

Krier M., et.al. *Assessing the Scaffold Diversity of Screening Libraries. J. Chem. Inf. Model.*, 2006, 46 (2), pp 512–524

Abstract:

Medicinal chemists have traditionally realized assessments of chemical diversity and subsequent compound acquisition, although a recent study suggests that experts are usually inconsistent in reviewing large data sets. To analyze the scaffold diversity of commercially available screening collections, we have developed a general workflow aimed at (1) identifying druglike compounds, (2) clustering them by maximum common substructures (scaffolds), (3) measuring the scaffold diversity encoded by each screening collection independently of its size, and finally (4) merging all common substructures in a nonredundant scaffold library that can easily be browsed by structural and topological queries. Starting from 2.4 million compounds out of 12 commercial sources, four categories of libraries could be identified: large- and medium-sized combinatorial libraries (low scaffold diversity), diverse libraries (medium diversity, medium size), and highly diverse libraries (high diversity, low size). The chemical space covered by the scaffold library can be searched to prioritize scaffold-focused libraries.

The authors give in-depth analysis of very few vendors' libraries including TimTec natural products and synthetic (general stock) compounds. The article discussed the use of computational/analysis tools in assessing scaffold diversity and provides good visuals to draw conclusions about available screening collections on the market.

TimTec library collections are assigned into two major categories: "diverse libraries (medium diversity, medium size)", and "highly diverse libraries (high diversity, low size)". TimTec diverse libraries have "more original and less-populated scaffolds". As a group, diverse libraries and highly diverse libraries "are really diverse in terms of scaffold architecture and generally present a larger choice of proprietary low-populated scaffolds".

High diversity result can be explained by the fact that TimTec does not use combinatorial chemistry approach and its molecules are hand synthesized. In addition, there is balanced combination of in-house synthesis and outside sourcing. Some overlap with other vendors turns out to be beneficial for one-point access to good diversity and for compounds re-supply.

Please contact us for more information about individual libraries scaffolds, fragments representation and diversity score.