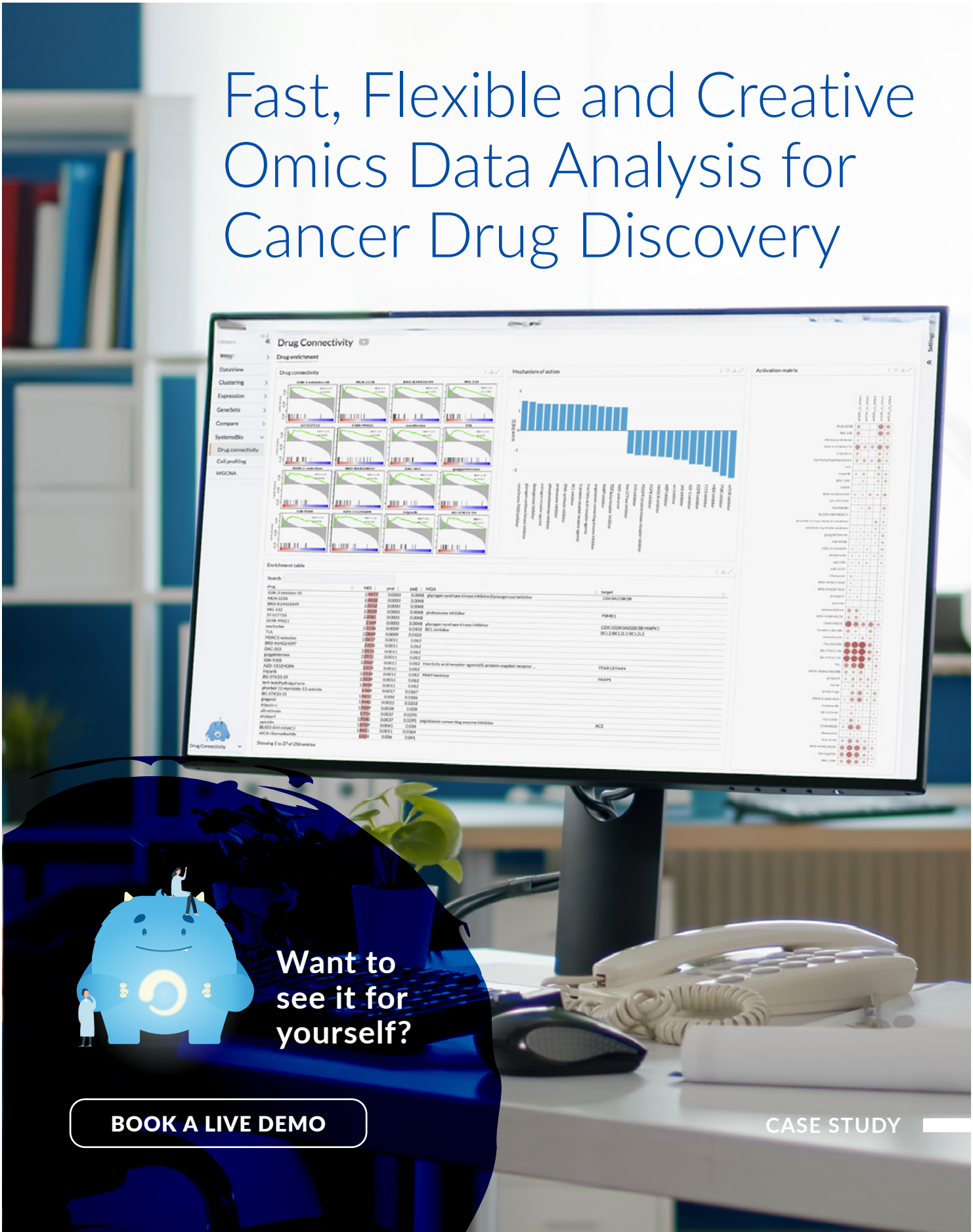
BigOmics Analytics®

Plexium

Fast, Flexible and Creative Omics Data Analysis for Cancer Drug Discovery



CASE STUDY

FAST, FLEXIBLE AND CREATIVE OMICS DATA ANALYSIS FOR CANCER DRUG DISCOVERY

Field: Cancer Drug Discovery
Location: San Diego, California

CLIENT PROFILE

Plexium is the premier, next-generation targeted protein degradation (TPD) company, leading the way in the rational design and discovery of monovalent protein degraders across a wide range of modalities. Plexium has developed a comprehensive approach toward Targeted Protein Degradation, powered by a proprietary best-in-class platform, designed to identify novel small molecules that induce selective degradation of drug target proteins through E3 ligase mediated proteasomal degradation. From molecular glues to monovalent degraders, Plexium is advancing a pipeline of novel targeted protein degraders for the treatment of cancer, neurodegeneration, and other diseases.

