

DreaMS: a universal large language model for interpreting tandem mass spectra

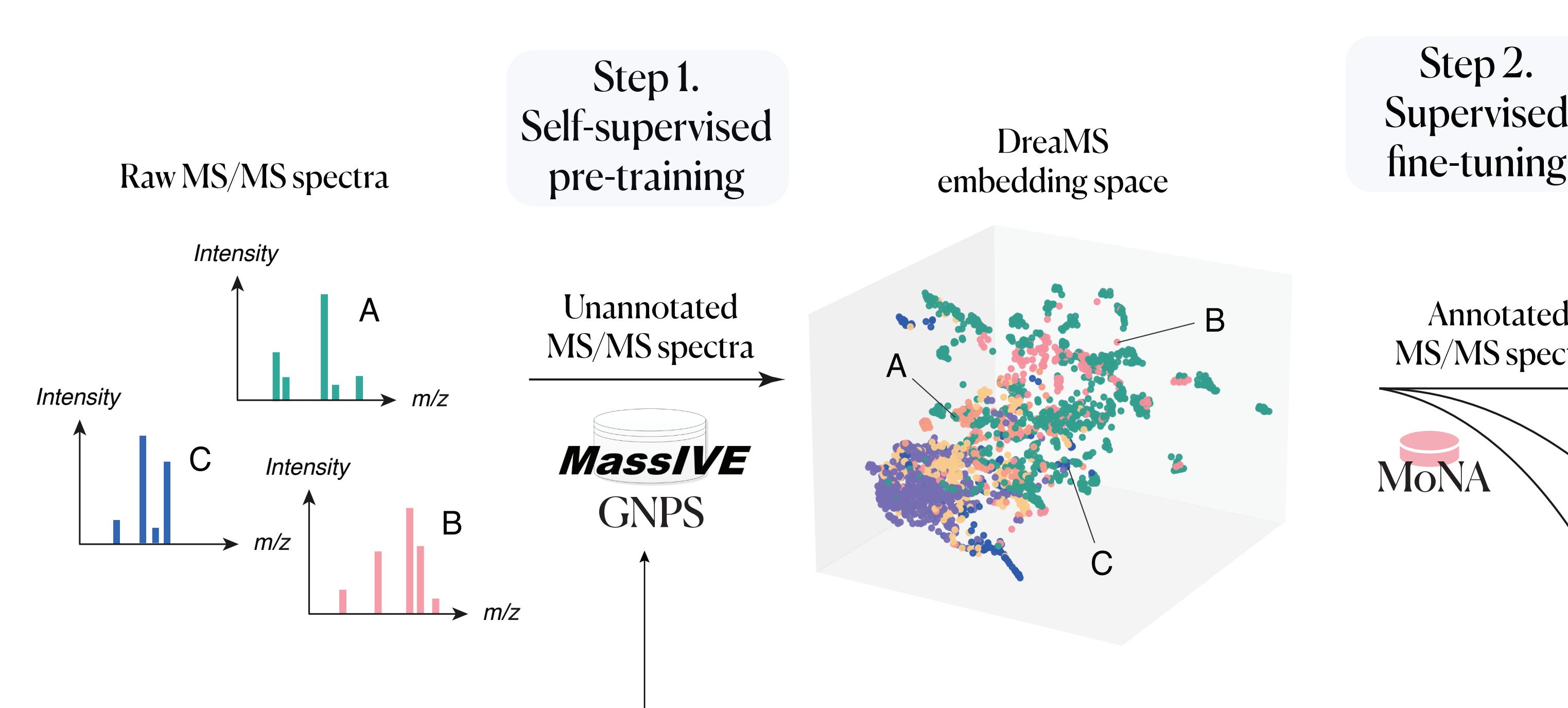
Roman Bushuiev^{1,2,*}, Anton Bushuiev^{2,*}, Raman Samusevich^{1,2}, Corinna Brungs¹, Josef Sivic², Tomáš Pluskal¹



