Abstract #33 | Poster

Optimising Base Editing to Identify Novel Translated Microproteins in B Cell Lymphoma.

E. G. Whittle[1], J. Krupka[1], G. Scevchenko[1], R. Jackson[1], D. J. Hodson[1]

[1] *Univeristy of Cambridge, Department of Haematology, Cambridge, United Kingdom*

The human genome contains ~20,000 annotated protein-coding genes, originally defined by an arbitrary size threshold of 100 amino acids. Ribosome profiling (Ribo-seq) has challenged this assumption by revealing widespread translation of thousands of unannotated, non-canonical, small open reading frames (smORFs), some encoding microproteins with potent biological functions. However, most remain uncharacterised at the functional level.

We performed Ribo-seq across >80 primary B-cell samples and lymphoma cell lines. Using an in-house smORF calling pipeline, we identified over 30,000 translated non-canonical ORFs arising from lncRNA and untranslated regions of coding RNAs. CRISPR perturbation of the top 1,000 smORFs revealed >100 with B-cell essential phenotypes. However, conventional CRISPR cannot distinguish whether phenotype arise from loss of translated microprotein or collateral damage to nearby structural or regulatory elements within host RNA. Base editing (BE) allows high-precision, high-throughput mutagenesis of start-codons, ablating smORF/microprotein translation with minimal collateral damage to surrounding RNA. BE application to smORF discovery has been limited by technical constraints and variable efficiency.

To overcome this, we developed a BE sensor screening strategy where each sgRNA is physically linked in cis to 45nt of its genomic target sequence, enabling direct quantification of editing by sequencing of the integrated cassette. A library composed of 404 sensor sgRNAs targeting 84 neutral genomic loci was screened in lymphoma cells expressing 12 different BE variants. SpG_ABE8e emerged as the most accurate and efficient editor across all tested variants. Sp_BE4max showed efficient editing but was limited by PAM availability. Systematic characterisation of BE parameters revealed a conserved 5-nt editing window, a minimum 7-day editing period and an unanticipated dependence of the 5’ flanking nucleotide for optimal efficiency. Deployment of a larger pilot library confirmed the precision and resolution of our optimised platform in established lymphoma cell models.

These methodological advances enabled us to construct the first genome-wide BE screen targeting the start codon of every potential smORF in the human genome. Ongoing work applies this platform in the lymphoid system to link non-canonical translation to function and identify microproteins that contribute to B-cell and B-cell lymphoma biology.