Qeios

Peer Review

Review of: "CAFE: An Integrated Web App for High-Dimensional Analysis and Visualisation in Spectral Flow Cytometry"

Simone Lucchesi[1]

1. University of Siena, Italy

The article introduces CAFE, a web app designed for the analysis and visualization of high-dimensional spectral flow cytometry data. The tool implements a pipeline that includes preprocessing steps, data visualization, and clustering based on the Leiden algorithm. Overall, the work is well-presented, highlighting CAFE's potential as a user-friendly tool for data analysis. However, I have some concerns and suggestions for improvement that are reported here.

The use of density-based downsampling prior to batch correction raises concerns. Could this step impact the effectiveness of batch correction, as data reduction might alter the distribution of cell populations? Additionally, is ComBat the most appropriate choice for single-cell batch correction? To address these questions, it would be helpful to visualize the data for each marker before and after batch correction. For instance, in Figure 3A, the purple region (Tem CD8?) in the lower part of the plot shows a different shape between the healthy and COVID groups. Could this difference be attributed to a batch effect artifact?

The preprocessing pipeline includes an automatic noise exclusion step. Could this step exclude rare and biologically relevant cell populations? To evaluate this, testing the tool on datasets containing rare populations would be useful. This would help assess the tool's performance across diverse conditions and ensure its robustness.

The authors report that the Leiden algorithm outperforms others, such as PhenoGraph, FlowSOM, and SPADE. A comparison of CAFE with state-of-the-art algorithms demonstrating its higher performance would be a strong selling point for potential users, especially considering the already

 doi.org/10.32388/H8DC3J

high number of algorithms available for analyzing flow cytometry data. Additionally, since manual gating remains the gold standard in cytometry analysis, a comparison with manual gating results would further validate CAFE's performance and reliability.

Minor suggestions:

- If not already implemented, adding 2D plots to visualize marker distribution across clusters would provide users with greater control and insight into their data.

- Providing more detailed information on parameter settings and their impact on the analysis would enhance user understanding and facilitate more effective use of the tool.

- The installation process might be complex for some users; simplifying it could increase the potential adoption of CAFE.

In conclusion, the article is well-written and clearly presents the CAFE workflow, though some additions are needed to fully demonstrate the tool's performance and clarify preprocessing steps. The user-friendly interface is a welcome addition for cytometrists, even if the installation process appears to be complex and could represent a barrier for many potential users. Addressing these concerns would strengthen the article and improve the accessibility of CAFE.

Declarations

Potential competing interests: No potential competing interests to declare.