CASE STUDY REPORT

Challenges and Strategies for the Application of Omics and Spatial Technologies

A Concise Report Featuring Insights From The Prominent Thought Leaders Of Omics 2023 and Spatial 2024

Introduction

Omics and spatial biology technologies represent revolutionary advancements in the field of life sciences, fundamentally transforming our understanding of biological systems. Omics technologies provide unprecedented insights into the genetic, transcriptomic, proteomic, and metabolic landscapes, enabling a deeper understanding of biological processes, disease mechanisms, and potential therapeutic targets.

Also transforming the landscape is the ability to analyse biological specimens within their spatial context. Spatial transcriptomics, spatial proteomics, and spatial metabolomics allow researchers to see and inspect the spatial organization of biomolecules within complex biological systems, shedding light on the spatial heterogeneity of tissues, cellular interactions, and disease microenvironments. The implications are far reaching, driving successes in the areas of personalized medicine, clinical development, drug discovery, and beyond.

This year, DeciBio predicted the Spatial Biology Market to grow at a rate of 21% per annum, reaching $2.27B by 2029. In this time, we expect to witness substantial innovation. Advances in high-throughput sequencing technologies, mass spectrometry, imaging modalities, and computational algorithms will continue to drive the development of more sensitive, accurate, and cost-effective omics and spatial biology tools. This will enable researchers to analyse biological systems with unprecedented resolution and throughput, facilitating discoveries that were previously unimaginable.

Moreover, the integration of omics and spatial biology data with artificial intelligence and machine learning algorithms will further enhance our ability to interpret complex biological information, identify biomarkers, and discover novel therapeutic targets. These synergistic approaches will accelerate the pace of biomedical research, leading to the development of more effective diagnostics, therapies, and precision medicine strategies.

Furthermore, the adoption of omics and spatial biology technologies is expected to expand beyond academic research settings to clinical diagnostics, pharmaceutical companies, biotechnology startups, and other industries. This broader adoption will fuel innovation, stimulate economic growth, and ultimately improve healthcare outcomes for patients worldwide.

Tom Cohen,
Senior Digital Content Editor, Oxford Global

Contents

Key Speakers Include

Maria Antonietta Cerone,
Principal Scientist,
Cancer Research UK

Sophia Tsoka,
Reader in Bioinformatics,
King's College London

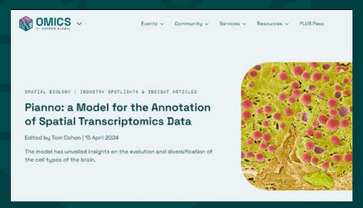
Martin Isabelle,
Associate Director,
Adaptimmune

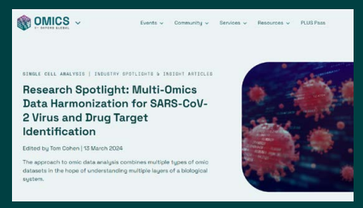
Catia Costa,
Senior Research Fellow,
University of Surrey

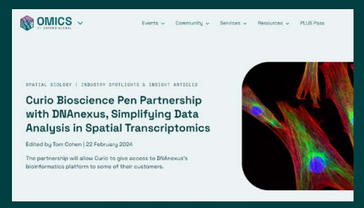
Explore Our Comprehensive Content Portal

Immerse yourself in cutting-edge scientific content - from online Monthly Science Exchanges, best practice Online Symposiums to eBooks and landscape reports providing a unique perspective on the latest R&D trends and challenges.

- Pianno: a Model for the Annotation of Spatial Transcriptomics Data
- Research Spotlight: Multi-Omics Data Harmonization for SARS-CoV-2 Virus and Drug Target Identification
- Curio Bioscience Pen Partnership with DNAnexus, Simplifying Data Analysis in Spatial Transcriptomics







Using Spatial And Multi-Omics Data

Grand View Research predicts the global multiomics market, which valued at USD 2.35 billion in 2023, to grow at a CAGR of 15.29% from 2024 to 2030, driven by the rising demand for single-cell multiomics and technological advancements. The market's growth is further supported by the increasing use of genomics, proteomics, metabolomics, and transcriptomics in healthcare, and it was positively impacted by the COVID-19 pandemic, which highlighted the role of multiomics in diagnosing and understanding the virus.

Source: Grand View Research

Utilizing Spatial Data In Biology

Why multimodal imaging?

Catia Costa (Senior Research Fellow, University of Surrey) and her team use multimodal imaging to get as much information as possible out of a sample. This is in service of understanding all the elements of a sample (metabolites, lipids, proteins, etc...) on every omics level and how they all relate to each other.

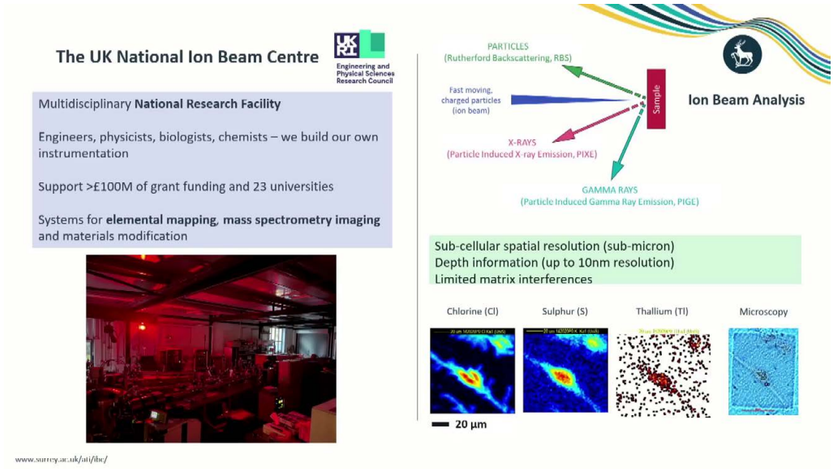
The Surrey Ion Beam Centre are developing analytical strategies to combine elemental mapping with mass spectrometry imaging. From this they hope to understand the impact of molecules like metal nanoparticles, elemental accumulation, and metal containing drugs on the local chemistry (e.g. metabolites, proteins, lipids).

