

Integration of Digital Pathology and Multi-Omics – The Good, The Bad, and the Ugly

Linh Tran

Incoming 2nd Year Biomedical Sciences student

University of Warwick

Word count: 1982

Introduction:

The seemingly separate worlds of digital pathology and multi-omics have both demonstrated promising results regarding their use, predominantly in research for cancer prognosis, as well as proposing improved methods for diagnosis compared to the current standard (Hu et al., 2023). However, the two worlds can collide via integration, unlocking the potentials for ground-breaking developments in a multitude of fields, including precision medicine (Ahmed, 2022)(Barsoum et al., 2019). The integration of the two suggests a revolutionary change in the way researchers from multiple disciplines can see, analyse and interpret patient data. These techniques are also able to accurately predict factors like survival; this invaluable information could aid in improving the quality of life of patients, and assist in the development of novel and effective therapies (J. Cheng et al., 2017). It is suggested that widespread implementation of these integrated techniques can also improve the work-flow for pathologists and other professionals alike, additionally widening the scope for collaboration between clinical and research fields (Jahn et al., 2020a).

Whilst both methods have garnered much support from organisations to individuals, these techniques are not without their pitfalls and serious ethical concerns, from IT

infrastructure concerns to implementation of artificial intelligence (McKay et al., 2022). This essay aims to examine the benefits of integration, scrutinise and rank the concerns and challenges by severity, and to outline proposed solutions and considerations that should be taken to these challenges in published literature.

Defining Digital Pathology and Multi-Omics:

Digital pathology and multi-omics each have their own advantages and challenges. To understand the implications of their integration requires appreciation of their respective functions and challenges.

The Royal College of Pathologists defines digital pathology as, ‘the acquisition, management, sharing and interpretation of pathology information – including slides and data – in a digital environment’ (The Royal College of Pathologists, 2023). It is based on the use of Whole Slide Imaging (WIS) which digitises the entirety of a slide, allowing it to be uploaded, stored, and shared amongst pathologists and researchers. It employs the use of specialised scanners and software to digitise the slide in high resolution (Farahani et al., 2015), and eliminates issues of physical slide storage, access to slides and slide preservation (Kumar et al., 2020). It is greatly used in education as it serves as a convenient and effective way to teach (Hassell et al., 2023). Figure 1. Shows an example of a digitised slide that may be used for a variety of purposes. However, it is not widely used as part of clinical practices yet, as no large scale clinical trials have been performed (Cross et al., 2018), and adequate IT infrastructure has yet to be established in the majority of hospitals due to significant costs

(Jahn et al., 2020a). These are only a few examples of the advantages and disadvantages of digital pathology; further examples will be elaborated on in relation to multi-omic strategies.