CHALLENGES

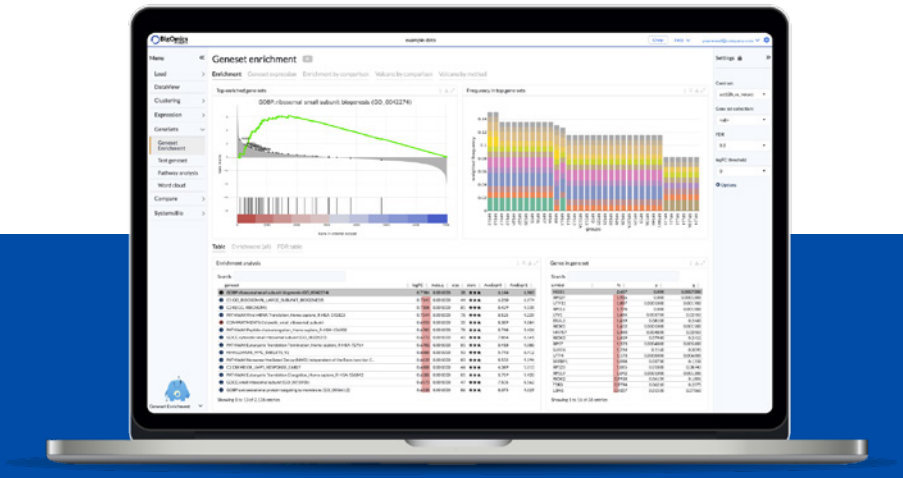
- **Prioritizing Speed:** The biotech industry requires fast data analysis as it has to compete with larger pharmaceutical companies.
- **Avoiding Siloed Data:** It is crucial for drug development to avoid isolated, siloed data that are difficult to combine.
- **Efficient Data Revisiting:** Fast access to data is essential, which requires the efficient retrieval of information from both old and new data sets.

SOLUTION

To implement Omics Playground so that data analyses can be carried out faster, more flexibly and more creatively.

RESULTS

- **Speed Enhancement:** Rapid data retrieval and reanalysis, reducing the time to answer specific questions from days to a mere 7 minutes.
- **Efficient Biomarker Identification:** Streamlined process for obtaining biomarkers in just 15 minutes, compared to the previous time-consuming workflow, which can take up to 3 days.
- **Simplified Data Extraction:** Easy extraction of meaningful insights from data. Simplification of complex data for easy understanding and efficient communication. A new, improved way of intuitively analyzing data that leads to better results.



BigOmics® Analytics

« The standout feature for me is the ease of obtaining biomarkers. It's almost too simple – just 4 clicks, and you're done. The process is so straightforward that it almost feels like cheating. »

KEN STEADMAN, PHD
COMPUTATIONAL BIOLOGIST

INTRODUCTION

Ken Steadman, PhD, Principal Scientist at Plexium uses Omics Playground extensively to analyze omics data for cancer drug development.

Ken's academic career began at the National Cancer Institute, where he examined the SEER (Surveillance, Epidemiology, and End Results) data correlated with expression profiles from SAGE (Serial Analysis of Gene Expression) databases, to identify new targets that drove intractable tumors. At the same time Ken and his colleagues were working on clinical trials of the Histone Deacetylase inhibitor Romidepsin, the Bates lab co-discovered the multidrug resistance gene ABCG2 and characterized its role in tumor multidrug resistance. The Bates lab were looking at the whole spectrum of cancer drug development, from discovering new targets to overcoming drug resistance to getting a drug approved in the clinic.

Ken then went to the Lombardi Cancer Center, where he worked in a clinical laboratory looking for ways to combine existing drugs to develop new treatments. He developed precision medicine clinical diagnostic SNPs (Single Nucleotide Polymorphisms) and associated their relationship to drug resistance.

More recently, Ken conducted prostate cancer research at Cedars-Sinai Medical Center investigating a new master regulator of neuroendocrine prostate cancer. He has worked through every stage of drug development, from in silico studies to identify new targets, all the way through clinical trials for FDA approval, and uses his experience from clinical research to set priorities for the development of new therapeutics.

How did you get started in bioinformatics?
I started my career as a pure molecular biologist and later learned how to use bioinformatics tools to help with the identification of new targets and hypothesis testing. If you pursue a simple question, you can usually repurpose data, and you quickly realize how much work has already been done. It's about using data orthogonally, in a way it wasn't initially intended to be used, and conducting integrated analyses to inform your questions. You can get a pretty good idea of whether something is going to be valuable or not by repurposing data.