The Surrey Ion Beam Centre

The Surrey Ion Beam Centre are part of the UK's National Ion Beam Centre and funded by the Engineering & Physical Sciences Research Council (EPRC). The centre builds its own instruments in house and work with both academia and industry. Their ion beam analysis lab is used for elemental mapping and mass spectrometry imaging.

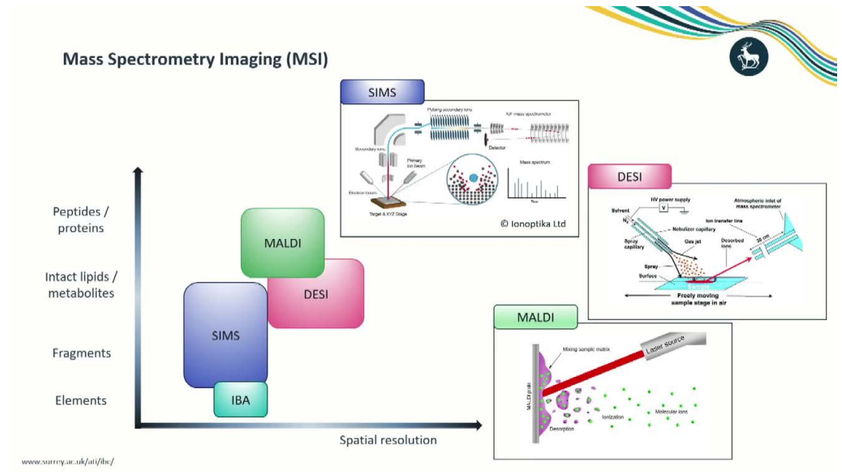
Ion beam analysis

A suite of techniques is available for ion beam analysis, the one that costa works with uses a proton beam. When the beam hits a sample, the team can detect a variety of reactions: the particle can bounce back, or the sample can produce x-rays and gamma rays. The team detect all this information simultaneously, so they have a well-rounded package of knowledge about the elemental composition of the sample.



Performing ion beam analysis alongside mass spectrometry imaging

There are a variety of techniques available for mass spectrometry imaging, such as secondary ion mass spectrometry (SIMS), desorption electrospray ionisation (DESI), and matrix assisted laser desorption ionisation (MALDI). The team wanted to use these techniques (elemental and molecular imaging) in parallel with IBA, but on the same tissue sample.



Challenges for integrating MSI and IBA for analysis of a single tissue sample

Costa outlined four of the main challenges that her team faced when trying to integrate these imaging methods together.

1. Different data handling strategies

Each technique offers different spatial resolutions and uses different data standards which can't be easily integrated.

2. Expertise and instrumentation are not normally co-located

Until very recently, the ion beam centre didn't have any mass spectrometry experts or instruments in their lab.

3. Different sample handling requirements

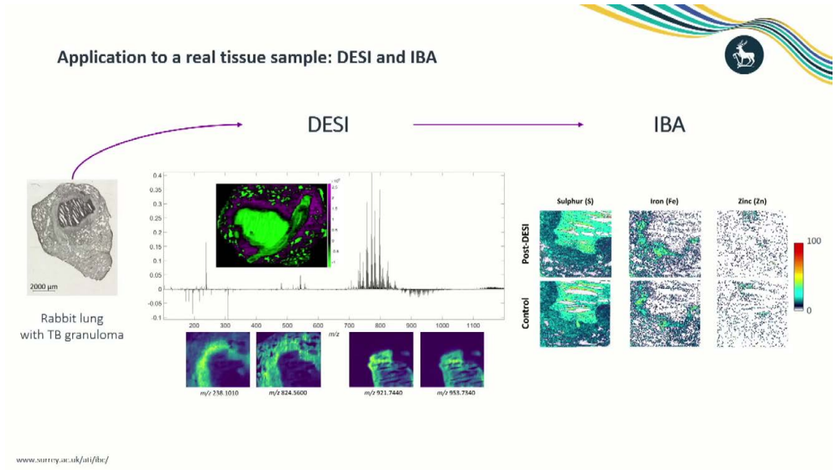
Different techniques require different handling requirements for the samples that they study. For example, MSI uses glass slides, but glass is cannot be used for IBA. To fix this they found PET films from Leica that are commercially available and compatible with IBA, DESI, and SIMS. However, the MALDI laser burned the PET slide, so they are currently working on this.

4. The techniques are destructive

Ion beam analysis (IBA) uses a proton beam which causes high levels of damage to the sample. The beam team have explored methods to reduce and mitigate proton beam damage (J. Am. Soc. Mass Spectrom. 2022, 33, 12, 2263−2272), but no method significantly reduced the damage.

