


# Bringing complexity into clarity with Qlucore Diagnostics: Clinical use of transcriptomic data to classify cancer patients

*By* Erik Söderbäck, Head Diagnostic Sales, Qlucore

CANCERS ARE COMPLEX and heterogeneous diseases that require analyses of a multitude of genetic, cellular, and tissue alterations for correct clinical diagnosis, for disease and risk stratification, for treatment selection, and for follow up. Ultimately, the most informative biomarkers should be used for analysis.

## The roles of DNA and RNA in Cancer Diagnostics

Molecular diagnostics of cancers have traditionally meant genotyping driver mutations and other DNA variants indicative of pathological causation. However, using information solely from DNA limits the potential of cancer diagnostics to the static information borne in inherited and somatic genetic variants. In contrast, transcribed RNA molecules reflect the phenotype and the dynamics of cancer cells and tissues. Moreover, RNA provides information about gene fusion events which are of great importance for the development of many cancer forms. Therefore, RNA analyses

are valuable, or even required, for cancer diagnostics.[1,2,3,4] For clinical purposes fluorescent in situ hybridization (FISH) and hybrid capture arrays have up until now been used to analyze RNA, but RNA sequencing is rendering the most informative and versatile analyses.

## Sequencing technology unlocks the information in DNA and RNA

During the last decade the introduction of massive parallel sequencing, or Next Generation Sequencing (NGS), has meant a paradigm shift in oncology allowing for more and better precision personalized medicine. The NGS technology and methodology development, accompanied by a sharp decline in costs, has brought a revolution regarding the quality and amount of genetic information that can be gathered for diagnostic purposes. Highly informative sequences can today be generated also from difficult patient samples such as Formalin Fixed Paraffine Embedded (FFPE) tissue and liquid biopsies (plasma, urine etc.). Moreover, the development and streamlining of bioinformatic tools and software have enabled clinical laboratories outside of university hospitals, with limited bioinformatic competence and resources, to start using NGS in oncology applications – although a lot remains to allow all cancer patients to be diagnosed with these advanced techniques.[5]

In this paper I describe how Qlucore (see About Company) is developing an IVDR compliant analysis software solution for clinical sub-classification of different cancer forms based on transcriptomic data generated from standard RNA sequencing procedures. Combining DNA sequencing with RNA sequencing provides a more complete and holistic profile of a cancer patient. For example, it was recently shown that adding transcriptomic analysis to standard DNA typing of Acute Myeloid Leukemia patients brought a 5-fold improvement in predicting drug response of patients.[4] Multiomic approaches for diagnostic purposes are anticipated to becoming more widespread in the future.

## RNAseq for analysis of gene fusions in cancer

Panels for targeted NGS RNAseq for gene fusion analysis have lately become relatively common tools for oncology applications in clinical laboratories. By including whole transcriptome data, however, gene expression profiling can be added to the analytical toolbox for oncology. Therefore, even if limited efforts have been made to use and harmonize transcriptome analysis for clinical routine diagnostics,[6] RNA sequencing procedures and analytical tools represent a forward step in development for personalized medicine in oncology.

> *"Gene fusion events happen when a gene is split and the opened ends of the chromosome ligate with other parts of the genome, either on the same or on a different chromosome."*

Gene fusion events happen when a gene is split and the opened ends of the chromosome ligate with other parts of the genome, either on the same or on a different chromosome. If the split gene end is translocated into another gene, it may result in expression of a chimeric protein. The gene merged into another gene is not controlled by normal cellular transcriptional regulation. If the translocation involves an oncogene the change in regulation may exert a cancerogenic effect on the cell.

Gene fusions have been shown to have great importance for the development of different cancer forms.[7] This is manifested by the many drugs approved for treatment of a variety of different cancers for which the presence of a particular gene fusion has been identified and shown to have clinical significance.[7,8]

Several vendors during recent years have launched RNA panels for NGS (based on, e.g., amplicon or hybrid capture technologies) that target RNA molecules resulting from specific gene fusion events for which there are approved treatments. On the other hand, hundreds of gene fusions not associated with specific treatments have been reported to have clinical implications in oncology, and the number is rapidly increasing.[9] The identification of numerous new clinically important fusions has meant that in research settings total RNAseq data is increasingly used for unbiased identification of all gene fusions present in a sample. Also, for clinical use, rapid medical development and scientific findings favor the introduction of full transcriptomic data for fusion analysis since the design of a targeted RNA panel will be outdated for clinical use in a relatively short time.