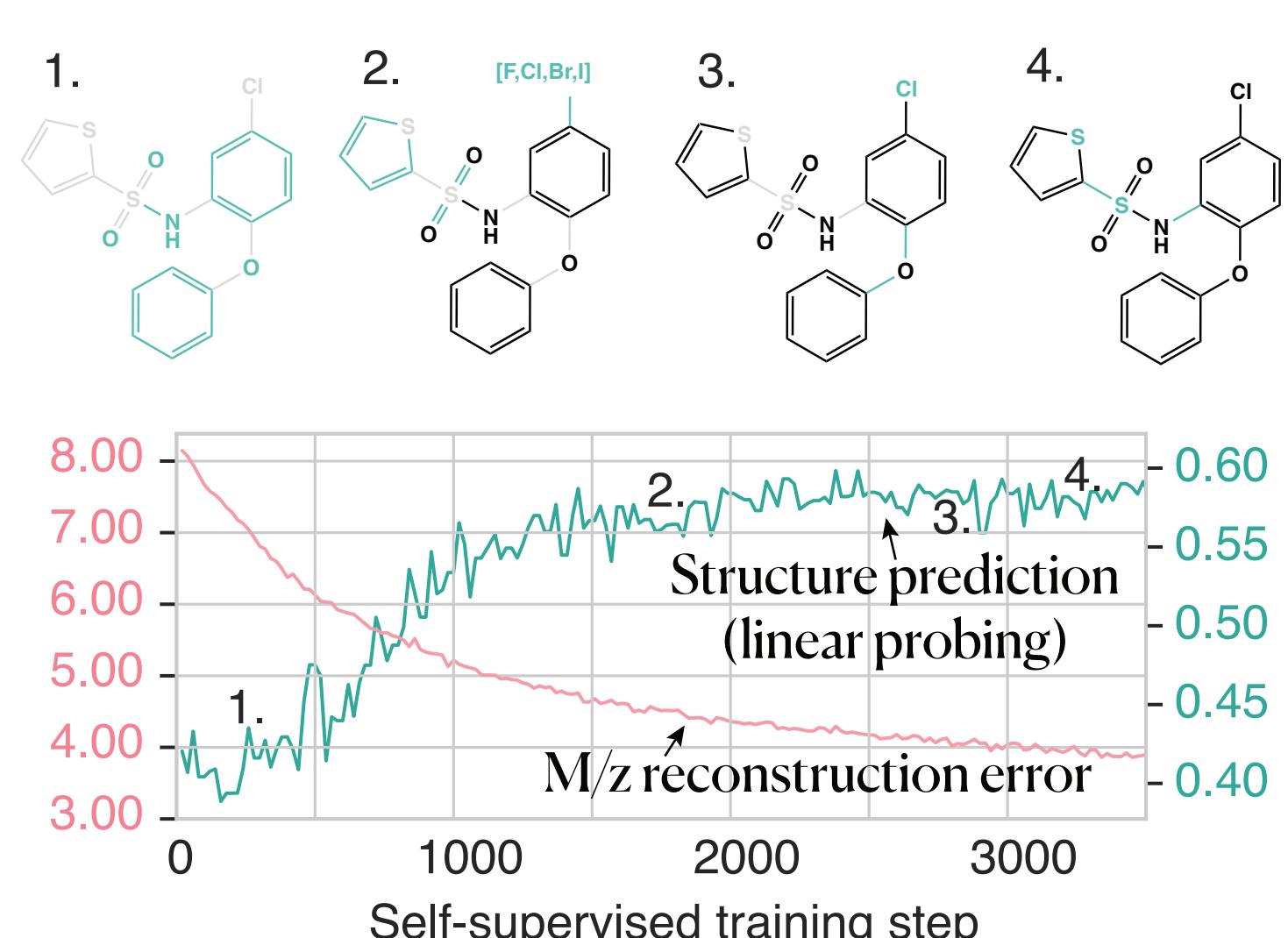