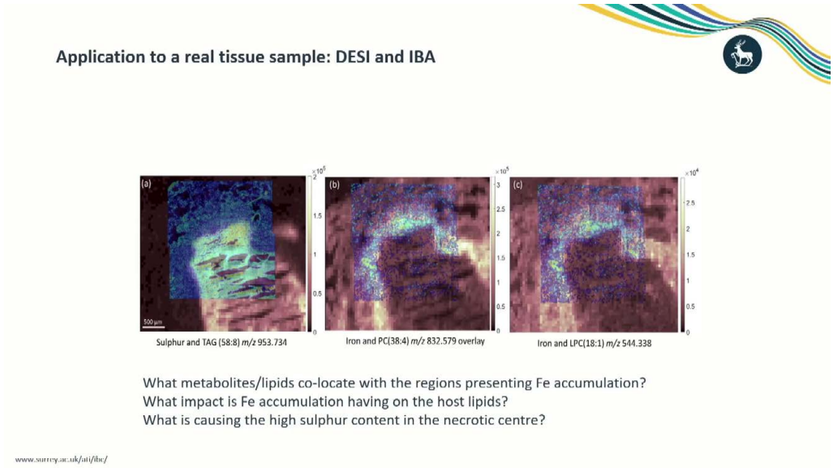
Furthermore, the fact that DESI uses a solvent led to the loss of ions like chlorine and potassium; to mitigate this, they changed to a fully organic solvent. They also did the mass spectrometry first, as it was less likely to cause damage before the other destructive techniques.

A protocol that worked for DESI and IBA



The protocol which saw DESI and IBA work together had no loss of elements. Furthermore, the combination of imaging allowed them to see a band of iron around the necrotic centre, where they also saw large sulphur deposits. The team wanted to know why this was and formulated the following questions:

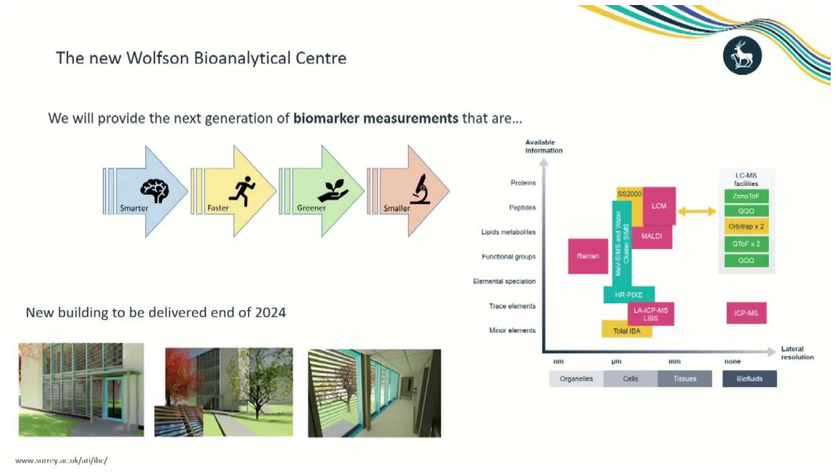
- What metabolites/lipids co-locate with the regions presenting Fe accumulation?
- What impact is Fe accumulation having on the host lipids?
- What is causing the high sulphur content in the necrotic centre?



Conclusion

The team have made a lot of progress in this area, and there have been a lot of challenges to overcome. Most of what has been achieved has been with DESI as it is the easiest to use, but the team are not going to give up on MALDI as the technique will allow them to better mass spectrometry of proteins. Furthermore, the team are continuing their research with SIMS and also non-MSI techniques like Raman.

The Surrey Ion Beam Centre has just been awarded a £2.9m grant from the Engineering & Physical Sciences Research Council (EPSRC). With those funds, they are planning to put SIMS instruments in the same chamber as IBA so that they can do both techniques straight after one another, meaning that they won't need to take the sample out. They also want to try MeV SIMS, for which there is data to suggest they will get a much higher yield of intact molecule. Furthermore, a new building, the Wolfson Bioanalytical Centre, is under construction and due to be completed by the end of 2024.



Analysis of Omics Data via Machine Learning

Sophia Tsoka (Reader in Bioinformatics, King's College London) began her talk by stating that when the bioinformatics team are presented with a set of data, the team needs to adapt their methodology in order to solve the particular problem at hand. Her team of data scientists aim to identify latent patterns (properties that are encoded in the data) and predict future events using analysis techniques. They are interested in offering methods based on the principles of machine learning (ML) not only in the interest of performance, but also interpretability: to flexibly represent data and challenges they deal with.

This of course this involves a vast amount of data, and advances in data analysis techniques and hardware used for computation have helped address this challenge. But there are also other problems which manifest: the scale of data is variable, from the atomic level upwards; the data comes in different formats, e.g., text, vectors, or numbers; and there are different inherent properties which need to be considered.

Bioinformatics and data science can address these challenges. This involves employing different methodologies that are based on mathematical optimisation, as well as newer approaches that are based on neural networks. Critically, these models need to be 'white boxes' – fully interpretable and explainable.

Computational Challenges in Biomedical Data Science

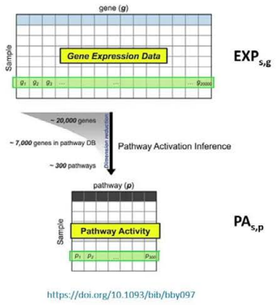
Tsoka introduced the so-called "large p, small n problem:" the typical data that her team analyse has tens to hundreds of samples and observations, but the features that get measured can number in the order of tens of thousands.

Feature selection is the standard method used to deal with this problem. Although feature selection methods are abundant in the literature, most of them neglect to consider the inter-relationships between the features. Tsoka stressed the fact that features are not independent, and they do not act in isolation of each other.

So Tsoka and her team are working on ways in which they can represent those relationships and molecular interactions (e.g. biochemical pathways) which allows for: Dimensionality reduction in a biologically meaningful manner, improvement in classification prediction performance, and better interpretability of the model.