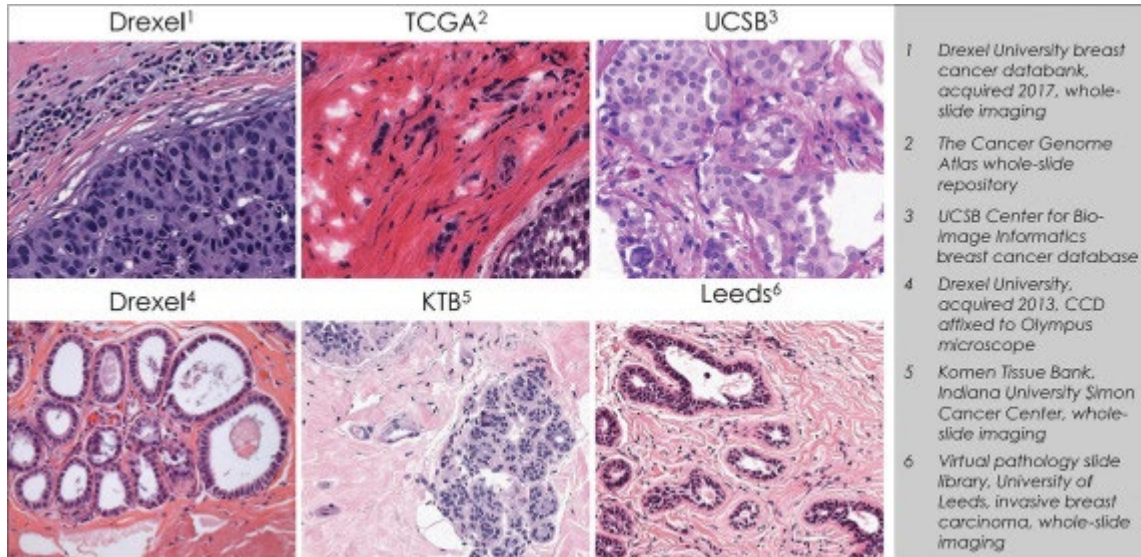


Figure 1: Multiple samples of haematoxylin and eosin stains that have been scanned using WSI (Aeffner et al., 2019).

Multi-omics refers to the integration of multiple datasets from the ‘-ome’ disciplines, namely genomics, epigenomics, transcriptomics, proteomics and more (Hasin et al., 2017a). These datasets can be analysed together and are used to characterise biological molecules involved in cellular processes with an extra level of depth. Multi-omics can be used to provide further insight into complex cellular mechanisms, and different approaches to disease research can be taken depending on the omic layer that’s chosen as a starting point. Figure 2. is a visualisation of the possible interactions between biological molecules and across omic layers.

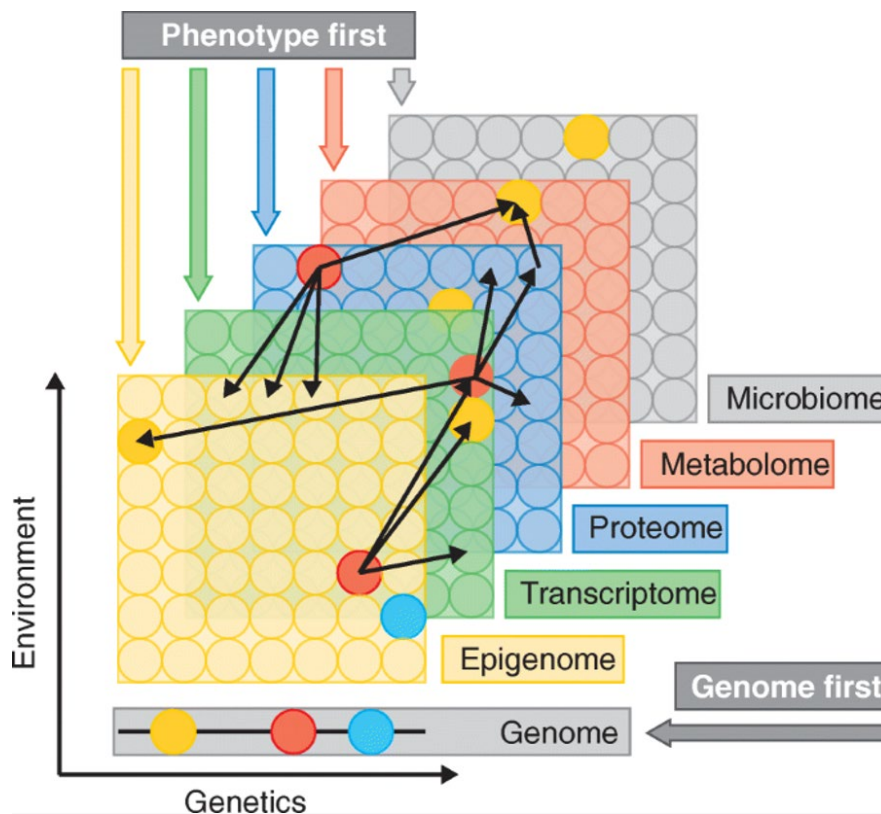


Figure 2: A visualisation of different omic layers and their interactions. Pools of omic data are represented by the circles, coloured arrows represent potential research starting points and black arrows represent potential interactions (Hasin et al., 2017).