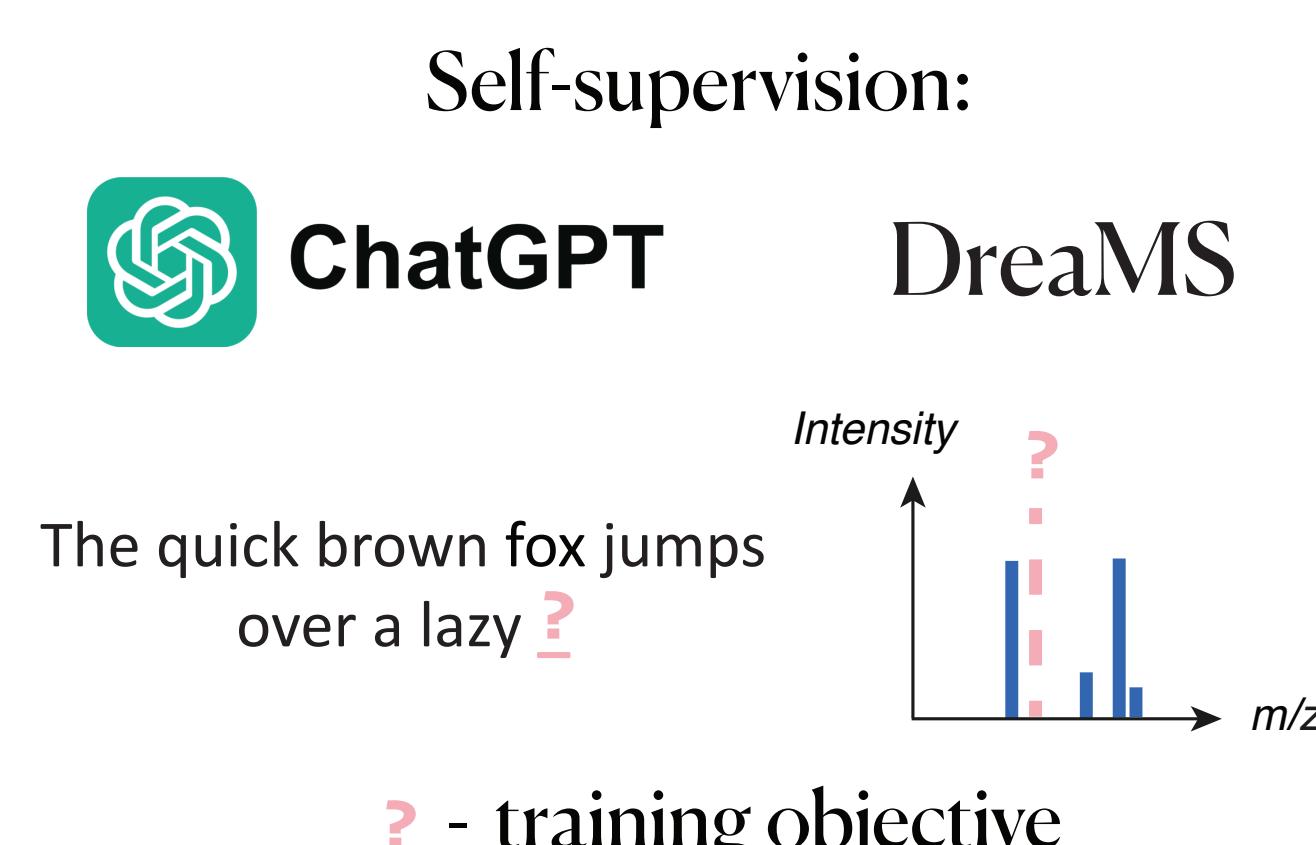
Paper

