IGX PLATFORM 2.0 RELEASE NOTES

March 31, 2020

NEW FEATURE: Single-cell support

Single-cell sequencing is increasingly applied for the elucidation of paired chain information, essential for the characterization of therapeutic molecules. The analysis of single-cell receptor sequencing data comes with an additional layer of complexity. The IGX Platform 2.0 supports the storage, annotation, and analysis of paired receptor chains. Users can now easily structure paired receptor data and perform highly customizable searches to explore, prioritize, and select clones.

NEW APP: IGX-Compare

Internal and external databases of receptor sequences contain a lot of valuable information. Querying databases can be a difficult task due to their size and complexity. Also, making sense out of a query result is a time-consuming process. With IGX-Compare, users can compare any two data scopes and easily visualize and prioritize matching clones. IGX-Compare can be used by querying a single clone against entire databases to see if there is more information on the clone of interest. It can also be used to find the overlap between two patient cohorts.
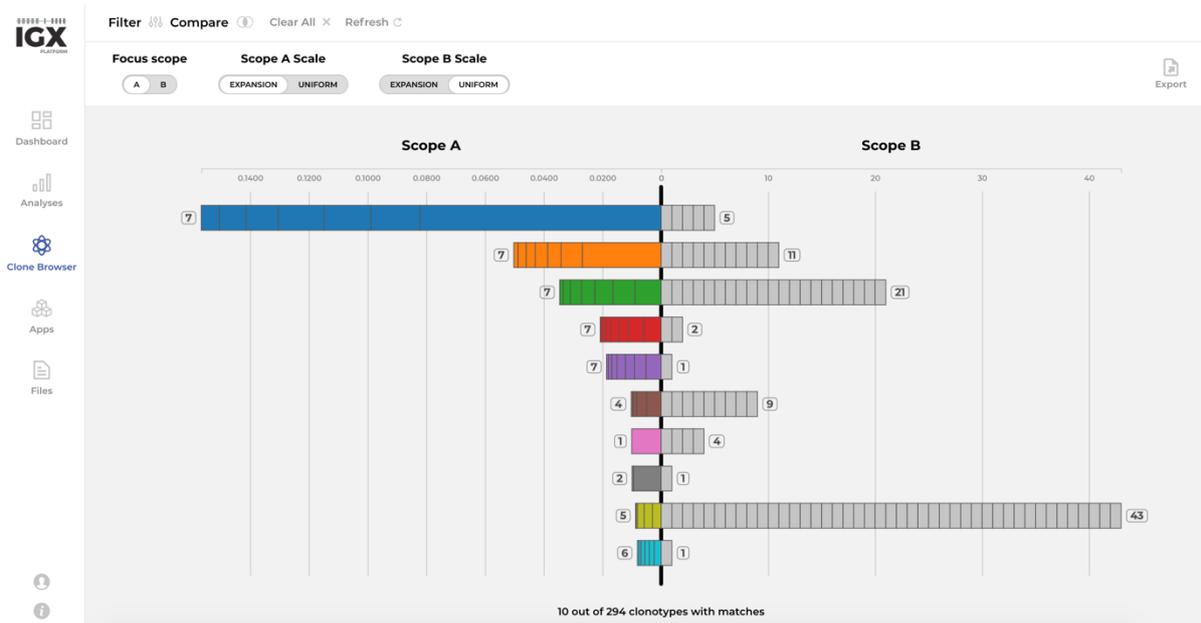

NEW FEATURE: Clonotype grouping

Research groups often have their own definition of a B or T cell clone, depending on the experimental setup or even the biological question that they aim to answer. The IGX Platform 2.0 supports various commonly used clonotype grouping definitions (e.g., CDR3 AA and CDR3 NT), making it compatible with a wide variety of use cases.