Instead of focusing on a single ‘-ome’, multiple ‘omes’ can be investigated in tandem, and more can be revealed about their interactions. They are often modelled as networks and can be visualised, where information flow can provide further insight into correspondences between omic groups (Hasin et al., 2017). These omic datasets are a wealth of information, made up of hundreds to thousands of samples, and are found in multiple databases that are easily available for use, such as FlyBase and GenBank (Pinu et al., 2019). Data for each ‘ome’ is obtained through a wide variety of technologies, such as microarrays, which can analyse thousands of genes every experiment (Karahalil, 2016). The potential for multi-omic technologies is great, ranging from biomarker discoveries for cancers to

improving classification of diseases (Eddy et al., 2020), (Xiao et al., 2022). Multi-omics can provide a holistic approach to understanding the mechanisms of disease. Despite these promising developments in this emerging field, like digital pathology, it too has its disadvantages and causes for ethical concern (J. Cheng et al., 2017).

The Good – The Best of Both Worlds

By combining the forces of these technologies, the results can be ground-breaking. When used with machine learning, predictions can be made regarding prognosis of various cancers, such as breast and ovarian cancers, both cancers having the highest incidence and mortality rates amongst women (Hu et al., 2023; Xiao et al., 2022). Together, they can also predict biomarkers for cancer from routine pathology images, which can greatly increase the number of early diagnoses in a greatly efficient manner, improving survival chances for patients. This also has potentials for aiding in the development of better cancer treatments (Arslan et al., 2022). The integration of digital pathology and multi-omics has also led to the identification of lung cancer hallmarks in histopathology images, providing a deeper understanding into the mechanisms of the cancer (Guramare et al., 2022). Figure 3. shows visualised data and histopathological samples that were used to predict the progression-free outcomes of patients with colorectal cancer.

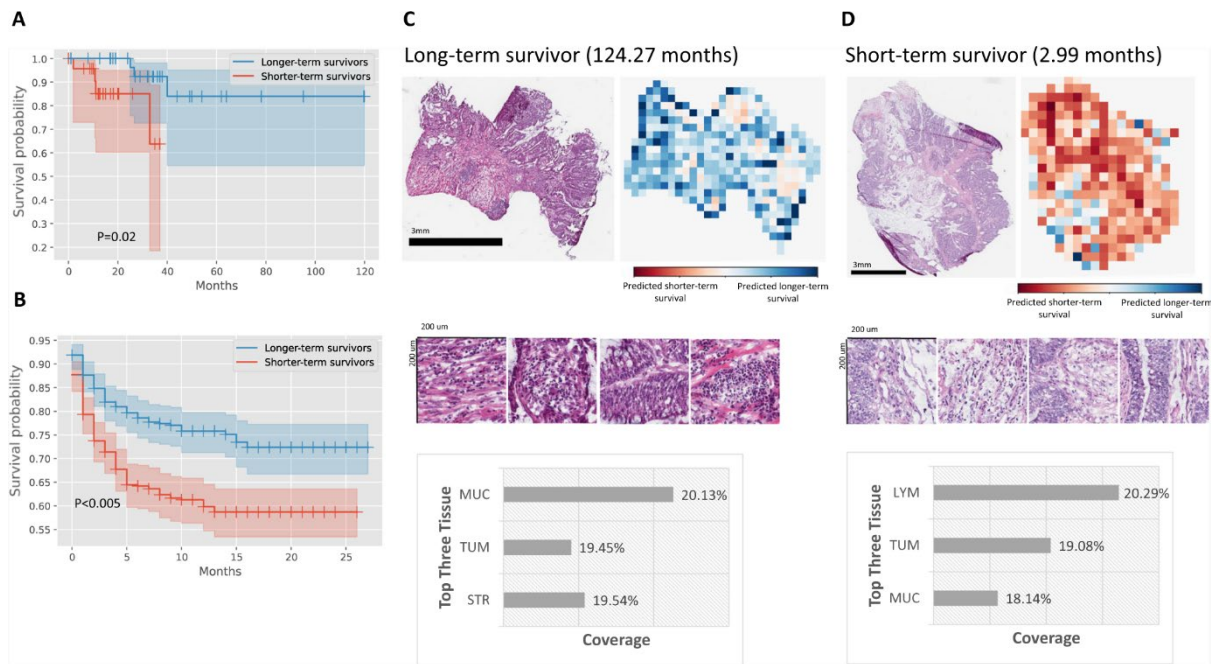


Figure 3: The machine learning Multi-omics Multi-cohort Assessment (MOMA) platform used histopathological and multi-omic data to successfully predict the progression-free survival outcomes of stage I and II colorectal cancer patients (Tsai et al., 2023).

