

# Bridging the gap: manage, analyze, visualize and interpret repertoire sequencing data using the IGX platform

Jorrit Posthuma, Lorenzo Fanchi, Henk-Jan van den Ham, Alvise Trevisan, Nicola Bonzanni

ENPICOM B.V., 's-Hertogenbosch, the Netherlands



## Immune repertoire sequencing

Lymphocytes form the core of the adaptive immune response. By expressing unique receptors that recognize specific antigens, B and T cells raise specific immune responses that are subsequently remembered by memory responses. The diversity of B and T lymphocytes within a host is called the immune repertoire and represents the total set of receptors that can recognize antigens. Since these receptors are enormously diverse, the nucleotide sequence of the antigen-recognizing part of these receptors can be used as a barcode to identify B and T cell clones (1, 2).

As a result of the high diversity of the immune repertoire, the low-throughput of Sanger sequencing provides only limited visualization of this variability. Next generation sequencing (NGS) platforms are ideally suited to extensively characterize and visualize the complexity and plasticity of the TCR and BCR repertoires (3, 4)



### Receptor identification

Obtain receptor sequences to characterize relevant cytotoxic T lymphocytes (CTLs)



### Antibody Discovery

Enable large-scale B-cell receptor repertoire profiling and the annotation of molecular-level immunogenetic data



### Vaccine Development

Inform the development of improved vaccines and provide surrogate markers to predict vaccine efficacy



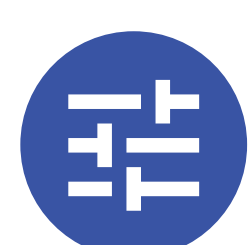
### Personalized repertoire monitoring

Quantify changes to an individual's repertoire as a result of disease, vaccination, or (immuno-)therapy treatment

## The ImmunoGenomiX platform

The ImmunoGenomiX (IGX) platform is an end-to-end immunosequencing data analysis platform designed to process, manage, analyze, visualize, and interpret immune repertoire data. Starting from the raw high-throughput sequencing data, the IGX platform delivers report- or publication-ready figures.

The IGX Platform is founded on innovative bioinformatics methods and has been coded from scratch using rock-solid software engineering. It has a modular structure which hosts multiple application for different repertoire sequencing analyses. Currently, two applications are available for the platform: IGX Profile, for clonotype analysis, and IGX Explore that consist of four different downstream analyses. Additional applications to answer immune repertoire-related questions will be added one-by-one.



### Flexible at the frontend

allowing customers to use their own sample preparation protocols and the Next-Generation Sequencing (NGS) technology of choice



### IGX Clone Collections

sequencing data, as well as metadata, can be used to perform powerful searches, select and filter clones that can be organized in clone collections for follow-up analysis



### IGX Tags

a powerful, tag-based, annotation system that allows users to extensively annotate repertoire sequencing data with structured metadata at a great level of granularity, including annotations for single clones such as affinity or avidity