Pathway Activity Inference

- Aggregating gene-level features into pathway-level data
  - Improves the "large-p small-n" problem
  - allows the number of features to be comparable to the number of samples
- Pathway space data are then used for sample classification
- Better biological interpretation, as cellular functions are reflected better by pathways rather than single genes
- Robust representation of data as variance across samples is reduced
- Inferring the weights of genes (and hence pathway activity) is formulated as a combinatorial optimisation problem



https://doi.org/10.1093/bib/bby097

Inferring weights of genes and then deriving pathway-based matrices is formulated as a combinatorial optimisation problem. The team then looked to modelling this problem using mathematical optimisation.

Mathematical Programming Optimisation

Mathematical Programming Optimisation

Mathematical programming is optimisation technique that
- models a problem with mathematical equations
- finds the best answer matching to the maximised or minimised the objective function value

feasible solution: set of values for the decision variables that satisfies all the constraints.

optimal solution: is one where there is no other feasible solution with a better objective function value

- advantage is adaptability and interpretability

$$\begin{aligned} min \quad & f(x) \\ s.t. \quad & g(x) \leq 0 \\ & h(x) = 0 \\ & x \in X \end{aligned}$$

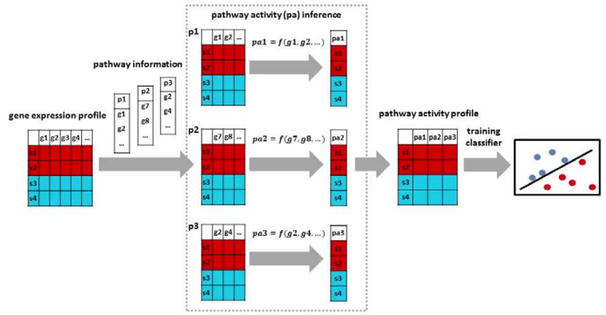
This technique starts with specified user constraints and a minimised or maximised objective function. They then model the problem using mathematical equations to find the best answer for the objective function value. Using this model, the feasible solution is a set of values for the decision variables that satisfies all the constraints. The optimal solution is one where there is no other feasible solution with a better objective function value.

Starting with a large omics dataset from a profile of some samples against proteins of interest, the first step in the process is to break up the matrix into pathway-specific matrices. The optimisation procedure follows:

1. Pathway activity definition: Summation of gene weights
2. Enforcement of non-overlapping limits: Phenotype ranges should not overlap
3. Pathway activity enclose limits: Each sample pathway activity is forced to lie within its phenotype range

The objective function in this instance is to minimise the number of samples allocated to the wrong class.

Disease Classification Using Pathway Activity



Once the team have a smaller matrix, they can then pass it through standard classifiers (e.g., random forest). Here, they can apply the classifier onto the pathway activity matrix and separate samples into the appropriate disease class. Tsoka commented that the mathematics for this process can be found in the appropriate references.

Optimisation-based Method - DIGS

- Pathway activity value **PAs**: *weighted linear summation of the expressions* of pathway genes:

$$PA_s = \sum_m G_{sm} \cdot (rp_m - rn_m) \quad \forall s$$

- Limitations on gene weights:

$$\begin{aligned} & rp_m \leq L_m \quad \forall m \\ & rn_m \leq 1 - L_m \quad \forall m \\ & \sum_m (rp_m + rn_m) = 1 \end{aligned}$$

- Limitations on number of active genes:

$$rp_m + rn_m \leq W_m \quad \forall m \qquad \sum_m W_m \leq NoG$$

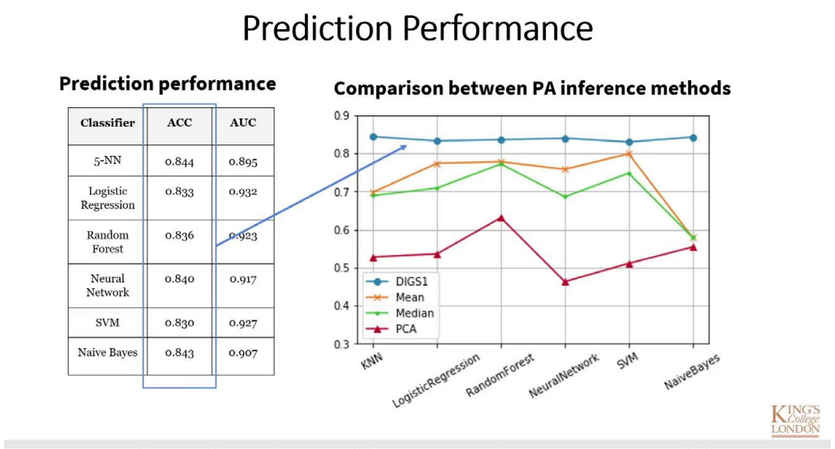
Yang L, Ainali C, Tsoka S, Papageorgiou LG, BMC Bioinformatics, 15(1):390, 2014

The pathway activity value is a weighted linear summation of the expressions of pathway genes. Here, the algorithm learns class ranges and is able to predict the class of a specific sample.

Implementation and Validation

For the machine learning model, Tsoka's team apply the standard-practice cross-

validation procedure. The dataset is randomly split: 70% of the data becoming a training set and 30% becoming a testing set – this process is repeated ten times. Then, training sets are split into pathway specific matrices – in this case, making 186 pathway specific matrices. The model is then trained on each pathway and produces gene weights. Using the gene weights, they calculate pathway activity values and generate the pathway activity matrix on which they train standard classifiers so that they are able to predict the classes of the samples.



The model is able to predict and match pathway activity that can separate samples into their appropriate phenotypes much more efficiently.