NEW ANALYSIS: V/J Gene Frequency plot

Gene segments play a major role in defining antigen recognition. IGX-Explore has a new analysis: the V/J Gene Frequency plot, an interactive visualization where users can uncover patient groups with a

preferential use of a particular gene, indicating the expansion of antigen-specific clones. Using the metadata annotations, users can group their samples and quickly compare gene usage across different patient cohorts.
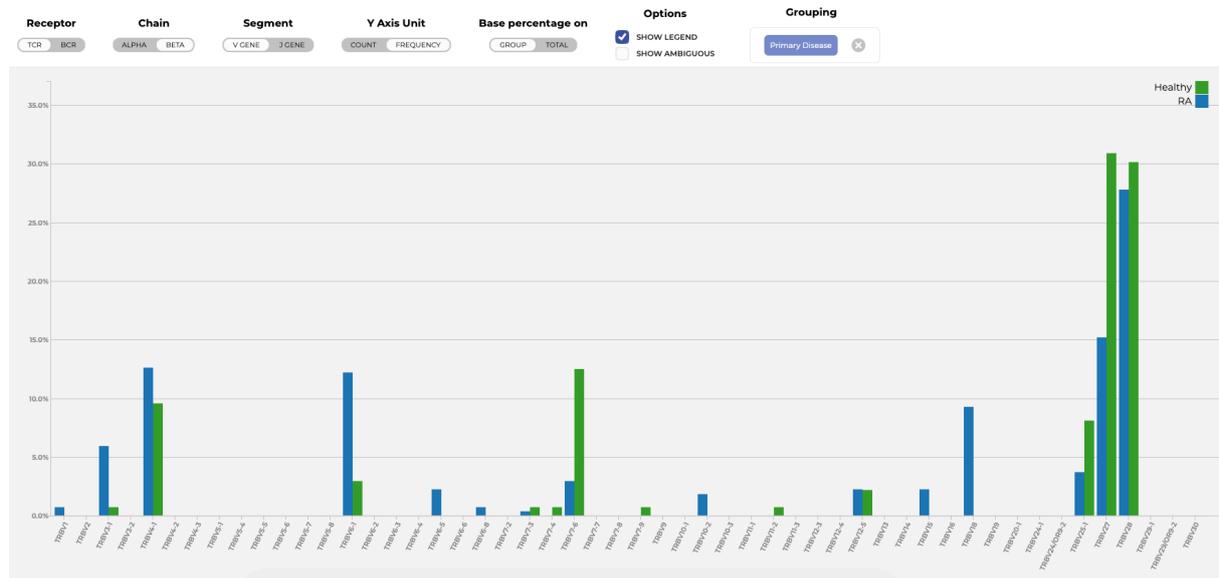
MORE EXTENSIVE DATA SUPPORT: Upload of readily annotated clones

We have extended integration capabilities allowing users to organize and analyze all their receptor sequences in a single environment. The IGX Platform 2.0 supports the upload of readily annotated clone tables (e.g., in CSV, TSV format) from common platforms such as Adaptive Biotechnologies and 10x Genomics. This way, users can continue analyzing their data through an intuitive user interface, integrate sequencing data from different technologies in one place, and compare their results with publicly available databases using IGX-Compare.

RICH METADATA UPLOAD

The metadata upload system features the following improvements:

- **Namespaces** – The search functionalities in the IGX Platform perform at its best when the 'type' of metadata tags is known (e.g. age is numeric, whereas disease diagnosis is comprised of character data). By creating separate namespaces, we offer organizations the possibility to structure their data even better and fully exploit the possibilities of the IGX data handling capabilities.
- **Excel file format support** – Metadata can now be attached to clone collections using common excel data formats. The first column of the file is used to specify name of the clone collection; any other column is then added as a tag to your data.
- **Batch** – A single metadata file can be used to annotate hundreds of samples in a matter of seconds.

BETTER PERFORMANCE: Scaling to hundreds of millions

The IGX Platform now supports the management and analysis of hundreds of millions of sequences, which is a must-have when providing a gateway to external Rep-Seq databases. We have made significant changes in the way we store and query data. More specifically, using smart partitioning and better indices the search space that is required for queries has decreased, speeding up the query process (e.g. for the identification of interesting clones).

NEW PAGE: Dashboard

Keeping track of the data that is being produced in different projects can be challenging. The new dashboard page in the IGX Platform provides a centralized overview of active analyses and summary statistics of the number of patients, sequences, or cells in each of the different studies, facilitating strategic discussions and reporting.