



The World Diversity Set 3 is a pre-plated library of 10,000 diverse screening compounds which makes full use of Specs' global acquisition program.

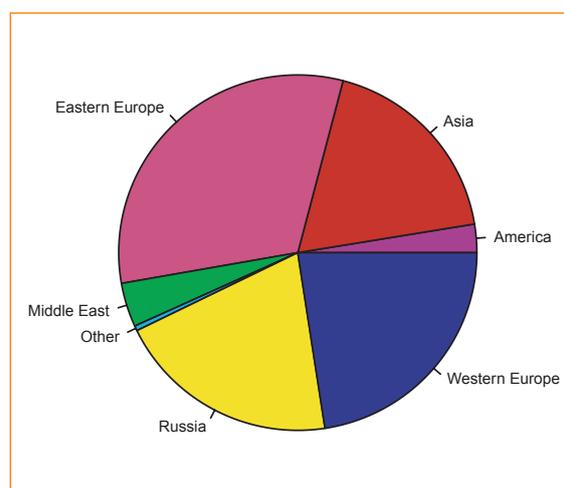
A total of 454 different producers from 47 countries worldwide have contributed to this diverse and unique set of compounds. In order to specifically maximize the geographical diversity of the library we have clustered the compounds using various chemical descriptors and favored producers from countries that have no predominant history as suppliers of screening compounds.

Selection method

In order to establish a drug-like library compounds were selected using structural filter alerts based on chemical properties and our stringent Specs screening guidelines. The following criteria were applied:

- Molecular weight < 500 Da
- Rotational bonds ≤ 10
- H-bond donors ≤ 5
- H-bond acceptors ≤ 10
- tPSA $\leq 140 \text{ \AA}^2$
- $\log P < 5$
- $\log S \geq -7$

For the filtered set Kier and Hall topological (valence) connectivity indices were calculated and a principle component analysis was applied. K-means clustering was used as well on 8 of the most relevant components which resulted in nicely spread clusters. An algorithm was applied on every single cluster with a bias for producers from countries that have no predominant history as suppliers of screening compounds. This resulted in a globally diverse library of compounds whose origin is shown in the geographical distribution chart below.



The diversity of the library was demonstrated by the generated Bemis-Murcko fragments and the frequency in which every unique framework occurred. Out of a total of 5,267 frameworks, 3,970 were unique. Furthermore 5,064 scaffolds occur 5 times or less and only 6 occur more than 50 times but these are

obvious structures like single six-membered rings and six-membered rings tethered by small carbon or hetero atom linkers. The various distribution charts are given below and are indicative of the diversity of this unique set.

