Propuesta de Trabajo Fin de Máster
Año académico 2020-2021
Máster en Métodos Computacionales en Ciencias

Proyecto Nº 15

Título: *Bioinformatic identification and prioritization of neoantigens to develop personalized immunotherapies*

Departamento/ Laboratorio *Departamento de Inmunología e Inmunoterapia*

Director 1 *Sandra Hervas*
Contacto: *mshervas@unav.es*
Codirector: *Ibon Tamayo*
Contacto*: itamayou@unav.es*

Resumen

Tumor immunotherapy has emerged as a new approach for eliminating malignant tumors. Recently it has been shown that tumor mutational burden (TMB), defined by non-synonymous single amino acid mutations, correlates with clinical response to immunotherapy. Tumor-specific neoantigens (NeoAg) derived from these somatic mutations are thought to be the target of antitumor T cell responses mediated by immunotherapy. However, TMB has been a better predictor of response than NeoAg burden. With the advent of Next Generation Sequencing (NGS), researchers can now rapidly sequence a patient's DNA and RNA and analyze these sequencing data to predict NeoAgs computationally. This process requires several steps, each involving the use of bioinformatics tools and complex analytical pipelines. Matched tumor–normal DNA sequencing data are processed and analyzed to call somatic mutations of various types. RNA sequencing (RNA-seq) data are used to quantify gene and transcript expression, and verify mutation expression. Bioinformatics algorithms have made considerable progress and researchers have been able to identify relevant NeoAgs for an antitumor response. However, there is no current consensus approach, and there are few established best practices (1). On the other hand, most of the pipelines to select the best candidate NeoAg are based on the affinity of the mutant sequence to the patient's Human leukocyte antigen (HLA). However, several studies show that NeoAgs predicted based on their affinity to patient's HLA molecules do not predict response to immunotherapy. Other properties that predict peptide immunogenicity, and thus NeoAg quality, have recently emerged. Among this quality metrics are (i) the differential agretopicity index (DAI), defined as the ratio of HLA affinity of the mutant peptide and the wild type peptide, (ii) the similarity of the mutant sequence to known immunogenic peptides and (iii) the dissimilarity of the mutated sequence to the non-mutated (reference) proteome (2). The design of bioinformatic pipelines that integrates multiple NeoAg quality criteria may improve the prediction of clinical response to immunotherapy treatments as well as the selection of highly immunogenic NeoAgs for personalized treatments, such as vaccines. In this study, we propose to design different bioinformatic pipelines that include different quality metrics in order to see which of them provide the NeoAg burden that best predicts response to immunotherapy. To validate our results, we will use a cohort of patients from the TCGA who have been classified as responders or non-responders to immunotherapies treatments as well as an own cohort of patients with hepatocarcinoma.

It is not necessary for the student to know about advanced statistics. The student will have the opportunity to learn programming in R, working with the cluster and several algorithms ready to use in the cluster.

(1) Richters MM eta al. Best practices for bioinformatic characterization of neoantigens for clinical utility. Genome Med. 2019 Aug 28;11(1):56.

(2) *Richman LP et al. Neoantigen Dissimilarity to the Self-Proteome Predicts Immunogenicity and Response to Immune Checkpoint Blockade. Cell Syst. 2019 Oct 23;9(4):375-382.*