
Towards a Common Transcriptional Signature of Vaccination : Robustness of a Comparative Study of 13 Vaccines

Arthur Hughes*^{†1,2}, Rodolphe Thiébaud^{1,2,3}, and Boris P. Hejblum^{1,2}

¹Univ. Bordeaux, INSERM, Inria, BPH, U1219, F-33000 Bordeaux, France – Université de Bordeaux (Bordeaux, France) – France

²Vaccine Research Institute [Créteil, France] – Université Paris-Est Créteil Val-de-Marne - Paris 12 – France

³CHU de Bordeaux, Service d'Information Médicale, INSERM, U1219, F-33000 Bordeaux, France – Université de Bordeaux (Bordeaux, France) – France

Abstract

Vaccination is the most cost-effective intervention in public health for reducing global morbidity and mortality, yet most vaccines have been developed empirically, resulting in limited understanding of the mechanisms which characterise effective responses. High-throughput technologies such as microarrays or bulk RNA-sequencing enable the investigation of transcriptional signatures of vaccination by simultaneously measuring thousands of molecules within biological samples. These data are commonly analysed using differential gene expression tools, or with supervised learning approaches to identify biomarkers predictive of downstream immune responses. A central goal of systems vaccinology is to identify common signatures which are shared across vaccines, which could then be targeted when developing novel vaccines. In order to compare these responses across multiple vaccines, the Human Immune Project Consortium has compiled an integrated public database of existing vaccine experiments including gene expression transcriptomic measurements and immune response profiling (Diray-Arce, 2022). Hagan et al. (2022) previously analysed these data, deriving a time-adjusted common signature characterised by the expression of plasma cell-related genes. However, the high dimensionality of these data and the multiplicity of valid analytical choices make such studies vulnerable to issues associated with the replicability crisis in science. Given the importance of these findings for systems vaccinology, we reassessed their robustness by adopting different analytical decisions. Using alternative and more principled differential gene expression analysis methods, we obtained broadly similar results to Hagan et al., but found that their interpretation depends strongly on arbitrary hyper-parameter choices. In contrast, supervised learning approaches yielded divergent results, from which we instead propose best-practice recommendations for robust biomarker discovery in vaccine research.

Keywords: transcriptomics, vaccine, systems biology, machine learning, precision medicine, geneset

*Speaker

†Corresponding author: arthur.hughes@u-bordeaux.fr