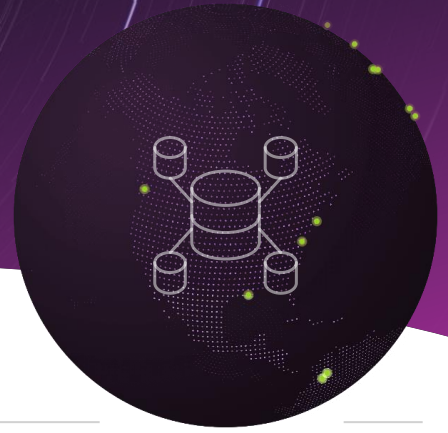# Introduction to Euformatics with Element Biosciences

The future appears to us through a series of doorways – there is always a door to the future, and behind that door the landscape changes, again and again. In genetics, the central dogma opened to antisense RNA, followed by a time when broken genes were anathema because splicing didn't exist. The simple Eppendorf tube was the home of sequencing. The human genome was only an idea. One door opened on big data, and with it came both physicists and computer scientists who changed the landscape of data analysis.

Euformatics is a software company founded to address the challenges that came with high throughput data processing such as NGS, and interpretation in the landscape of clinical research and diagnostics. The company develops tools for ensuring there is no sequence garbage going in and no garbage reports coming out. Quality control, internal validation of the analytical NGS methods, and external proficiency testing or EQA were among the company's first products to market. These were followed by variant interpretation software, evolving among more complex challenges in genomics and as new variant types become amenable to sequencing.

The first customers of Euformatics worked on small gene panels, limited by chemistry and technology. Today, clinical research customers work on very large panels, clinical or full exomes, while the research projects where Euformatics is involved study whole human genomes.

In this changing landscape, Euformatics software bridges the gap between research and healthcare, translating discoveries into more actionable outcomes. Laboratories have also chosen to collaborate with Euformatics to identify and assess their next purchases, evaluating NGS kits and sequencers. In such exercises the company has been able to give an unbiased assessment of performance and quality and document the progress brought in by new players in the clinical research NGS space such Element Biosciences/AVITI, which has made sequencing faster, more efficient and accessible to more scientists.

Element Biosciences is a pioneering life science company democratizing access to advanced biological tools, driving impactful discoveries to benefit humanity. We are focused on developing disruptive DNA sequencing and multiomics technology for research markets.

Throughput considerations

When a laboratory uses an AVITI sequencer together with exome chemistry and aims for high confidence analytics at a read depth beyond 100x, it means hundreds of microscopic volumes of solutions will be handled by a lab robot before hitting the flow cell of the sequencer. And when the sequencer stops at the end of a run, there is half a terabyte of raw data that high-quality move from the sequencer through a pipeline of demultiplexing, read alignment, variant calling, quality control steps and eventually variant annotation and prioritisation. Frequently these steps will be on different computing platforms, meaning that data transfer time becomes part of the operation on top of computing itself. High throughput therefore becomes a project requiring a considerable amount of different expertises, from wet lab and liquid handling robotics to algorithms and IT hardware. Understanding and collaborating with each other is key.

Euformatics uses cloud and on site high-throughput computing technologies to maximise the efficiency of algorithms, their security and shareability across platforms. Automation is implemented as much as possible to reduce human errors and to allow a process that can take hours to progress supervised but without the need for human intervention. Much of this automation can be developed together with the sequencing instrument provider, but some might need customisation depending on customer-specific work.

Quality

Euformatics has at a very early stage of the NGS era been involved in supporting EMQN and GenQA with computational aspects of external quality assessment (EQA) globally. In parallel, it developed automated quality control (QC) tools for individual NGS-based laboratories. The QC and validation tools can be used as separate modules or integrated via API connectors to other Euformatics or high-quality tools used or developed by the laboratory. Automating QC has allowed Euformatics to relieve the laboratories from a considerable amount of necessary control steps required to provide confidence in the quality of the analysis results. Quality is no accident.

Christophe Roos, Euformatics Chief Scientific Officer, comments the data quality that Element Bio AVITI produces:

"The AVITI platform offers an exceptionally high quality of the sequence observed not only as a high Phred score also as very few missed bases, high uniformity and low error rate. The possibility to operate some of the chemistry on the flow cell with the Trinity workflow also reduces handling effects and leads to significantly higher sample-to-sample reproducibility, lower read depth requirements even for more tricky variant types such as Indels and CNVs. Overall, it allows sequencing more samples for the same costs, improving the cost per sample while maintaining a very high-quality output."