I was studying microRNA and competitive endogenous RNAs interactions so I need to learn network analysis. I'm still a biologist, but a biologist who understands bioinformatics and can set up a pipeline. If you know how everything is connected, you can design your project to get the best answer.

« During my PhD, the main reason I wanted to do my own bioinformatics was because it's quicker and easier and you're not dependent on someone else. »

How did you get introduced to the world of omics data analysis?
At the time, this data was not really known as omics data, it was called informatics or big data. But even then, SEER had years of survival rate data for patients across all tumor types – a huge amount of data. We came across SAGE and thought: what if we could link survival to expression profiles and create fingerprints? If you look at it as an axis of intractability, we hunted down the edges of that axis and determined what is easy to treat, what is hard to treat and how do they differ? In this way, we identified and prioritized potential new targets.

What challenges did you encounter while analyzing transcriptomics data when you entered the biotech industry?
Now that I work in a biotech company and deal a lot with different types of omics data, especially transcriptomic and proteomics data, the priority has changed. The emphasis is on finding efficient ways to extract meaningful information.

« Speed is key. Plexium is a small company, and one way to be competitive is to be agile and make decisions quickly. »

It's about having a clear hypothesis, focusing your efforts, and designing the right experiment to get the most informative answer. I focus on integrating compound binding data, proteomics and transcriptomics datasets and take into consideration the differential kinetics between proteomics and transcriptomics in interpreting the analysis. Often we have to make compromises and weigh things against each other: How can we get the best data so that we can make an informed decision quickly?

« Integration science is the most "teamy" of the team sciences. No man is an island. You have proteomics experts, transcriptomics experts, the RNA core, the proteomics core, all working together, so it is necessary to find compromises to generate data quickly. »

What did you like most about academia and working in a biotech company?
In the academic world, I enjoyed the freedom of conducting basic science research that was of course aligned with my PI's area of scientific expertise – you can't just pick anything at random. I also really liked my time working in a government lab because funding can be more straight forward and you don't have to spend as much time writing grants, and still get to work with a great team of scientists.

I like discovering an interesting problem and asking myself whether I can do something with it. That really helps me to focus on experimentation and project planning. How can I design experiments that answer the question without wasting months of time to generate data? How can I approach this question in a way that I can answer 80 percent of the question in two weeks instead of making it a lengthy project of four months? It's a challenge, but I can be much more agile this way.

Compared to the academic world, what I like about working in industry, especially in a smaller biotech company like Plexium, is that we move fast, act agile and make decisions quickly. We deal with a lot of different types data, especially transcriptomic and proteomics data. It is really about focused teamwork, and designing experiments to get the best answer quickly. If I discover an interesting problem, I think about how I can design an experiment that answers key questions in a couple of weeks, rather than making it a resource consuming project.

One of the main challenges in research is that of "siloed data", i.e. the fact that different datasets are often scattered around and it is difficult to maintain an overview. How does that challenge manifest itself in a biotech company and what other challenges are there?
Experiments usually focus on a specific question. Conducting a massive, generalizable experiment is impractical because it takes a lot of time and you would potentially go into an omics black hole where you're endlessly sifting through data. It requires focus. However, when designing an experiment, obtaining a singular answer and then leaving it untouched does not fully extract the value, labor time, and costs invested.

Therefore, the emphasis is on avoiding siloed data and instead generating data that can be revisited. Drug development involves the evolution of compounds and targets. The ability to quickly revisit old data and compare it to new findings is crucial. Additionally, quick access to data is important, for example when preparing for presentations or writing papers.

With a homegrown dataset, decisions about the importance of certain elements have to be made early on, and it could easily consume half a day to revisit old data, retrieve the code, and recall important details. This shows that it is really important to extract value from experiments rather than just addressing one question and never look at it again.

How did you start collaborating with BigOmics?
We found ourselves at a nexus, generating a lot of data, and asking how we were going to analyze it all. We needed more than our existing homegrown systems and thought it would be really useful to have a third-party solution that could help us speed up data recovery and reanalysis based on old and new datasets.

It's not just about generating data, it's about extracting meaningful insights. The last mile, simplifying the data for others to understand, is particularly crucial, and our inhouse solution wasn't effective in addressing this last mile.

« It felt like a customized solution tailored precisely to what I needed. »

When we came across BigOmics and started working together, both companies underwent a process of growth as we tackled increasingly complex problems. The BigOmics platform evolved to become more and more in-depth and focused on what I wanted and needed to get out of it. If there was a need, I could simply express it and either submit a ticket or discuss it. It felt like a highly customized solution. The platform helped me address the challenges I was facing with data analysis.