## RNAseq for gene expression profiling in cancer

Gene expression profiling entails quantitative analysis of RNA molecules in a sample and provides information about development and dynamics of a cancer. For example, expression profiles can be used to distinguish sub-types of different cancer forms and to stratify the clinical

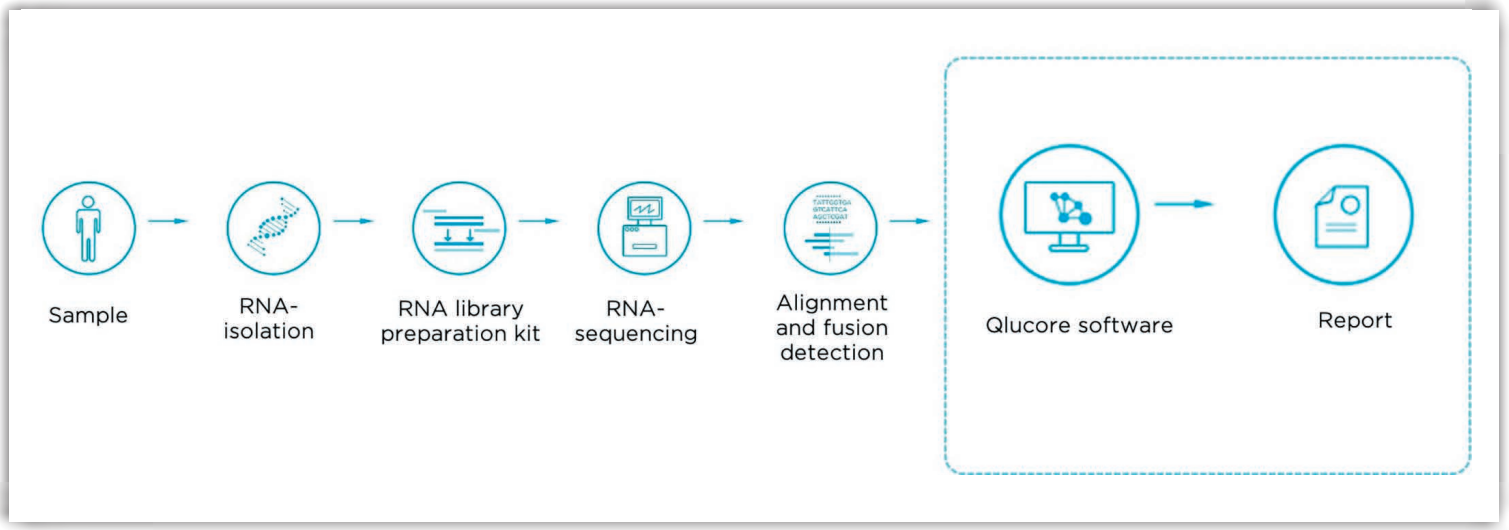


**Figure 1:** A Standard Operating Procedure used for the laboratory workflow upstream of Qlucore Diagnostics.

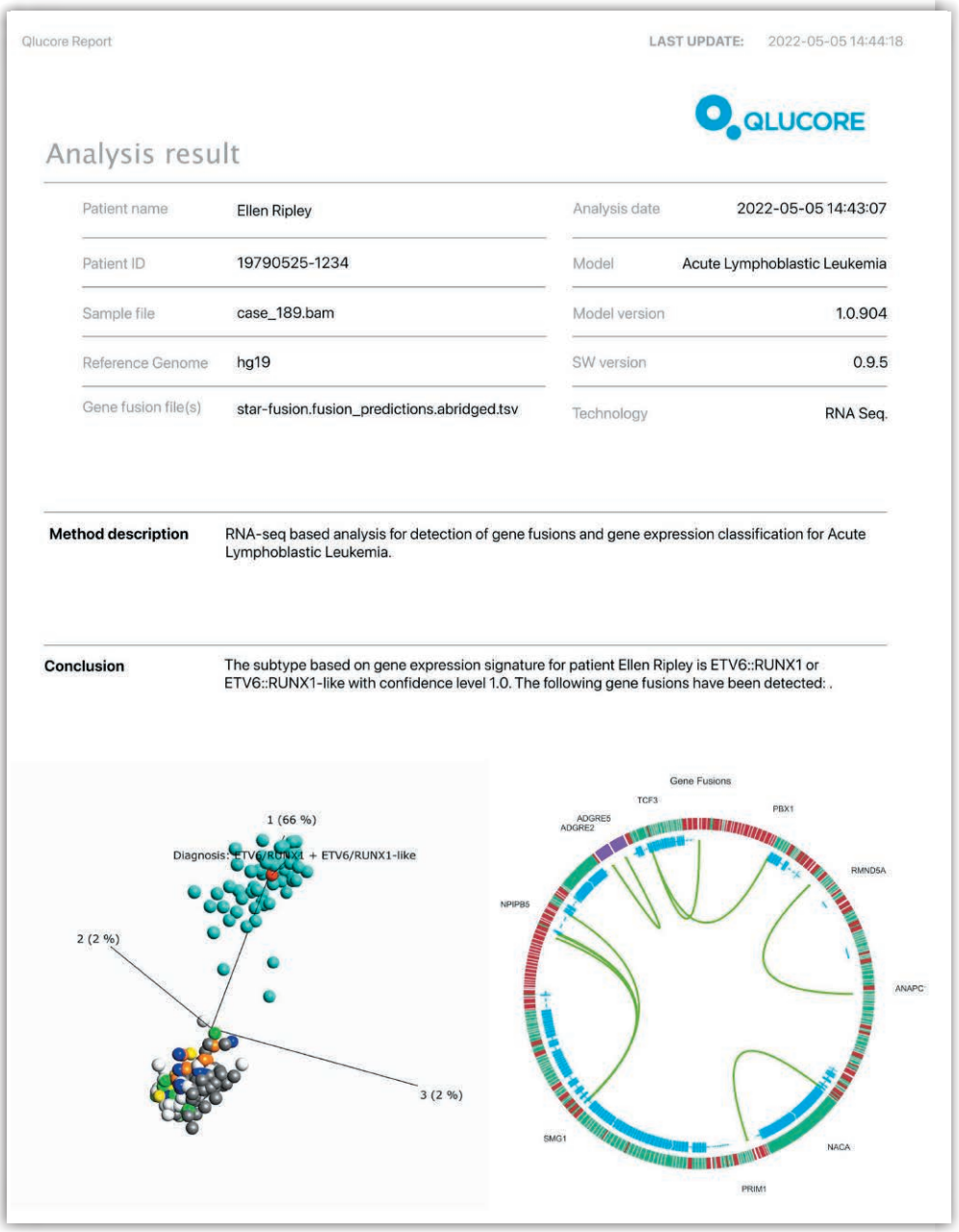


**Figure 2:** First page of an example report from the ALL classification model of Qlucore Diagnostics. To aid clinicians in their understanding of the reported results the gene expression classifier model puts the patient in context with reference sample data in an interactive principle component analysis plot based on the expression signatures and presence of gene fusions. The classified sample, in this case an ETV6/RUNX1 subtype of ALL, is indicated in red in the plot and the other colored dots indicate reference samples representing all subtypes of the disease which were used to train the classification model in a machine learning process. The circular plot presents gene fusions identified in the patient sample. Before reporting, fusions may be filtered for published clinical importance using the built-in Mitelman database. The report functions and layout are customizable.

handling of patients. In a study of metastatic breast cancer patients, the use of expression profiling clearly distinguished the patients with poor prognosis. Regarding the patients with BRCA1 mutations the authors concluded that "gene expression [profiling] will outperform all currently used clinical parameters in predicting disease outcome."[10]

Another example is Acute Lymphoblastic Leukemia (ALL), the most common acute leukemia in children. Analyses of gene fusions and classification based on transcriptomic data are needed to classify ALL patients into one of the 6 subclasses identified by WHO as decisive for clinical handling.[2,11]

## Qlucore Insights and Qlucore Diagnostics

Qlucore recently launched Qlucore Insights (QI), a software package designed for analyzing clinical data. In parallel the company develops the IVDR compliant Qlucore Diagnostics (QD, for release in 2023) built on the experiences from QI. The general software platform of QD is disease agnostic – that is, different disease-specific analysis models can be connected. The first applications for QD will be diagnostic tools for cancer sub-classification using transcriptomic data in a streamlined analysis solution.