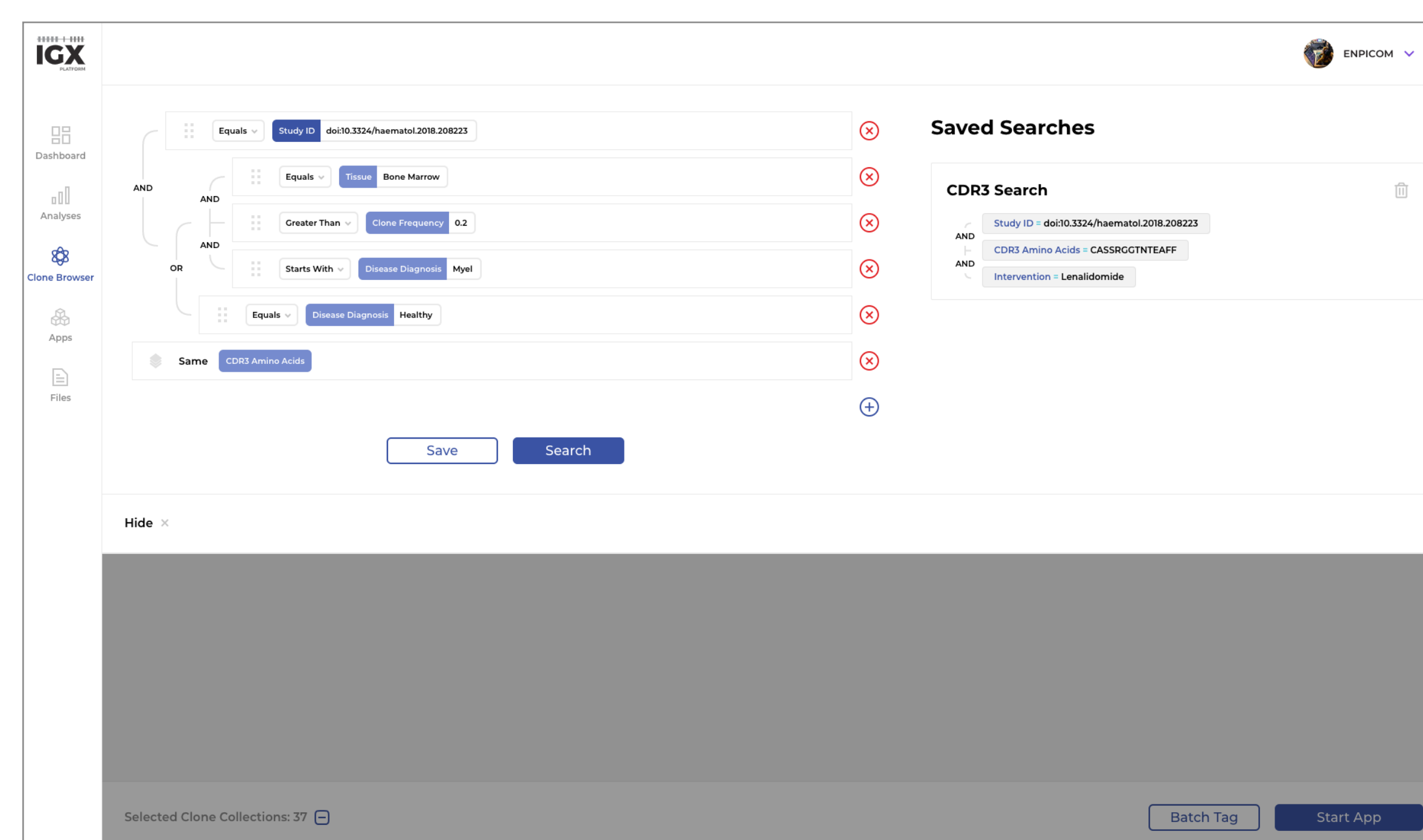
### Easy-to-use

no skilled bioinformatics staff required, the IGX platform runs in a secure cloud environment and features an intuitive and simple graphical user interface



## IGX Clone Browser

Finding and selecting relevant clones can be likened to searching for a needle in a haystack. As repertoire sequencing tends to generate vast amounts of data, it is of critical importance to have robust systems which enable the user to explore, prioritize, and select clones based on structured metadata and sequencing information. The IGX Clone Browser is a powerful tool to explore immune sequencing data and was designed from the ground up to provide a structured overview of all data, while at the same time allowing powerful and highly customizable searches.