What was the impact after using Omics Playground?
The most important aspect is speed. If I can quickly revisit the data, I can answer a question within just 7 minutes. Answering questions has become much easier because I can quickly refer back to relevant information. It's now a seamless experience to compare one dataset to another, like assessing the performance of a drug in one case versus another drug used in a slightly different scenario. Previously, this would have been a day-long process just to ensure that the right approach was chosen.

« No more questions like. “How am I going to write this code?” I can accomplish that in 7 minutes now. I get a figure, create a slide deck, and it’s done. Boom. It's a real time-saver. »

Which features do you like best about the platform?
Obtaining a biomarker has become trivial at this point. With correct data structuring, loading, setup, and contrast adjustments – boom – the biomarkers are there, and you can move forward.

« The standout feature for me is the ease of obtaining biomarkers. It's almost too simple – just 4 clicks, and you're done. The process is so straightforward that it almost feels like cheating. »

Biomarker analysis now takes a maximum of 15 minutes, – a huge time saving. Before using Omics Playground, it would take me about 3 days to complete the same tasks. Now, I can have the biomarkers in 7 minutes, and within another 7 minutes, I can confirm them, explore additional insights, and analyze them against old data. A 15-minute turnaround is a substantial time saver, and time is one of the main aspects when doing research in the biotech sector.

What do you explore first in a data set? What drives you to explore the data?
I focus primarily on mechanism of action and biomarkers. First thing I do is I want to see the structure of the data. From a pure bioinformatics perspective, you look for p-values and look at it quantitatively. From a biologist's perspective, you're looking for qualitative pathways, almost like themes that emerge, and therefore the ability to look at the data from a holistic point of view. Of course biomarkers are important, but I like to jump back and forth between the most significant gene changes and the geneset enrichment to give those genes context.

I might have a handful of genes, go straight into the dataset and look at them to immediately see what kind and how much variation is observed and then compare qualitatively with a previous experiment: Did I modulate those target genes? Then I go back and maybe do a Spearman correlation from that data to see if there's anything new in that theme that I'm looking for. The process is almost like a stream of consciousness analysis of the data. I jump from module to module until something interesting pops up, and then I make notes on the data. I also like not only taking screenshots, but opening images organized across tabs so I can put a story together – something I would never do in R. I couldn't possibly put together a storyline with actual figures in them. With Omics Playground it's an iterative loop, a little more free-form. I may discover something new because I am not focusing on making sure the next line of code works.

« It's liberating, especially as a biologist, when I don't have to limit myself to whether I can translate and implement my question into a line of code without having to rewrite it ten times because I forgot a comma. Instead, I let the platform inform me of things I didn't expect to find. »

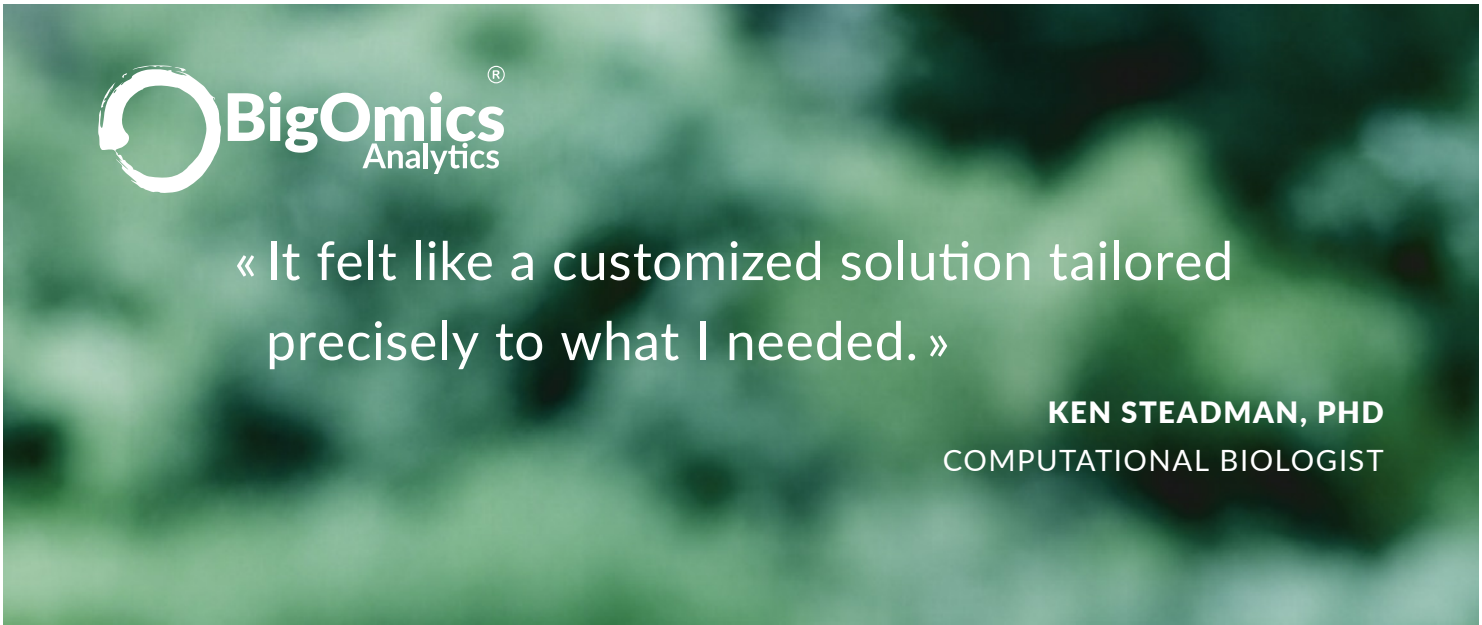
When working with new data, have you ever had the experience that the platform suddenly gave you unexpected insights?
That's pretty common, and you collect themes because when you're working on the same target in the same chemical space, you develop those themes. There's a certain amount of serendipity that you can get from looking at data this way that you can't get from plowing through R. There really is no such thing as serendipity in R. Every change requires a new piece of code that you are going to write, a new module that you have to load, so there is no such thing as a chance encounter. Perhaps this will one day be possible with AI applied to R.