Even further, this integration has also led to the development of machine learning models that can accurately predict survival outcomes, transcription subtypes and genetic deviations that play a role in the development of lung adenocarcinoma (Chen et al., 2021).

Data derived from multi-omics can be used to profile a tumour's microenvironment which helps provide a deeper level of understanding of the interactions that occur (Van Oekelen & Laganà, 2022). When considered in tandem with the physical manifestations of the tumour as seen in histopathological images, the magnified level of understanding and breadth of data available makes the integration of digital pathology and multi-omics an incredibly promising field, with seemingly limitless possibilities. The examples previously listed are only a few of the developments that have been made in this field. Whilst the benefits and potentials are evident, the challenges need to be addressed as there are multiple causes of concern with these technologies before further developments can be made.

The Bad – Issues with Integration and Infrastructure

One of the prevalent issues for both digital pathology and multi-omic technologies is the lack of infrastructure required to sustain these techniques on a wider scale. Digital pathology requires specialised hardware and software that come at a hefty price, as well as specialised training required in order to use scanners and software (Griffin & Treanor, 2017). The same is for multi-omics regarding use of analysis software. As well as this, data storage is another issue, as both require very large amounts of storage to save their vast amounts of data (Subramanian et al., 2020). Both methods have their own respective challenges, such as multi-omics requiring a workforce of multiple disciplines to conduct analysis, and digital pathology requiring adequate quality control and the need for standardisation (Jahn et al., 2020).

When considering integration, this gives rise to new challenges; one of the most prominent being the integration of omic and non-omic data. The nature of datasets is that it is often heterogeneous, which makes pinpointing the exact mechanism or cause of a disease incredibly difficult. The nature of the disease itself may also be heterogenous making the task of identifying specific molecular mechanisms responsible for the disease even more complex (de Maturana et al., 2019). Heterogeneity also contributes to the problem of bias in both digital pathology and multi-omics, resulting in correlations being found that are not a direct result of a causative agent (Tarazona et al., n.d.). Bias is also amplified through extrinsic factors, such as the pressure of discovering novel biomarkers. It has been suggested that bias

is a significant cause of preventing reproducibility of results using multi-omics. A myriad of other issues are also a cause for concern, from missing data from omic datasets, statistical methods resulting in false positives, methodology faults and limited omic literacy amongst the scientific community (Lay et al., 2006).

These issues are only amplified when integration is introduced to the equation.

Solutions to these problems have been proposed, such as developing more robust criteria before experimentation to avoid bias (Lay et al., 2006), but without adequate standardisation of the fundamental data itself, and more established methods on handling said data, these issues will remain prevalent.

The Ugly – Ethics and the Potential Misuse of Life-Intrusive Information

Technical issues considered, the more pressing causes for concern with the integration of digital pathology and multi-omics are privacy risk, informed choice, equity, and the implementation of AI in these practices.

Regulations on data protection regarding multi-omics (Williams & Anderson, 2018a), digital pathology and the use of AI are not yet well established (J. Y. Cheng et al., 2021), and this leaves the potential of third-party involvement that do not have the participants' best interests considered. The positive developments of this integration also have their negatives, where instead of using genomic and epigenomic for good, it can be used for malicious motivations. Much emphasis is placed on regulating data privacy surrounding genomics,

however the same cannot be said for the other omes. Epigenomic data has the potential to be used as a way to deduct a participant's lifestyle choices, use analysis techniques that could reveal information such as drug use, their medical conditions and more (Dupras & Bunnik, 2021). This data can also be used, particularly by health insurance companies, to discriminate against those with medical conditions based on their genomic data. Genomic can also worryingly be used to discriminate against those from other ethnic groups and sexes. Information found in genomic and epigenomic data increases the risk of reidentification, and this is amplified when integrated with different omic datasets and non-omic data (Safarlou et al., 2021).