Summary

Tsoka concluded with a summary of her presentation. Current analyses are not simply supported by data, rather they are defined by data. Interpretable machine learning methods are needed to represent feature relationships and allow tracing of decision-making rules. Here, mathematical optimisation offers flexible and interpretable modelling. They also found that informed machine learning offers more realistic models and improved disease classification performance.

The team continue to explore this area, work is currently being done to apply this method to scRNA seq data and cell type annotation tasks.

Personalised Medicine

The Cancer Research UK Stratified Medicine Programme As A Model For Delivering Personalised Cancer Care

Maria Antonietta Cerone (Principal Scientist, Cancer Research UK) began by stating the primary mission of Cancer Research UK: "to create a world where everyone is free of the fear of cancer." CRUK's research spans all cancer types and the entire research pipeline. They work on all areas of research; however their main areas are cancer biology, treatment, and early diagnosis.

The Stratified Medicine Programme 2 (SMP2) is the first precision medicine study set up and run in the UK's NHS. It was set up in 2014 following its pilot study (SMP1). CRUK set out to provide genomic screening to patients with stage III/IV non-small cell lung cancer (NSCLC) who are not eligible for surgery or radiotherapy. Then, after they are screened, eligible patients can then be enrolled into the National Lung Matrix Trial (NLMT).

SMP2

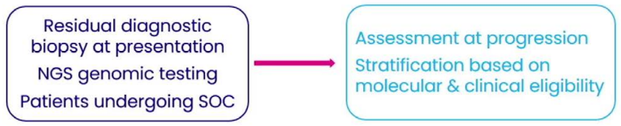
SMP2 focuses on lung cancer as it is the second most common cancer in men and women and the most common cause of cancer death in the UK (21% of all cancer deaths from 2017-2019). Over the last decade, overall incidence rates have been stable in the UK (13% in female and 12% in males) and overall survival has not improved in the last 50 years (<5%).

SMP2 recruited from a UK wide network of consenting patients. After that, a residual diagnostic biopsy was sent to one of three NHS Genomic labs. They used a bespoke 28 gene panel which was later replaced by a commercial panel from Illumina. While the patient was going through genomic testing, they would undergo the standard of care treatment. When the results were ready, they would be fed back to the NHS clinical hubs. Therefore, by the time a patient relapses, the molecular result will be available to the oncologist, and they can determine whether the patient is eligible for the NLMT.

NLMT

NLMT is a phase II, adaptive "signal seeking" umbrella study looking at advanced NSCLC patients. The trial takes the form of a Bayesian, modular design under one protocol to allow for arm drop-in/drop-out. Essentially, the study matched specific genetic aberrations to a new therapeutic agent. Those aberrations acted as predictive biomarkers for response to the drug.

The SMP2 NGS panel test genes required for NLMT



NGS panels used

Bespoke 28 gene panel
SMP2v01/SMP2v02
Designed by illumina/SMP2/ Pharma partners/NLMT

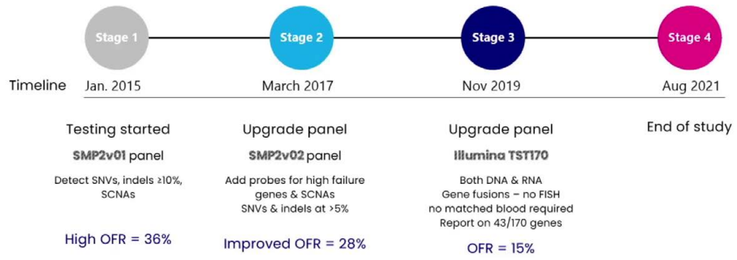
illumina **TST170**

Genes required for NLMT cohorts

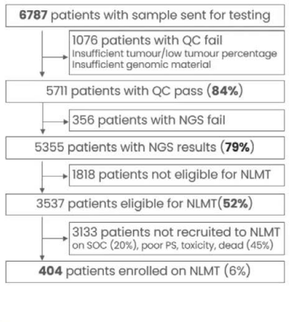
AKT1	ALK	BRAF	CCND1
CCND2	CCND3	CCNE1	CDK2
CDK4	CDKN2A	EGFR	FGFR2
FGFR3	Her2	HRAS	KRAS
MET	NF1	NRAS	NTRK1
PIK3CA	PTEN	RB1	RET
ROS1	STK11	TSC1	TSC2

Genomic report available at relapse to ensure enrolment in NLMT

SMP2 adapted & evolved during its lifetime

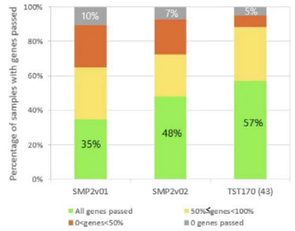


SMP2 screened around 7,000 NSCLC patients



- > 10,000 patients consented to SMP2
- ~ 7,000 patients tested
  (20% not tested - poor sample quality)
- Overall average QC failure rate < 20%
- NGS success rate: 79% for all patients with sample sent
  > 80% of sample QC pass
- 52% of patients molecularly eligible for NLMT
- Overall median testing TAT: 19 working days
- Overall median TAT from consent to results: 131 days (from consent to sample sent most variable)

The panel upgrade improved overall NGS results



Samples passed at least 50% genes:	65%	73%	88%

SMP2 screened around 7000 NSCLC patients from 2015-2021. However, in the end only 404 patients (6% of those screened) were enrolled in the NLMT. This was mainly due to failure of sample quality control, toxicity issues, or mortality.

