Hybrid precision medicine approach combining epigenetic biomarkers and machine learning to stratify chronic lymphocytic leukemia patients

Manuel Cossio[1], Marta Kulis[1] & Ignacio Martin Subero[1].

Patient stratification in cancer has always been a multidisciplinary challenge. In the beginning, the classification was based mainly on clinical findings and specific chromosomal aberration encountered in some types of neoplasias that were related to a particular clinical development. With the advances in sequencing technologies, a variety of target genes and recurrent mutations for each cancer type were also studied. However, only a reduced group of this type of feature was able to contribute to the prediction of the clinical outcome in certain neoplasms. In addition to genetic changes,, alterations in the epigenomic landscape have arisen as a crucial molecular trait of cancer cells with powerful links not only related to pathogenesis and evolution but also with clinical behavior of almost every tumor type. Thanks to epigenomic studies, the importance of DNA methylation was discovered and understood, mainly due to its reliability, tissue invariance and the possibility of tracking the cell of origin of the patient's neoplasia. Consequently, several efforts were directed to these particular marks to implement them as biomarkers to allow patient stratification in a very simple, robust and fast laboratory setting. In this abstract, we will present the validation of a previously-published strategy (Queiros et al., Leukemia 2015) that uses 5 epigenetic biomarkers and a machine-learning algorithm to classify chronic lymphocytic leukemia (CLL) patients into three groups with differential cell of origin and distinct clinical features This strategy based on epigenetics rather than genetics seems to improve the current classification of CLL patients based on the mutational status of the immunoglobulin genes.

Isolated DNA of a large cohort of patient samples was bisulfite converted and pyrosequenced using specific in-silico designed primers. Methylation levels for each of the 5 CpG regions were quantified and later the entire 224 CLL cases were classified with a supervised machine learning classifier (support vector machine), into three epigenetic CLL subgroups.

The method correctly separated the patients into three distinct groups, named as n-CLL (40%), i-CLL (16%) and m-CLL (44%). m-CLL and i-CLL were highly enriched for IGHV mutated cases whereas n-CLL with IGHV unmutated. i-CLLs presented a higher proportion of patients in Binet stage B than m-CLL. Time to treatment analysis showed a clear separation between the three groups, with m-CLL being of good prognosis, n-CLL of bad clinical behavior and an intermediate state in the case of i-CLLs.

Based on these results and recent data published in the literature, we conclude that using epigenetic biomarkers together with artificial intelligence for the stratification of patients in CLL seems to be an excellent clinical asset in the precision medicine era.

[1] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain.