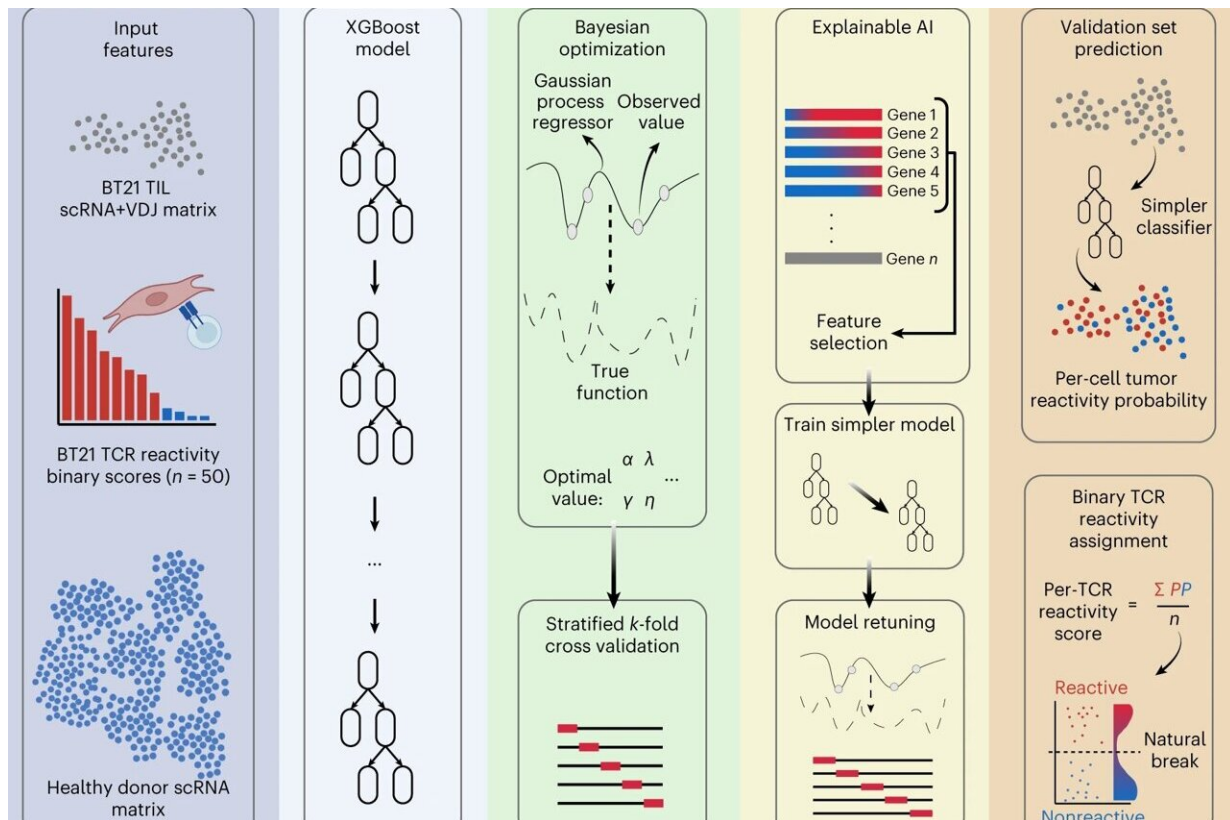# Machine learning classifier accelerates the development of cellular immunotherapies

March 15 2024

PredicTCR50 classifier training strategy. ScRNA data from healthy donors, as well as scRNA + VDJ and experimentally derived tumor-reactivity data for the 50 most frequent TCR TIL clonotypes from sample BT21, were used to train an intermediate model using XGBoost. Due to the sparse nature of scRNA data, we optimized this intermediate model by first performing Bayesian optimization to tune hyperparameters with stratified k-fold cross-validation. Subsequently we identified the top features (that is, genes) in this intermediate model using explainable AI SHAP, and then trained a simpler model using only these features

to prevent overfitting to the training data. This simpler model was retuned as before and then applied to the remaining BT21 TIL data. Per-cell reactivity probabilities calculated by the classifier were averaged for each TCR clonotype, and the Fisher–Jenk natural break was used to determine the appropriate minimum threshold for calling TCRs as tumor reactive. Credit: *Nature Biotechnology* (2024). DOI: 10.1038/s41587-024-02161-y

Making a personalized T cell therapy for cancer patients currently takes at least six months; scientists at the German Cancer Research Center (DKFZ) and the University Medical Center Mannheim have shown that the laborious first step of identifying tumor-reactive T cell receptors for patients can be replaced with a machine learning classifier that halves this time.

Personalized cellular immunotherapies are considered promising new treatment options for various types of cancer. One of the [therapeutic approaches](link) currently being tested is so-called "T-cell receptor transgenic T-cells." The idea behind this is that immune T cells from a patient are equipped in the laboratory to recognize the patient's own unique tumor and then reinfused in large numbers to effectively kill the [tumor cells](link).

The development of such therapies is a complicated process. First, doctors isolate tumor-infiltrating T cells (TILs) from a sample of the patient's tumor tissue. This cell population is then searched for T-cell receptors that recognize tumor-specific mutations and can thus kill tumor cells.

This search is laborious and has so far required knowledge of the tumor-specific mutations that lead to protein changes that are recognized by the patient's immune system. During this time, the tumor is constantly mutating and spreading, making this step a race against time.

"Finding the right T cell receptors is like looking for a needle in a haystack, costly and time-consuming," says Michael Platten, Head of Department at the DKFZ and Director of the Department of Neurology at the University Medical Center Mannheim. "With a method that allows us to identify tumor-reactive T-cell receptors independently of knowledge of the respective tumor epitopes, the process could be considerably simplified and accelerated."

A team led by Platten and co-study head Ed Green has now presented a new technology that can achieve precisely this goal in a recent publication. As a starting point, the researchers isolated TILs from a melanoma patient's brain metastasis and performed single-cell sequencing to characterize each cell. The T cell receptors expressed by these TILs were then individually tested in the lab to identify those that were recognized and killed patient tumor cells.

The researchers then combined these data with training a machine learning model to predict tumor-reactive T cell receptors. The resulting classifier could identify tumor-reactive T cells from TILs with 90% accuracy, work in many different types of tumor, and accommodate data from different cell sequencing technologies.

"PredicTCR enables us to cut the time it takes to identify personalized tumor-reactive T cell receptors from over three months to a matter of days, regardless of tumor type," said Ed Green.

"We are now focusing on bringing this technology into clinical practice here in Germany. To finance further development, we have founded the biotech start-up Tcelltech," adds Michael Platten. "PredicTCR is one of the key technologies of this new DKFZ spin-off."

The research is published in the journal *Nature Biotechnology*.

More information: C. L. Tan et al, Prediction of tumor-reactive T cell receptors from scRNA-seq data for personalized T cell therapy, *Nature Biotechnology* (2024). DOI: 10.1038/s41587-024-02161-y

Provided by German Cancer Research Center