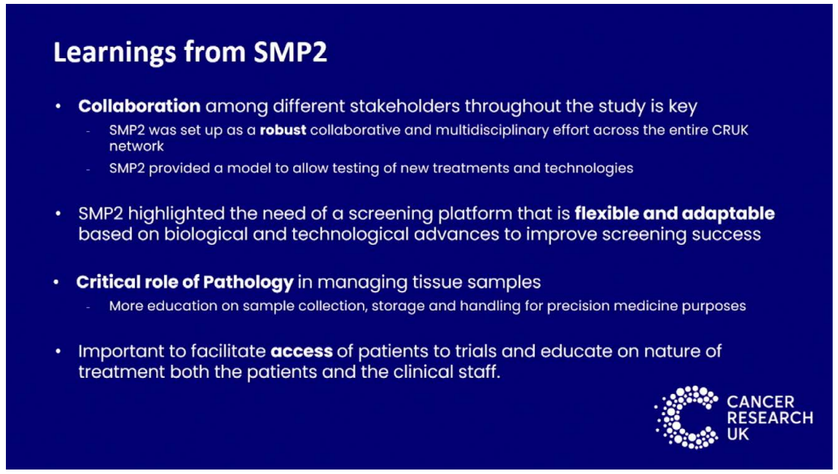
An important characteristic was that SMP2 adapted and evolved over the course of its lifetime. One of those critical evolutions was in upgrading the panel over time. The first upgrade in 2015 dropped the overall gene failure rate by 8 percentage points. Upgrading to the commercial Illumina panel in 2019 dropped the failure rate even further, down to 15%.

There were, however, some limitations with the new panel. They saw worse performance for somatic copy number alterations (SCNAs) for some genes frequently amplified or deleted in NSCLC. However, these aberrations were not very common and could be seen in other ways. So the team felt that the benefits of exploring more targets with the new panel outweighed the negative of a potential loss of sensitivity.

Conclusion

SMP2 demonstrated that routine genomic testing could be delivered at scale in a clinically relevant timeframe within NHS. This took advantage of extensive infrastructure spread throughout the country to ensure access for all patients. It also required continuous engagement of all stakeholders (clinicians, laboratory personnel, nurses, pathologists, pharma and technology partners, and patients themselves). And finally, it involved a flexible approach to the type of panel and analysis used, working collaboratively to implement those changes successfully.

SMP2 has been transformative of the testing UK landscape and paved the way for what is now becoming routine care in the UK. The learnings from SMP2 have helped to successfully set up genomics and other omics-based platform studies.

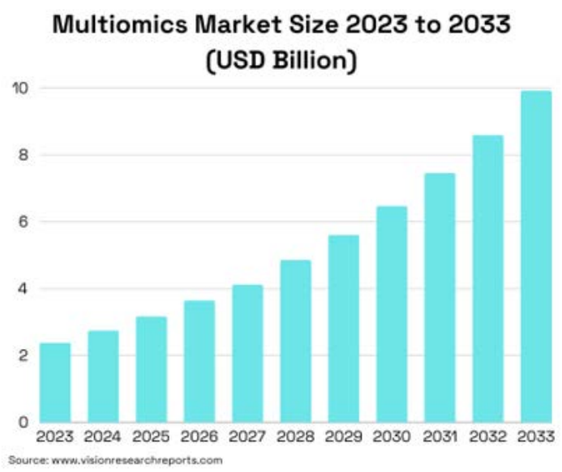


Market Report On Multi-Omics Techniques Personalised Care

One of the biggest challenges that healthcare will face with an older population are the increasing incidence of cancer. According to Cancer Research UK, new cases of cancer will hit 28 million a year by 2040 worldwide. That's an increase of 55% from 2020's figure of 18 million cases.

This dramatic increase underlies the need for early diagnostic tools for routine screening and more personalised approaches to the treatment of cancer. Non-invasive liquid biopsy has been one such focus in the area, driving the development of both diagnostic and response biomarkers.

Furthermore, reporting from Custom Market Insights predicts the global personalised cancer therapy market will grow at a CAGR of 8.2% until 2033, reaching a height of $707.1 billion. Big pharma companies like Roche, Novartis, BMS, Pfizer, and MSD are expected to maintain strong standings in this field but startups and biotechs will also enjoy a significant share of the sector.



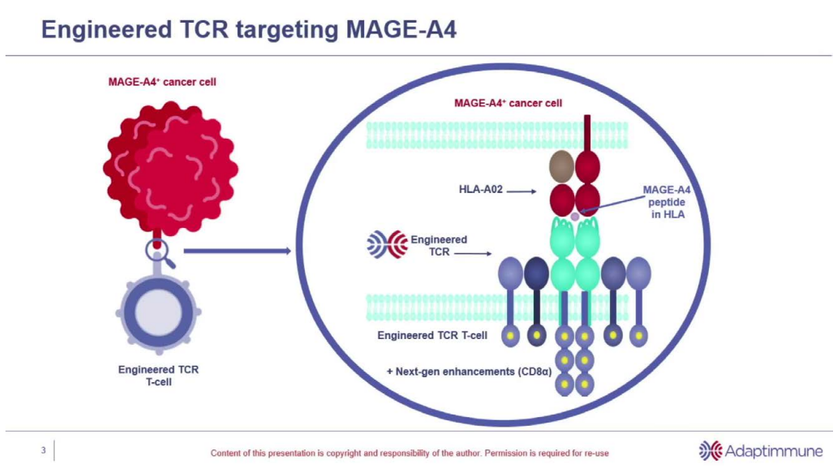
Using Spatial Technologies for Translational Medicine

In March 2024, DeciBio reported that the spatial biology market is projected to grow from ~$860M in 2024 to ~$2.27B by 2029 at an annual growth rate of 21%. The sector is currently dominated by 10x Genomics, NanoString, Akoya Biosciences, and Bio-Techne, together holding about 60% of the market.