Literacy amongst research participants regarding multi-omic technologies and digital pathology is likely to be limited, and this can result in participants unknowingly giving consent to partake in research, as well as having their data used, without fully understanding the implications of doing so (Williams & Anderson, 2018). There are also concerns surrounding the health equity implications of these technologies, with non-representative data that currently exists possibly having influence on machine learning training algorithms for these methods, introducing bias and inaccurate data, which can cause significant harm to at-risk populations (Williams & Anderson, 2018). But perhaps the most popular cause for concern is the use of AI in both digital pathology and multi-omics, which in itself brings up a multitude of ethical questions (McKay et al., 2022), from black box models (Wen et al., 2023), data transparency and accountability (Srivastava, 2023).

Conclusion

The potentials of integration should be recognised as it could be revolutionary in multiple research fields, particularly in cancer prognosis. It can provide us a newfound depth of understanding of the biomolecular processes responsible for disease. Despite this, further advancements cannot be made without addressing the serious challenges that are posed.

Through collaboration, education and establishing appropriate regulation on a wide scale, the use of digital pathology and multi-omics can be safely used for the betterment of potentially millions of patients worldwide.

Bibliography:

- Aeffner, F., Zarella, M. D., Buchbinder, N., Bui, M. M., Goodman, M. R., Hartman, D. J., Lujan, G. M., Molani, M. A., Parwani, A. V., Lillard, K., Turner, O. C., Vemuri, V. N. P., Yuil-Valdes, A. G., & Bowman, D. (2019). Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association. *Journal of Pathology Informatics*, *10*(1), 9. https://doi.org/10.4103/JPI.JPI_82_18
- Ahmed, Z. (2022). Precision medicine with multi-omics strategies, deep phenotyping, and predictive analysis. *Progress in Molecular Biology and Translational Science*, *190*(1), 101–125. <https://doi.org/10.1016/BS.PMBTS.2022.02.002>
- Arslan, S., Mehrotra, D., Schmidt, J., Geraldles, A., Singhal, S., Hense, J., Li, X., Bass, C., Kather, J. N., Pandya, P., & Raharja-Liu, P. (2022). Deep learning can predict multi-omic biomarkers from routine pathology images: A systematic large-scale study. *BioRxiv*, 2022.01.21.477189. <https://doi.org/10.1101/2022.01.21.477189>
- Barsoum, I., Tawedrous, E., Faragalla, H., & Yousef, G. M. (2019). Histo-genomics: digital pathology at the forefront of precision medicine. *Diagnosis (Berlin, Germany)*, *6*(3), 203–212. <https://doi.org/10.1515/DX-2018-0064>
- Chen, L., Zeng, H., Xiang, Y., Huang, Y., Luo, Y., & Ma, X. (2021). Histopathological Images and Multi-Omics Integration Predict Molecular Characteristics and Survival in Lung Adenocarcinoma. *Frontiers in Cell and Developmental Biology*, *9*, 720110. <https://doi.org/10.3389/FCELL.2021.720110/FULL>
- Cheng, J. Y., Abel, J. T., Balis, U. G. J., McClintock, D. S., & Pantanowitz, L. (2021). Challenges in the Development, Deployment, and Regulation of Artificial Intelligence in

