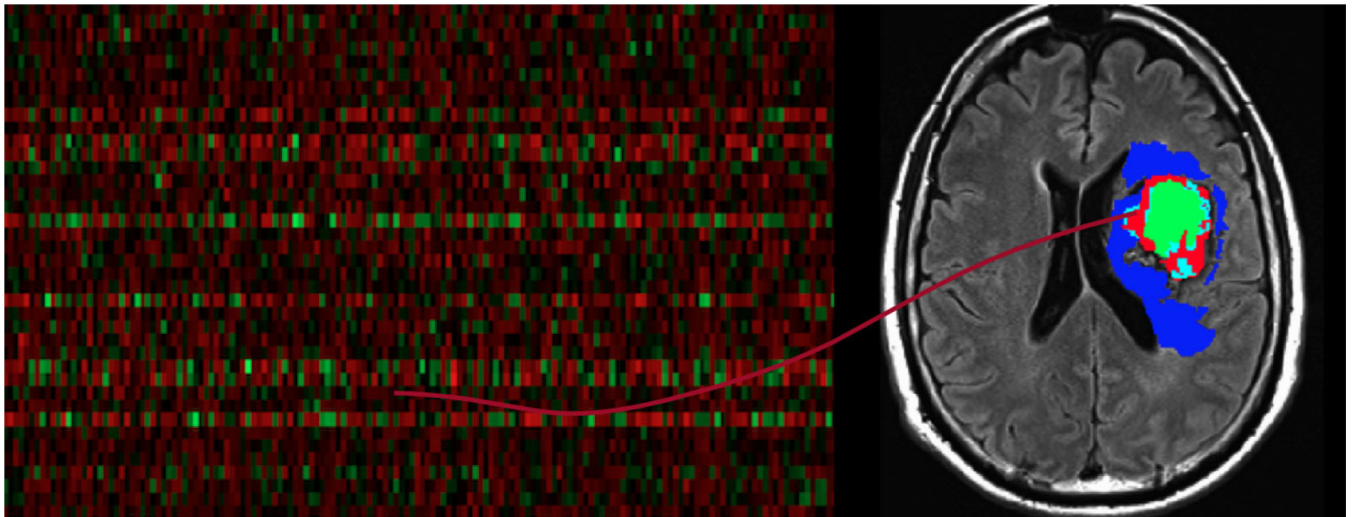
Figure 4. "Radiotranscriptomic" maps of biological disorders
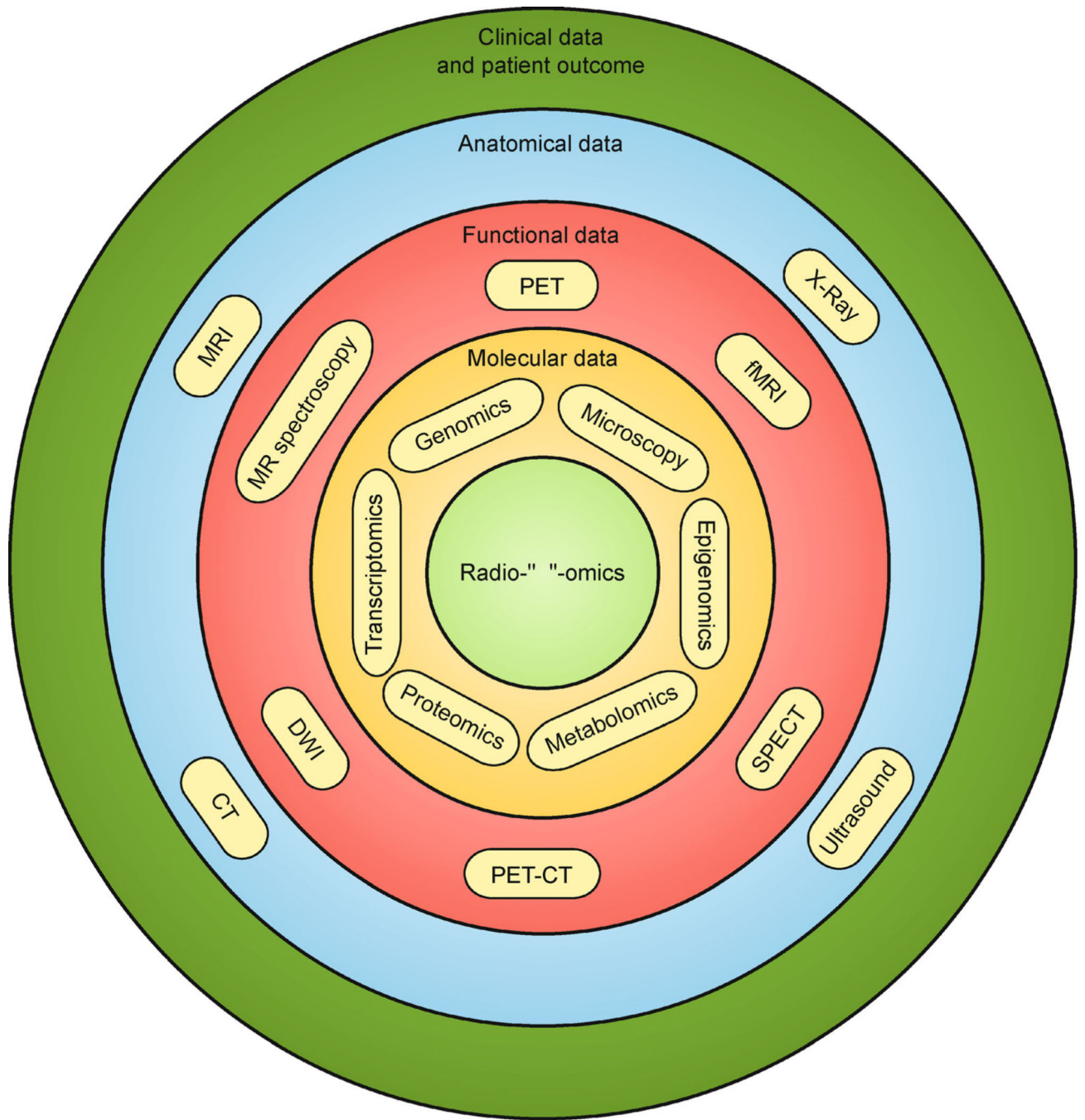Heatmap presents varying expression levels in correlation with a GBM-associated MR imaging annotation.

Figure 5. Integration of molecular, functional, and anatomical data in Radio-" "-omics