Source: DeciBio

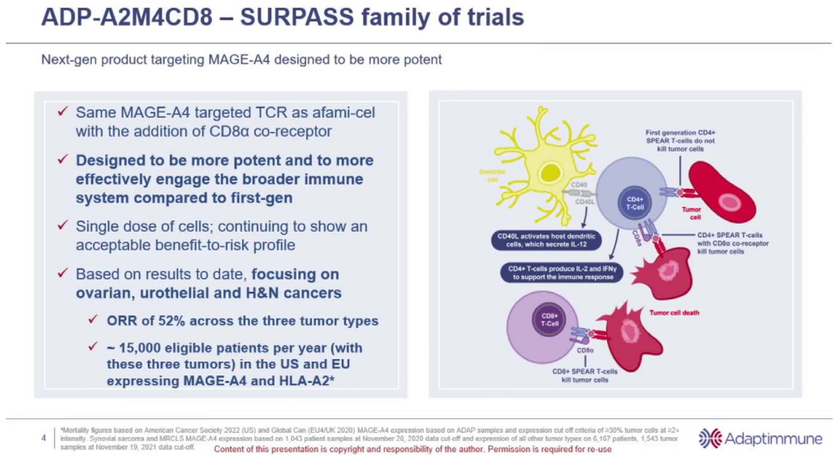
Targeting The Tumour Microenvironment – Assessing The Impact Of Cell-Based Therapies In Solid Tumours Through Multiplex And Spatial Biology

Adaptimmune is an immunotherapy company working on engineering T cell receptors to increase the affinity between the engineered T cells and cancer cells. Martin Isabelle (Associate Director, Tumour Profiling and Mechanistic Biology, Translational Sciences, Adaptimmune) outlined how his team use spatial tumour profiling to support Adaptimmune's investigations.



SURPASS family of clinical trials

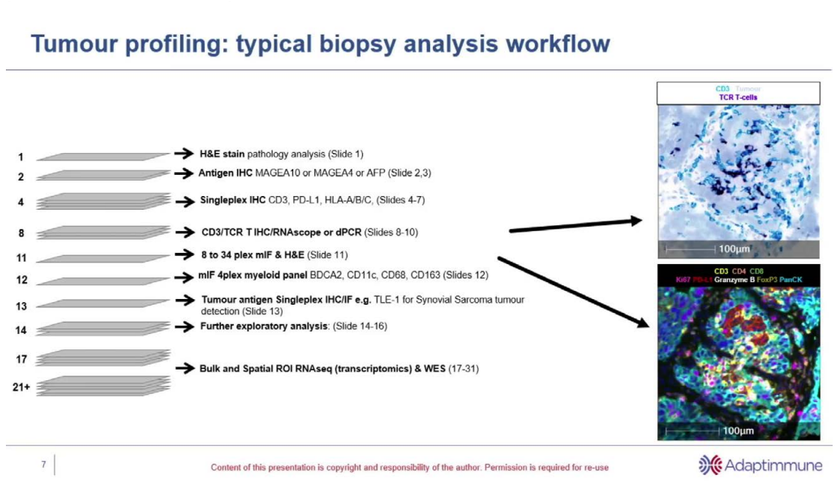
AdaptImmune's SURPASS family of clinical trials is aimed at testing this next generation of TCR T cell therapies. These therapies include enhancements like CD8alpha to increase the cytotoxic capability of the T cells. These T cells are engineered to target MAGE-A4 due to it being a common cancer marker. The team are simultaneously investigating other targets. The SURPASS trials will study a next-generation product targeting MAGE-A4 in ovarian, urothelial, and head & neck cancers, as these are the indications which have the best response rates.



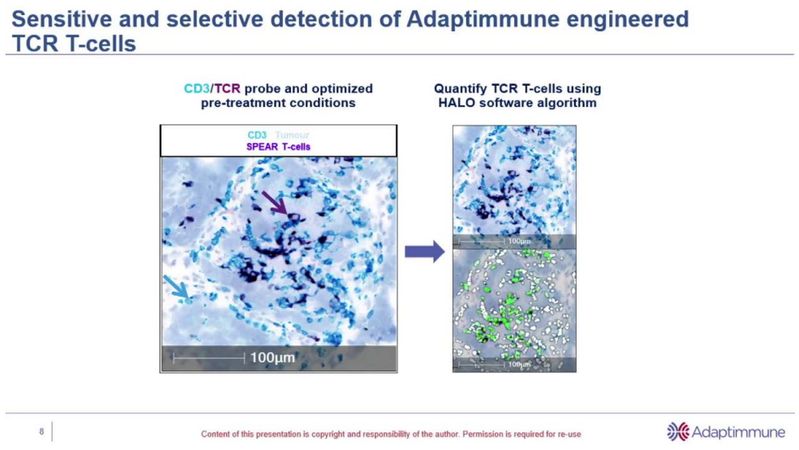
Translational assays deployed within tumour profiling team

The translational sciences team receives clinical samples based on post-infusion samples from patients. The team then profiles those samples and is tasked with figuring out if engineered T cells infiltrate and invoke anti-tumour response (by themselves or via endogenous non-engineered T cells). Here, the group must answer; 'Can we detect T cells in the tumour?' and 'Is there evidence of immune activation or suppression?'