GitHub

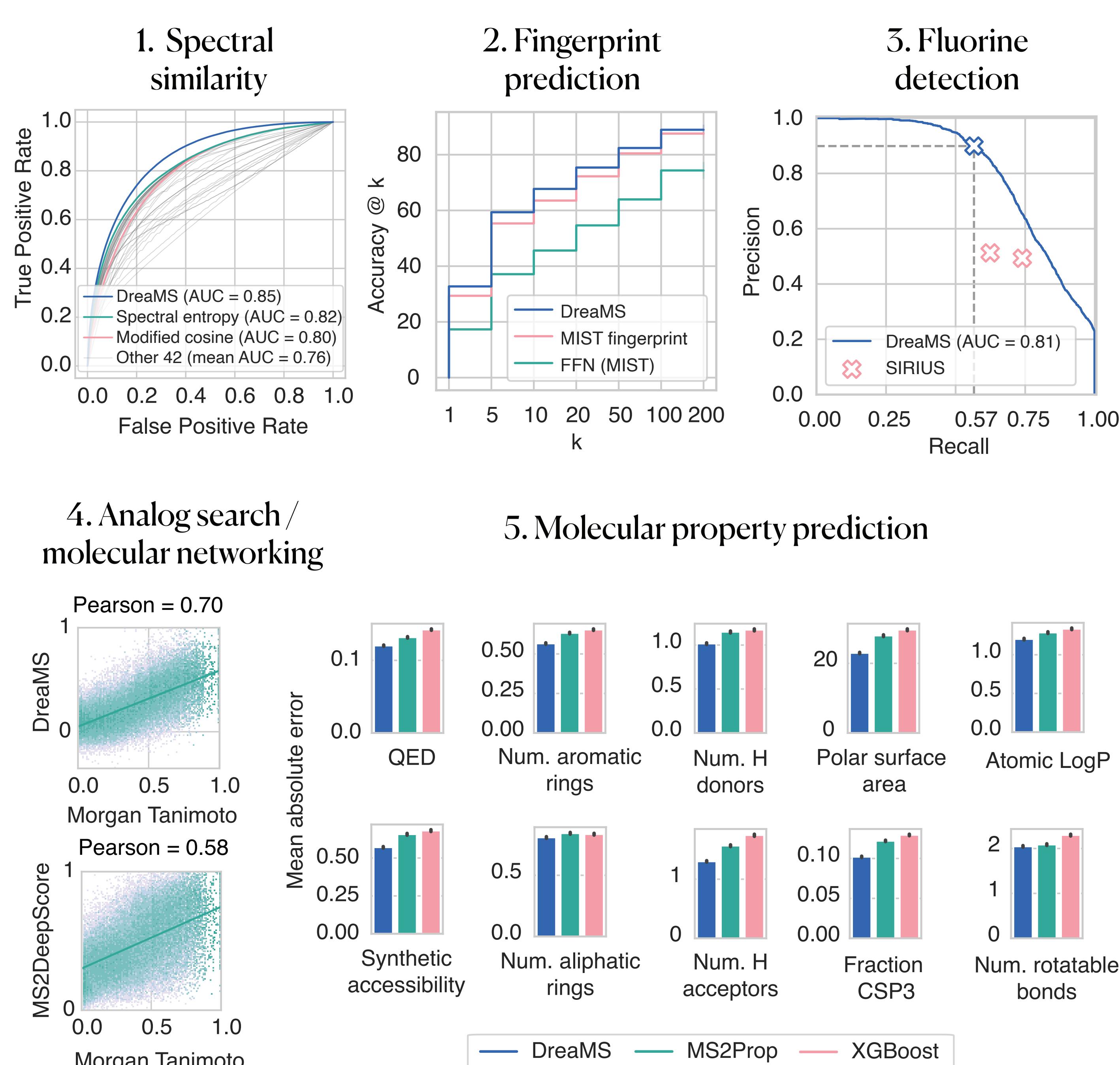
Method overview



Molecular structures emerge when DreaMS learns to