Flexibility

In a lab context, flexibility is the capability to mix and match test requirements such as capture or amplicon-based kits, different sequencing chemistry and in silico analysis. For Euformatics, this translates into modular software development corresponding to the steps in laboratory workflows. It means development of interconnectivity of the modules via APIs, but also to other software tools already available in the NGS laboratories. Flexibility requires high security and strict formalism for data modelling and application of public or de facto standards in file formats and terminology to allow smooth data exchange. To keep up to date with, or even be a bit ahead of the evolution of NGS applications in clinical research, Euformatics participates in international research projects and collaborates with various organisations active in the field of genomics, such as database providers (e.g. ClinGen), standardisation organisations (e.g. GA4GH), and of course other NGS innovators.

According to Vincent Lehmann, General Manager at Ouilab,

"Within the Ouilab Group we have different ways of working and it has been great to see that Euformatics, together with AVITI, conform to our ways instead of us having to change our processes to comply with the tools' requirements. For example, some labs prefer to analyse and store all data locally while others are comfortable using cloud-based solutions. Genomics Hub has proven to run very flexibly in different IT setups".

Automation

Robotics takes care of much of the wet lab work. Today, most data handling happens without much human interaction. Euformatics modules can be set up to react to the end of the sequencing procedure and sample demultiplexing so that it can automatically kick off the secondary pipeline (raw data alignment and variant calling), quality control with coverage analysis, as well as variant prioritisation used in reporting. The APIs providing access to the different modules can also be the contact points for other software available in the laboratory, such as internally developed analysis tools, data exchange, or laboratory management systems.

In the tertiary analysis automation is built into every layer of the bioinformatics stack and workflow. This Full Stack AI approach starts from the OS layer where the security patches are installed automatically. The Database layer learns the data (volume, distributions, skews) and optimizes the queries based on the data statistics. The Database further studies the querier in the background and based on the learnings optimizes the queries even further. AI Flags are used for automating the search for the needle in the haystack i.e. finding the pathogenic variants in the correct disease context. Finally, GenAI can be used in the analysis to give additional insight about the patient.

Francisco Garcia, SVP of Informatics at Element Biosciences, comments on the role of bioinformatics automation:

"Data that comes from the AVITI system is only as good as the ability to manage it. Clinical research is crucial to developing potential new treatments or diagnostics for disease, and Euformatics Genomics Hub has been an invaluable partner in creating new, fully automatic clinical research workflows that can effectively work with the information arising from AVITI."

Cost

The standard narrative in the NGS industry has been that the cost of sequencing has fallen considerably over the last 15 years. A major driver for the cost reduction per sequenced sample has been the development of higher throughput sequencers. The cost per megabase of DNA sequence has dropped faster than Moore's law would imply and has brought the cost of human genome sequencing from $10 million in 2007 to below $ 1,000 now.

In the clinical context, however, the cost of sequencing has to take into account the lab's actual sample flow. Very few labs can completely fill an Illumina NovaSeq X, a very high-capacity sequencer, multiple times a week to achieve the lowest marketed sequence cost. Also, clinical samples have a certain urgency due to patients needing a diagnosis soon to start treatment so the laboratory cannot wait to fill the sequencer. Thus, flexible mid-throughput systems such as AVITI can show the way to the right balance between speed and cost.

Jean-Marc Costa, medical biologist at Cerba Healthcare, concluded that

"The combined flexibility of the Element Bio AVITI and Euformatics bioinformatics suite allow us to run sequencing frequently while maintaining affordable per-sample analysis cost. This results in faster time-to-diagnosis compared to a situation where we would need to wait longer to fill a higher throughput instrument".

Cerba HealthCare

✓ Conclusion

The nature of research in general, and the constantly widening scope of molecular genetics and genomics in particular impose high requirements for software flexibility. It is common to build teams covering multiple scientific disciplines–depending on perspectives automation and flexibility will be antagonists. For clinical work high level quality standards such as ISO-13485 and IVDR have to be applied. Euformatics, in collaboration with Element Biosciences, can provide expert support with complete systems or adaptable parts thereof to NGS-based laboratories aiming for high efficiency and quality. Be confident when opening new doors to the future.

Euformatics   Element Biosciences