Agnostic mutual exclusivity analysis for synthetic lethal pair identification

Francesca Lorenzin[1], Davide Prandi[1], Tarcisio Fedrizzi[1], Alessio Locallo[1], Paola Gasperini[1], and Francesca Demichelis[1,2].

[1] Centre for Integrative Biology, University of Trento
[2] Englander Institute for Precision Medicine, New York-Presbyterian Hospital, Weill Cornell Medicine, New York

Prostate cancer is a clinically and genetically heterogeneous disease and despite the big advances over the years for the management of the disease, molecular targets and biomarkers still need to be identified for more effective patients' treatment and stratification. Large sequencing efforts (e.g. The Cancer Genome Atlas (TCGA), Stand Up to Cancer (SU2C), ICGC) provide information that can be exploited towards personalized therapy. Mutual exclusivity analysis of genomic lesions can be performed on these large collections of data to nominate potential synthetic lethal (SL) pairs that can serve as new targets for specific molecular subtypes. We embraced such large-scale search by first addressing technical issues that might confound signal of interest, including DNA purity, aneuploidy and allele-specific events, preventing the systematic investigation of the SL pairs. Here we present a two-step approach to perform large SL analyses and preliminary data on prostate cancer mutual exclusivity pairs to be validated *in vitro*. We developed the SPICE (Synthetic Lethal Phenotype Identification through Cancer Evolution analysis) pipeline to uniformly process thousands of samples that are then used by FaMe (Fast Mutual Exclusivity algorithm) to explore millions of aberrations for the nomination of mutual exclusive pairs.
The SPICE pipeline is very flexible in terms of data input (sequencing assay and design) and provides purity- and ploidy-adjusted somatic copy number alteration (SCNA), allele specific copy number, somatic nucleotide variant (SNV) and level of clonality per lesion. Afterwards, FaMe models each type of aberration as a binary matrix and implements fast matrix multiplication and massive parallelization to test for mutual exclusivity of hundreds of millions of pairs in few minutes. To shortlist nominated pair for *in vitro* validation filters based on expression data, interaction and drug target databases are used.
So far, we processed 6,289 TCGA tumor samples belonging to 26 tumor types. Specifically, for prostate cancer 496 samples from TCGA and 417 samples from SU2C were corrected for purity and ploidy and analyzed. We estimated SCNA of more than 37,253 genes across tumor types and observed that 50% of the tumors present a different copy number status in at least 5% of the genes. Moreover, the allele specific analysis uncovered frequent copy number neutral loss (CNNL) that ranges from <1% (LAML, PRAD, UVM, THCA) to >12% (LUAD, LUSC). Further details on the general features of these analyses and on newly identified mutually exclusive pairs in prostate cancer will be presented.

The authors declare no conflict of interest.

This work was supported by the European Research Council grant ERC-2014-CoG648670.