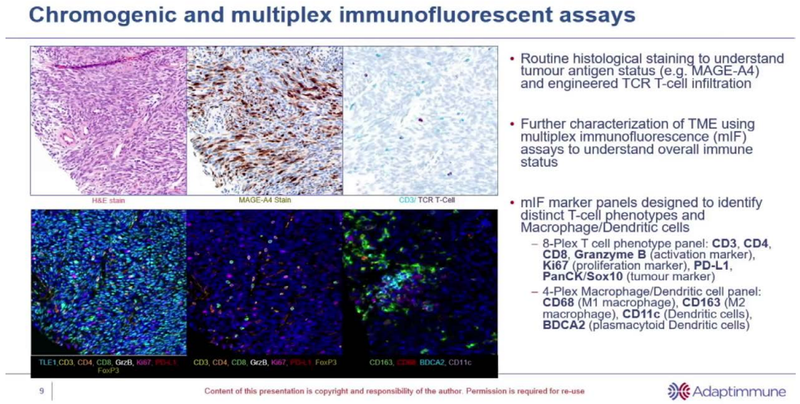
RNAscope is one of the assays that they use to detect T cells. Because the TCR has been engineered, the team has access to its sequence which can then be applied to RNAscope. Furthermore, they also look at CD3 so that they can find the total T cells - both endogenous and engineered. From there, they use Ultivue 8-plex immunofluorescence panels to further profile the cells within the tumour.



CD3/TCR panel



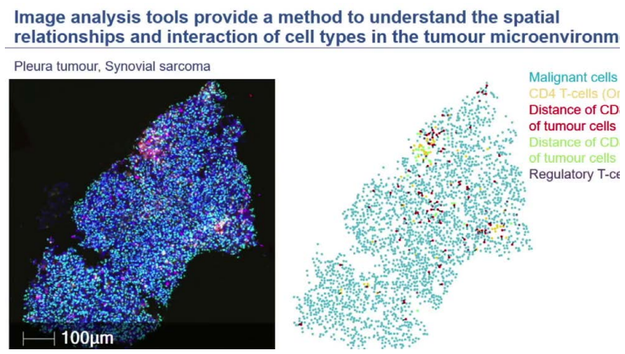
Isabelle then showed images from the CD3/TCR panel (image shown above). Purple areas indicate the engineered T cells, while areas in cyan are the CD3 cells. The team then look for the overlap of the CD3 and the engineered TCR T cells and use an algorithm called HALO to quantify the TCR T cells.



The Ultivue platform enables histological assessment using multiplex immunofluorescence and H&E staining on the same slide. The group use tumour antigen annotation to find regions of interest and build up their algorithms. This has ended up being a time consuming part of the process as thresholding every image with each marker bottlenecks the ability of the team. Therefore, they are currently looking for a way to automate this – either using third party or in house software.

Spatial Plotting

A nice feature of the technology is the ability to plot the tissue spatially. Their software allows for proximity analysis, so by performing nearest neighbour comparisons of tumour cells and cytotoxic CD8 cells, they can classify CD8 cells that are more or less than 10 microns from a tumour cell. Doing this for a number of samples allows them to build up a nice characteristic that includes the cells and their spatial metrics.



Conclusion

Understanding engineered TCR T-cell product and tumour microenvironment in clinical biopsies is vital for the translational team at Adaptimmune. Using Ultivue mIF panel has increased the company's in-house capability to stain for multiple different markers in a multiplex assay. This provides greater histological staining on a single slide.

Interesting insights into the tumour microenvironment across Adaptimmune tumour biopsies are now emerging using image analysis software such as Indica lab's HALO Highplex FL module. This has allowed the company to provide further context to the tumour microenvironment and the immune cell relationships present in tumour biopsies.

On top of this, there is scope to increase Adaptimmune's in house histology capabilities even further. This could include more multiplex panels involving RNAscope or IHC/IF and RNAscope/IHC/IF combined, expanding from 8plex to 35plex multiplex imaging, and Incorporating spatial transcriptomics into routine work.

Report Summary

This report has explored many key themes in the use of multi-omics and spatial technologies, primarily focusing on innovative methodologies and their application in understanding and treating diseases, particularly cancer.

The Surrey Ion Beam Centre are pioneering the integration of spatial imaging and mass spectrometry to comprehensively analyse biological samples. By combining elemental mapping with mass spectrometry imaging, they aim to understand the impact of various molecules on local chemistry. Challenges include integrating different imaging techniques due to varying data handling strategies, expertise, sample requirements, and destructiveness of techniques like ion beam analysis.

We also saw how machine learning can be used to identify patterns and predict outcomes from omics data. Here, interpretability and flexibility of machine learning models are critical. Here challenges like the "large p, small n problem" can be addressed through mathematical optimization techniques.

The Cancer Research UK Stratified Medicine Programme (SMP2) aims to provide genomic screening and personalized treatment options for non-small cell lung cancer patients. The program demonstrates the feasibility of routine genomic testing within the NHS, with continuous evolution and adaptation to improve panel efficacy and patient outcomes.

Finally, Adaptimmune uses spatial tumour profiling to support immunotherapy investigations. The SURPASS family of trials focus on testing T cell therapies targeting specific cancer markers, with translational assays deployed to detect T cell infiltration and immune response within tumours. The use of multiplex imaging techniques provides valuable insights into the tumour microenvironment and immune cell relationships.

As the multi-omic and spatial technologies sector expands, it will be vital to continue using interdisciplinary approaches to analyse the vast amounts of data produced. Furthermore, taking advantage of technological advancements like ion beam analysis and machine learning will not only boost efficiency, but take the sector across boundaries into new areas. Lastly, collaboration within and between companies – whether working hand-in-hand or joining data sharing initiatives – will further enhance biomedical research and personalized medicine for better patient outcomes.