Sample quality in sequencing is critical, so regardless of the disease, the upstream laboratory workflow starts with a Standard Operating Procedure (SOP) to generate the RNAseq data from Illumina platforms used in the analysis model (**Figure 1**).

*"Besides the cancer classification modelling in QD, the RNAseq data is also used for unbiased gene fusion analysis in a dedicated analysis model. All identified fusions can be reported; alternatively, the clinical importance of the fusions can be evaluated according to clinical guidelines using the built-in Mitelman fusion database."*

The software is installed locally on a client's system (behind a firewall) and is not dependent on internet access. To promote ease of use, the user interaction for setting up an analysis is kept to a minimum: importing RNAseq BAM files from Illumina sequencers, connecting sample information, and choosing the relevant cancer classification model. The data analysis typically takes a few minutes on a regular PC. Samples can be analyzed in batches.

Besides the cancer classification modelling in QD, the RNAseq data is also used for unbiased gene fusion analysis in a dedicated analysis model. All identified fusions can be reported; alternatively, the clinical importance of the fusions can be evaluated according to clinical guidelines using the built-in Mitelman fusion database.[12] Identified fusions are also used in the cancer classification models. Hence the Qlucore platform can classify cancers not only by sequence data but also by sequence *and* gene fusion data.

When the analysis is finished, a fully customizable clinical PDF report is produced automatically. The report can be integrated into standard hospital data management systems for consideration by clinicians. An example of the first page of a report is shown in **Figure 2**. In Qlucore Insights, which supports the same type of data

analyses in a research setting, the user can interact and control the results before setting up the final report. PCA plots help the user to understand the results from the statistical analysis. Circular plots and a built-in genome browser help the user to understand sequence details (e.g., breakpoint coverage analysis) about identified gene fusion events. The classification results are presented as confidence scores in a table format as well as visualized graphically.

## Development of cancer sub-classification models

As of today, a classification model for Acute Lymphoblastic Leukemia and a general Gene Fusion analysis model are validated and implemented in QI. Models for classification of several other cancer forms are in development (e.g. non-small cell lung cancer, bladder cancer, breast cancer). All models will be validated and documented for IVDR compliance and thereafter integrated in QD.

The disease-specific classification models are developed in collaboration with leading pathologists who have deep biological knowledge about a certain disease and access to patient samples. Reference cancer samples relevant for the sub-classes of the particular cancer of interest are chosen by the collaborating pathologists from hospital biobank repositories. The collaborator isolates the RNA and, subsequently, generates the RNAseq data with the use of an SOP. Qlucore specialists obtain access to the RNAseq data and optimize classification algorithms based on the most suitable statistical analyses and AI-based machine learning. The collaborators are also engaged in the verification and clinical validation of the developed classification model.

## Roll out plans for QI/QD

QI is commercially available for early access collaborating customers and partners. QD, initially with the classification model for childhood leukemia (ALL), is planned for commercial launch in 2023. Qlucore plans to continue the development of classification models for all cancers where sub-classifications aid the clinical assessment and treatment of patients. Not only transcriptomics but also other "omics" data may be used in the classification models (e.g., DNA methylation and proteomics) with these software tools. Moreover, the concept is not limited to cancer since classification of patients for personalized medicine using omics data is relevant for other types of diseases. Qlucore also foresee that the QD software is an ideal platform to stratify patients for companion diagnostics. Pharma companies will be approached as soon as the IVDR solution is available. JoPM



## Summary

- Qlucore is developing software solutions for clinical classification of cancers based on transcriptomics data. Two software packages, Qlucore Insights (early access collaborations) and Qlucore Diagnostics (2023 release), for RUO and CE-IVDR respectively, have been developed on the same technical platform. The two packages will be further developed in parallel for different purposes.
- The software is installed behind the firewalls of clinics and hospitals and used on-site.
- An SOP specifies how the RNA samples are to be handled, extracted, and prepared for RNAseq and analysis. Standard reagents for RNAseq are used, as are base-calling tools for sequence analysis.
- The software systems are built on a disease agnostic platform for which disease specific AI-powered statistical classification models are implemented after optimization and verification.
- The classification models are developed in collaboration with leading pathologists with deep biological knowledge and access to reference samples for a certain cancer type.
- A disease agnostic module for unbiased gene fusion analysis is included in the software systems. Reported fusions may be clinically annotated using the built in Mitelman fusion database.
- Customizable clinical reports can be generated automatically or after user interaction. The PDF reports may be implemented in the standard data handling system of a hospital (e.g., LIMS and data storage).
- Qlucore provides an easy to use, streamlined "data-to-report" workflow allowing personnel not trained in bioinformatics to perform the analysis.
- Data re-analysis and reporting can be performed using updated classifier models on any new or updated biomarkers and clinical guidelines.

