Proyecto PID2021-122726NB-I00 financiado por MCIN/AEI/10.13039/501100011033/ y por FEDER Una manera de hacer Europa

Identificación del proyecto:

Evolución de péptidos no canónicos y su papel como neoantígenos en cáncer

Descripción del proyecto:

In recent years we have learnt that a much larger number of non-canonical open reading frames (ORFs) than previously anticipated are translated. These ORFs are located upstream or downstream of canonical coding sequences, in alternative coding frames, or in transcripts annotated as long non-coding RNAs (lncRNAs). The encoded protein products tend to be shorter than 100 amino acids and are termed non-canonical peptides. In humans alone, 7,264 such non-canonical translation events have been described, and the number is likely to increase substantially in the coming years. In contrast to canonical coding sequences, non-canonical translated ORFs are frequently species- or lineage-specific. This is because the transcripts themselves show limited phylogenetic conservation, as is the case of the vast majority of annotated lncRNAs, or because the ORFs have only recently formed in already existing transcripts. This highlights frequently occurring but poorly understood de novo emergence processes, whereby coding sequences originate from previously non-coding ones.

Here we propose to perform innovative research to shed new light into the evolution and function of different types of non-canonical ORFs and their encoded products. The first aim is to characterize the translation activities upstream and downstream of canonical ORFs using a set of closely related yeast species as a model system. We will use a novel approach in which we will combine Nanopore RNA-Seq long reads, which enable the recovery of 5 and 3 untranslated regions (UTRs), and ribosome profiling data. We expect to obtain the first complete catalog of non-canonical peptides in Saccharomyces cerevisiae and, by comparing to similar experimental data from other species, together with information from genomic syntenic regions, reconstruct the most common evolutionary paths leading to their emergence. The second aim is to characterize proteins generated de novo in the human lineage, focusing on those encoded by lncRNAs and novel transcripts. For this, we will take advantage of the increasing number of ribosome profiling experiments performed in human tissues and cell types, combined with extensive transcriptomics and proteomics data from multiple species. We expect this research to be a milestone in our understanding of de novo gene formation in humans. Finally, we will investigate the de novo formation of peptides in tumors and their potential to elicit immune responses. There is growing evidence that many lncRNAs and novel transcripts show tumor-specific expression, being absent from healthy adult tissues. This is an important source of neoantigens that remains poorly characterized. We will quantify the number of potentially immunogenic non-canonical peptides arising in

different cancer patient cohorts, and determine if it can help better predict a patients response to immunotherapy.

Financiación: AGENCIA ESTATAL DE INVESTIGACION

217.800,00€

Este proyecto está cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER). "Una manera de hacer Europa"