What developments in biotechnology are you most looking forward to in 2024?
Everyone hopes that AI will solve everything, and I think that certain areas are better suited to this than others. I believe that AI can help very well in the search for biomarkers. You can use completely different types of data, and AI will be able to draw conclusions that even the smartest scientist could never come up with, but it will be quite difficult.

AI is already being overused in other areas and could lead to garbage data. I recently read a paper on AI that already reads AI. This brings us full circle, ChatGPT reading data generated by ChatGPT. There is a real danger here of generating and spreading misinformation.

And then, of course, the TPD field is interesting. We are waiting for the first tranche of patient data to come back soon. That's going to be very exciting. It's a brand new field with some big questions that still need to be answered. So, AI for biomarker discovery is certainly very promising and hopefully for development as well.

At BigOmics we support biotech and pharma teams in omics data analysis by providing highly interactive and user friendly analytics for transcriptomics and proteomics data. If you would like to learn more about the Omics Playground platform, visit our website at bigomics.ch or reach out and we'd be happy to walk you through the platform and its functionalities.



BIGOMICS SOLUTIONS

CENTRALIZED, COST-EFFECTIVE DATA ANALYSIS STREAMLINES SCALING AND IMPROVES PRODUCTIVITY

- RNA-Seq data are analyzed through peer-reviewed algorithms, so scientists can quickly identify the most promising therapeutic targets without requiring any coding knowledge.
- Moreover, more than 50,000 public gene sets and pathways such as GO, REactome, Hallmark, and Msigs can be accessed, as well as drug connectivity and drug sensitivity databases with more than 30,000 drug expression profiles.
- As more experiments are performed, the newly added data sets can be compared to previous results and more than 6,000 public data sets, providing the necessary context for scientific breakthroughs.
- All omics data is in one place, so scientists spend less time retrieving it for new analysis.

INTERACTIVE VISUALIZATIONS GIVE LEADERS A 360-DEGREE VIEW OF R&D

- Omics Playground helps managers and executives have an eye on day-to-day progress while also providing a cohesive overview of all different research projects.
- Since Omics Playground has improved the reproducibility of data, it's much easier for leaders to forecast and set informed timelines.
- Across numerous projects, many decisions have to be made rapidly yet confidently. With the full history of their experiments at their fingertips, executive can make prompt, data-driven decisions.

A UNIFIED DATA ANALYTICS SYSTEM ALLOWS FOR INTERDISCIPLINARY COLLABORATION

- Repetitive and time-consuming data analysis iterations can be avoided because biologists are able to gain insights from their data directly from the platform.
- Bioinformatics resources are conserved, allowing them to process requests faster.
- Everyone can easily access the same version of data and share results in team meetings, as scientists can present directly from the Omics Playground.

Want to see it for yourself?

BOOK A LIVE DEMO

LinkedIn

bigomics.ch
hello@bigomics.ch

BigOmics® Analytics