I Congresso Brasileiro de
Jovens Pesquisadores em
Matemática Pura e Aplicada

Statistical methods for preprocessing, integration, and quality assessment of synthetic lethality data

Samara Kiihl*

*UNICAMP, Campinas, Brasil

Resumo

Two genes are considered to be synthetically lethal (SL) when cells carrying a loss-offunction mutation in either of these genes are viable, but cells with loss-of-function mutation in both of the genes are not. In yeast, biologists have been observing the phenotype of mutant strains with two genes knocked out in search for SL, however, with over 12 million gene pairs to explore, only a small percentage of all possible interactions have been studied. We use data from the protocol called dSLAM (diploid-based synthetic lethality analysis by microarray) that measures growth of the mutants strains using microarrays. We describe some of the challenges associated with microarray data and propose a statistical model to improve prediction of SL pair. The model allows borrowing information across arrays and across genes to improve robustness and precision of the estimates. We use BioGRID, a curated database of genetic interactions, to demonstrate the advantages of our model by comparing it to naive approaches. Methods to speed up the search of SL pairs are crucial to identify all possible SL genes and may have an impact to overcome failure of cancer treatments targeting only one gene.