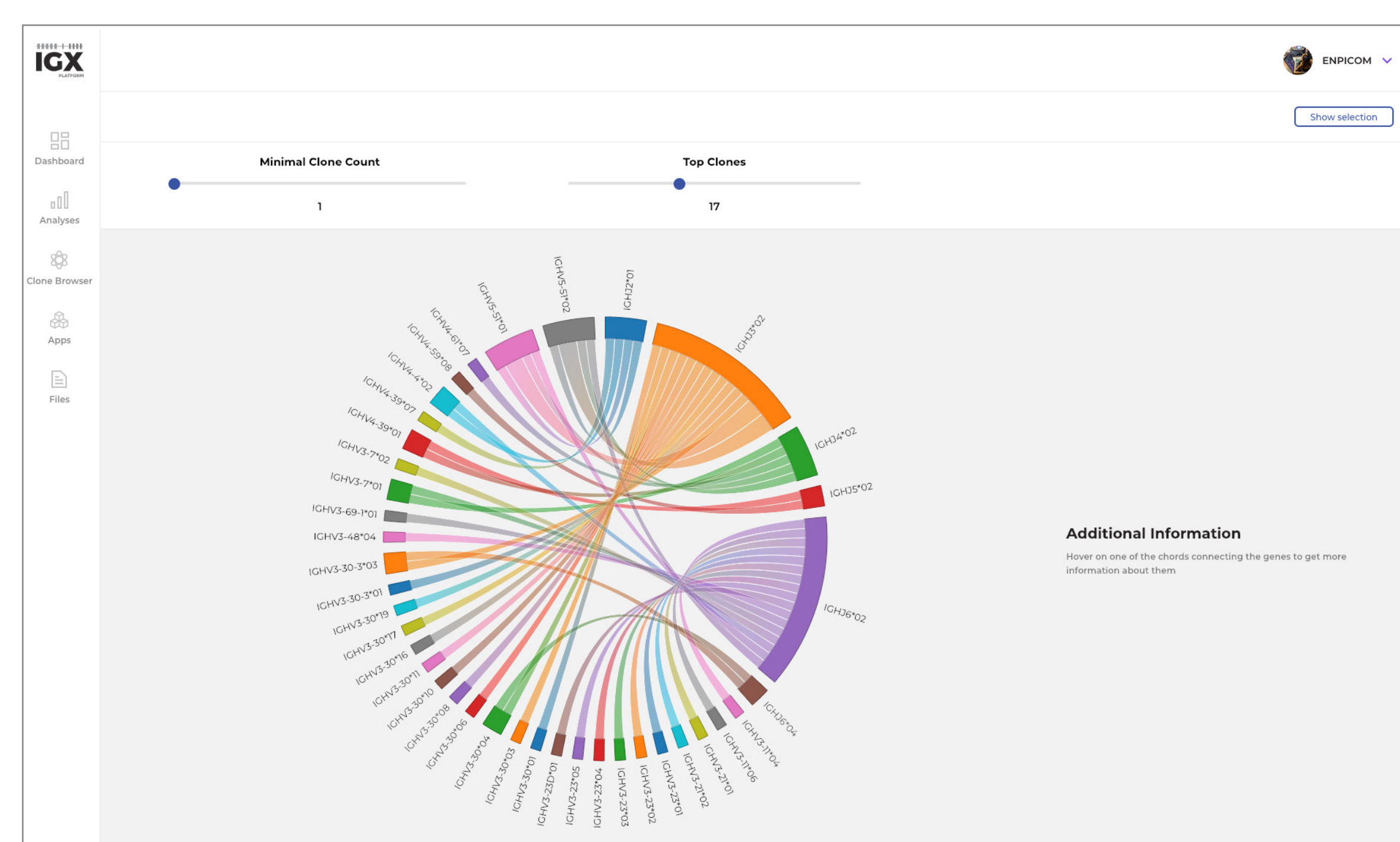


IGX Clone Browser filtering capabilities

## Technologies and frameworks

All individual component services, as well as analysis applications, are containerized and communicate using well-defined interfaces over HTTP. All communication to and from the platform is mediated by an API Gateway that also provides terminal TLS/SSL encryption, certified by an external, independent authority. Containerization enables strong versioning and reproducibility of the analysis results. All containerized services are orchestrated by Kubernetes. Computationally intensive parts of the analysis are run as Jobs (run-to-completion) and can use granular scaling up and scaling out of resources in order to optimize execution performance.

Both the front-end and back-end are coded in TypeScript, a modern, multi-paradigm language, which brings the benefits of static typing, including in particular more reliable run-time execution, to the power and versatility of JavaScript. On the front-end, TypeScript is coupled to the React UI library, the Redux state container and the Redux-Saga middleware for managing asynchronous actions. On the back-end, TypeScript runs on the Node.JS framework.



V/J gene usage plot generated by IGX Explore

## References

1. Liu et al, Cell 2017, 10.1016/j.cell.2017.01.014
2. Liu et al, Cell Biol Tox 2018, 10.1007/s10565-018-9426-0
3. Kurz et al, Blood 2015, 10.1182/blood-2015-03-635169
4. Roschewsky et al, Lan Oncol 2015, 10.1016/s1470-2045(15)70106-3