- Anatomic Pathology. *The American Journal of Pathology*, 191(10), 1684–1692.
<https://doi.org/10.1016/J.AJPATH.2020.10.018>
- Cheng, J., Zhang, J., Han, Y., Wang, X., Ye, X., Meng, Y., Parwani, A., Han, Z., Feng, Q., & Huang, K. (2017). Integrative Analysis of Histopathological Images and Genomic Data Predicts Clear Cell Renal Cell Carcinoma Prognosis. *Cancer Research*, 77(21), e91–e100. <https://doi.org/10.1158/0008-5472.CAN-17-0313>
- Cross, S., Furness, P., Igali, L., Snead, D., & Treanor, D. (2018). *Best practice recommendations for implementing digital pathology Unique document number G162 Document name Best practice recommendations for implementing digital pathology.*
- de Maturana, E. L., Alonso, L., Alarcón, P., Martín-Antoniano, I. A., Pineda, S., Piorno, L., Calle, M. L., & Malats, N. (2019). Challenges in the Integration of Omics and Non-Omics Data. *Genes*, 10(3). <https://doi.org/10.3390/GENES10030238>
- Dupras, C., & Bunnik, E. M. (2021). Toward a Framework for Assessing Privacy Risks in Multi-Omic Research and Databases. <https://doi.org/10.1080/15265161.2020.1863516>, 21(12), 46–64.
<https://doi.org/10.1080/15265161.2020.1863516>
- Eddy, S., Mariani, L. H., & Kretzler, M. (2020). Integrated multi-omics approaches to improve classification of chronic kidney disease. *Nature Reviews Nephrology* 2020 16:11, 16(11), 657–668. <https://doi.org/10.1038/s41581-020-0286-5>
- Farahani, N., Parwani, A. V., & Pantanowitz, L. (2015). Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International*, 23. <https://doi.org/10.2147/PLMI.S59826>
- Griffin, J., & Treanor, D. (2017). Digital pathology in clinical use: where are we now and what is holding us back? *Histopathology*, 70(1), 134–145.
<https://doi.org/10.1111/HIS.12993>
- Guramare, M., Javed, S. A., Agrawal, N., Griffin, M., Abel, J., Montalto, M., Beck, A. H., Mountain, V., Wapinski, I., Resnick, M., Biddle-Snead, C., & Conway, J. R. (2022). P2.15-01 Digital Pathology Uncovers Multi-Omic Hallmarks of Lung Cancer in Histopathology Images. *Journal of Thoracic Oncology*, 17(9), S158.
<https://doi.org/10.1016/J.JTHO.2022.07.261>
- Hasin, Y., Seldin, M., & Lusic, A. (2017a). Multi-omics approaches to disease. *Genome Biology* 2017 18:1, 18(1), 1–15. <https://doi.org/10.1186/S13059-017-1215-1>
- Hasin, Y., Seldin, M., & Lusic, A. (2017b). Multi-omics approaches to disease. *Genome Biology* 2017 18:1, 18(1), 1–15. <https://doi.org/10.1186/S13059-017-1215-1>
- Hassell, L. A., Absar, S. F., Chauhan, C., Dintzis, S., Farver, C. F., Fathima, S., Glassy, E. F., Goldstein, J. A., Gullapalli, R., Ho, J., Koch, L. K., Madory, J. E., Mirza, K. M., Nguyen, P. N., Pantanowitz, L., Parwani, A., Rojansky, R., Seifert, R. P., Singh, R., ... Bui, M. (2023). Pathology Education Powered by Virtual and Digital Transformation: Now and the Future. *Archives of Pathology & Laboratory Medicine*, 147(4), 474–491.
<https://doi.org/10.5858/ARPA.2021-0473-RA>