## About Qlucore

Qlucore is a software company based in Lund (Sweden) with a US office in New York. Qlucore has been "bringing complexity into clarity" (company motto) by offering analysis solutions to the research market since 2007. Its software package, Qlucore Omics Explorer, has tools for many types of omics analysis for research applications. With strong competences in statistic analysis and result visualization, paired with a unique architecture allowing very fast computation, the company recently launched a software platform, Qlucore Insights (QI), dedicated for clinical data. In parallel the company develops the IVDR compliant Qlucore Diagnostics (QD) built on the experiences from QI. QD will be commercially available in 2023. The two software packages are built on the same technical platform and will be available for Research Use Only and Diagnostics, respectively.

## Erik Söderbäck

Erik has a PhD in plant physiology and microbiology. After a decade of molecular studies of bacterial nitrogen fixation, he joined the company Pyrosequencing and has spent the last 20 years on scientific and application support of NGS applications. Before recently joining Qlucore he spent the last 15 years in QIAGEN as specialist and market developer for the product portfolio of Clinical Genomics.

**References**

1. Kwa M., Makris A., and Gustavsen, and Esteva S.J. Clinical utility of gene-expression signatures in early-stage breast cancer. Nature Reviews, Clinical Oncology 2017(74)1-16
2. Lilljebjörn H., Orsmark-Pietras C., Mitelman F., Hagström-Andersson A., and Fioretos T. Transcriptomics paving the way for improved diagnostics and precision medicine of acute leukemia. J. Seminars in Cancer Biology, 2021(Oct 1)1044-579
3. Cieslik M., and Chinnaiyan A.M. Cancer transcriptome profiling at the juncture of clinical translation 2017(96)93-1094
4. Preprint: Baranovskii, A., Gunduz, I.B., Franke, V., Uyar, B., and Akalin, A. Multi-omics alleviates the limitations of panel-sequencing for cancer drug response prediction. bioRxiv preprint doi: https://doi.org/10.1101/2022.06.15.496249
5. Mateo J., Steuten L., Aftimos P., Andre F., Davies M., Garralda E., Giessler J., Husereau D., Martinez-Lopez I., Normanno N., Reis-Filho J.S.D., Stefani S., Thomas D.M., Westphalen C.B., and Voest E. Delivering precision oncology to patients with cancer. Nature Medicine 2022(28)658-665
6. Vaidehi Jobanputra V., Wrzeszczynski K. O., Buttner R., Caldas C., Cuppen E., Grimmond S., Haferlach T., Mullighan C., Schuh A., and Elemento O. Clinical interpretation of whole-genome and whole-transcriptome sequencing for precision oncology. J. Seminars in Cancer Biology 2021(7:3)1-9
7. Nikanjam M., Okamura R., Barkauskas D.A., and Kurzrock R. Targeting fusions for improved outcomes in oncology treatment. Cancer 2019(126)1315-1321
8. FDA Table of Pharmacogenomic Biomarkers in Drug Labeling: https://www.fda.gov/drugs/science-and-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling
9. Mertens F., Johansson B., Fioretos T., and Mitelman F. The emerging complexity of gene fusions in cancer. Nat Rev Cancer. 2015(15:6)371-81
10. Van der Vliet L.J, Dai H., van de Vijver M.J., He Y.D., Hart A.A.M., Mao M., Peterse H.L., van der Koy K., Marton M.J., Witteveen A.T., Schreiber G.J., Kerkhoven R.M., Roberts C., Linsley P.S., Bernards R., and Friend S.H. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002(415)530–536
11. Kimura S., and Mullighan C.G. Molecular markers in ALL: Clinical implications. Best Pract Res Clin Haematol. 2020 (33:3).
12. Mitelman F., Johansson B. and Mertens F. (Eds.) MitelmanDatabase of Chromosome Aberrations and Gene Fusions in Cancer (2020)., https://mitelmandatabase.isb-cgc.org