



Consensus Diversity Plots: A Complete Diversity Analysis of Chemical Libraries for Drug Discovery

Mariana González-Medina,¹ Fernando D. Prieto-Martínez,¹ John R. Owen,² José L. Medina-Franco¹

¹ Facultad de Química, Departamento de Farmacia, Universidad Nacional Autónoma de México (UNAM), Avenida Universidad 3000, Mexico City 04510, Mexico.

² High-Performance Computing Research Group, ECIT Institute, Northern Ireland Science Park, Queens Road, BT3 9DT Belfast, United Kingdom.

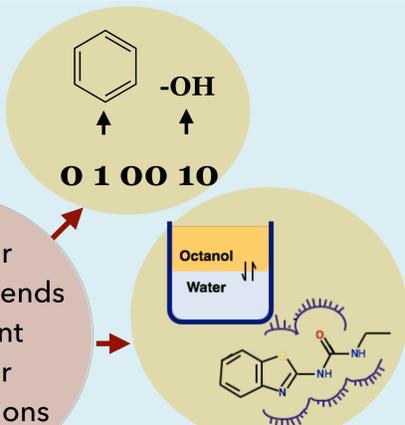
mgm_14392@comunidad.unam.mx, medinajl@unam.mx



Molecular diversity

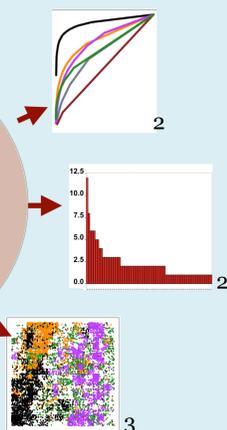
Measuring the diversity of compound data sets is relevant for drug discovery and many other areas of chemistry

Molecular diversity depends on different molecular representations



Molecule
Scaffold

Cheminformatic analyses of diversity are performed using multiple criteria



The assessment with each method is analyzed independently and is not straightforward to provide an evaluation of the "global diversity"

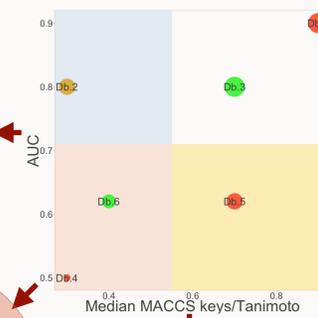
Goal

Visualize in few dimensions the diversity of several chemical libraries considering simultaneously multiple molecular representations

About the CDPs

Scaffold diversity plotted along the y axis

Each data point represents a data set and its size



Fingerprint diversity plotted along the x axis

It can be divided into four quadrants that classify data sets as high/low diverse considering both, fingerprints and scaffolds

Additional diversity criterion can be mapped with a categorical color

Molecular properties or complexity

Results

Consensus Diversity Plots Version 1 (CDPs - V.1)
Shiny app

Consensus Diversity Plots Version 1 (CDPs-V.1)

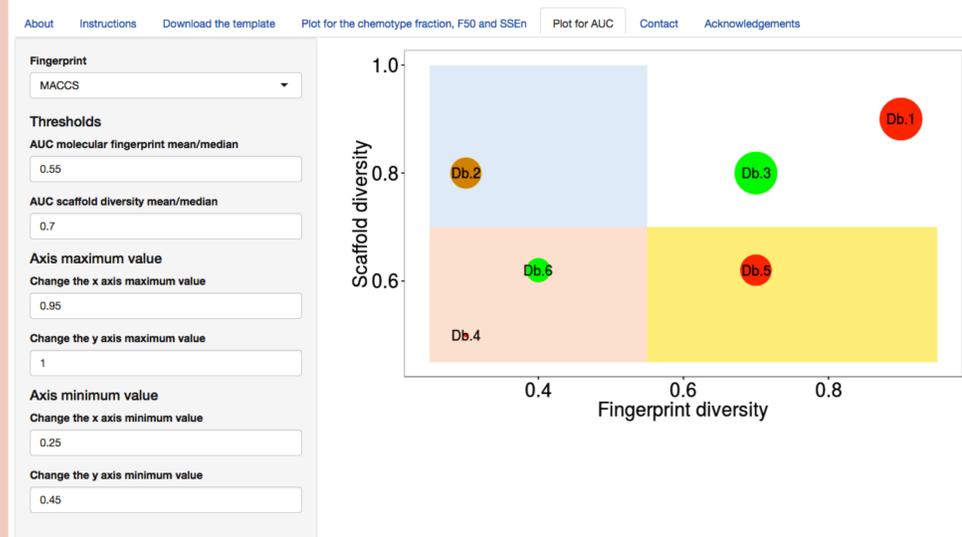


Figure 1. Example of a CDP produced with the online version of the CDPs

Figure 1 depicts an example of a CDP produced with the online version of the CDPs, freely available at:



Conclusion

CDPs are general data mining tools that represent in 2D multiple metrics to measure the diversity of compound data sets with different size and composition

Perspectives

Include the possibility to automatically calculate some diversity metrics

Add complexity as a property to be depicted by the color scale

We want to keep improving this app! If you have suggestions or if you experience an error, please visit our website and send us an email

References

[1] Xu Y-J, et al. J Chem Inf Comput Sci. 2002, 42(4): 912-926

[2] Medina-Franco JL, et al. QSAR Comb Sci. 2009, 28(11-12): 1551-1560

[3] Owen JR, et al. J Chem Inf Model. 2011, 51(7):1552-1563

[4] González-Medina M, et al. J Cheminform. 2016, 8, 63

Acknowledgments

UNAM: PAPIPE PE200116 and PAIP 5000-9163