- Hu, J., Lv, H., Zhao, S., Lin, C.-J., Su, G.-H., & Shao, Z.-M. (2023). Prediction of clinicopathological features, multi-omics events and prognosis based on digital pathology and deep learning in HR + /HER2 – breast cancer. *Journal of Thoracic Disease*, 15(5), 2528–2543. <https://doi.org/10.21037/JTD-23-445>
- Jahn, S. W., Plass, M., & Moinfar, F. (2020a). Digital Pathology: Advantages, Limitations and Emerging Perspectives. *Journal of Clinical Medicine*, 9(11), 1–17. <https://doi.org/10.3390/JCM9113697>
- Jahn, S. W., Plass, M., & Moinfar, F. (2020b). Digital Pathology: Advantages, Limitations and Emerging Perspectives. *Journal of Clinical Medicine 2020, Vol. 9, Page 3697*, 9(11), 3697. <https://doi.org/10.3390/JCM9113697>
- Karahalil, B. (2016). Overview of Systems Biology and Omics Technologies. *Current Medicinal Chemistry*, 23(37), 4221–4230. <https://doi.org/10.2174/0929867323666160926150617>
- Kumar, N., Gupta, R., & Gupta, S. (2020). Whole Slide Imaging (WSI) in Pathology: Current Perspectives and Future Directions. *Journal of Digital Imaging*, 33(4), 1034. <https://doi.org/10.1007/S10278-020-00351-Z>
- Lay, J. O., Liyanage, R., Borgmann, S., & Wilkins, C. L. (2006). Problems with the “omics.” *TrAC Trends in Analytical Chemistry*, 25(11), 1046–1056. <https://doi.org/10.1016/J.TRAC.2006.10.007>
- McKay, F., Williams, B. J., Prestwich, G., Bansal, D., Hallowell, N., & Treanor, D. (2022). The ethical challenges of artificial intelligence-driven digital pathology. *The Journal of Pathology: Clinical Research*, 8(3), 209–216. <https://doi.org/10.1002/CJP2.263>
- Pinu, F. R., Beale, D. J., Paten, A. M., Kouremenos, K., Swarup, S., Schirra, H. J., & Wishart, D. (2019). Systems biology and multi-omics integration: Viewpoints from the metabolomics research community. *Metabolites*, 9(4). <https://doi.org/10.3390/METABO9040076>
- Safarlou, C. W., Bredenoord, A. L., Vermeulen, R., & Jongsma, K. R. (2021). Scrutinizing Privacy in Multi-Omics Research: How to Provide Ethical Grounding for the Identification of Privacy-Relevant Data Properties. <https://doi.org/10.1080/15265161.2021.1991041>, 21(12), 73–75. <https://doi.org/10.1080/15265161.2021.1991041>
- Srivastava, R. (2023). Applications of artificial intelligence multiomics in precision oncology. *Journal of Cancer Research and Clinical Oncology*, 149(1), 503–510. <https://doi.org/10.1007/S00432-022-04161-4/FIGURES/3>
- Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Multi-omics Data Integration, Interpretation, and Its Application. *Bioinformatics and Biology Insights*, 14. https://doi.org/10.1177/1177932219899051/ASSET/IMAGES/LARGE/10.1177_1177932219899051-FIG1.JPEG
- Tarazona, S., Arzalluz-Luque, A., & Conesa, A. (n.d.). Undisclosed, unmet and neglected challenges in multi-omics studies. *Nature Computational Science*. <https://doi.org/10.1038/s43588-021-00086-z>

- The Royal College of Pathologists. (2023). *Digital Pathology*. The Royal College of Pathologists. <https://www.rcpath.org/profession/digital-pathology.html>. Date accessed [30.08.2023].
- Tsai, P. C., Lee, T. H., Kuo, K. C., Su, F. Y., Lee, T. L. M., Marostica, E., Ugai, T., Zhao, M., Lau, M. C., Väyrynen, J. P., Giannakis, M., Takashima, Y., Kahaki, S. M., Wu, K., Song, M., Meyerhardt, J. A., Chan, A. T., Chiang, J. H., Nowak, J., ... Yu, K. H. (2023). Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. *Nature Communications*, *14*(1). <https://doi.org/10.1038/S41467-023-37179-4>
- Van Oekelen, O., & Laganà, A. (2022). Multi-Omics Profiling of the Tumor Microenvironment. *Advances in Experimental Medicine and Biology*, *1361*, 283–326. https://doi.org/10.1007/978-3-030-91836-1_16
- Wen, Y., Zheng, L., Leng, D., Dai, C., Lu, J., Zhang, Z., He, S., & Bo, X. (2023). Deep Learning-Based Multiomics Data Integration Methods for Biomedical Application. *Advanced Intelligent Systems*, *5*(5), 2200247. <https://doi.org/10.1002/AISY.202200247>
- Williams, J. K., & Anderson, C. M. (2018a). Omics research ethics considerations. *Nursing Outlook*, *66*(4), 386–393. <https://doi.org/10.1016/J.OUTLOOK.2018.05.003>
- Williams, J. K., & Anderson, C. M. (2018b). Omics research ethics considerations. *Nursing Outlook*, *66*(4), 386–393. <https://doi.org/10.1016/J.OUTLOOK.2018.05.003>
- Xiao, Y., Bi, M., Guo, H., & Li, M. (2022). Multi-omics approaches for biomarker discovery in early ovarian cancer diagnosis. *EBioMedicine*, *79*, 104001. <https://doi.org/10.1016/J.EBIOM.2022.104001>