BEAR: a novel virtual screening method based on Large-scale bioactivity data

Jaeok Lee[2], Jiyeon Kang[2,*], Hwa Jeong Lee[2], Wankyu Kim[1,3,*]

[1]*KaiPharm , Seoul, Republic of Korea*

[2]*College of Pharmacy, Ewha Womans University, Seoul 03760, Republic of Korea*

[3]*Department of Life Science, Ewha Womans University, Seoul, Republic of Korea*

● Data-driven drug discovery exploits a comprehensive set of big data to provide an efficient path for the development of new drugs. Currently, publicly available bioassay data sets provide extensive information regarding the bioactivity profiles of millions of compounds. Using these large-scale drug screening data sets, we developed a novel in silico method to virtually screen hit compounds against protein targets, named BEAR (Bioactive compound Enrichment by Assay Repositioning).

● Assay Repositioning : The key idea of BEAR is to reuse bioassay data for predicting hit compounds for targets other than their originally intended purposes, i.e. Assay Repositioning

● The BEAR approach differs from conventional virtual screening methods :

1) it relies solely on bioactivity data and requires no physicochemical features of either the target or ligand.

2) Accordingly, structurally diverse candidates are predicted, allowing for scaffold hopping.

● BEAR shows stable performance across diverse target classes (mean AUC = 0.87 for ~1,000 targets), outperforming conventional machine learning methods.

● BEAR predicted dual inhibitors of P-gp & BCRP (ABC transporters) at a high hit rate (9 hits / 72 tested)