reconstruct masked m/z values

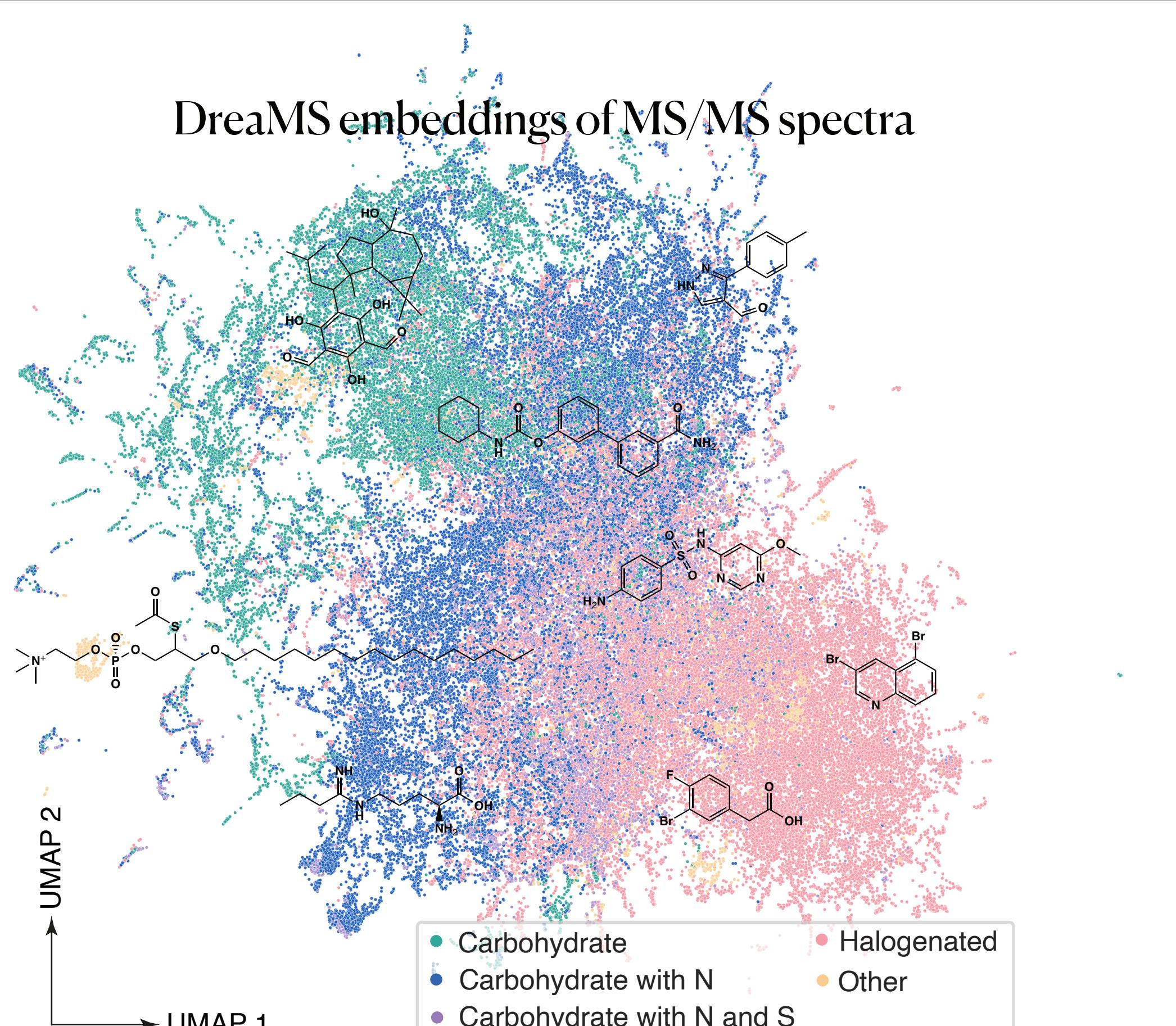


State-of-the-art method for:



DreaMS embeddings

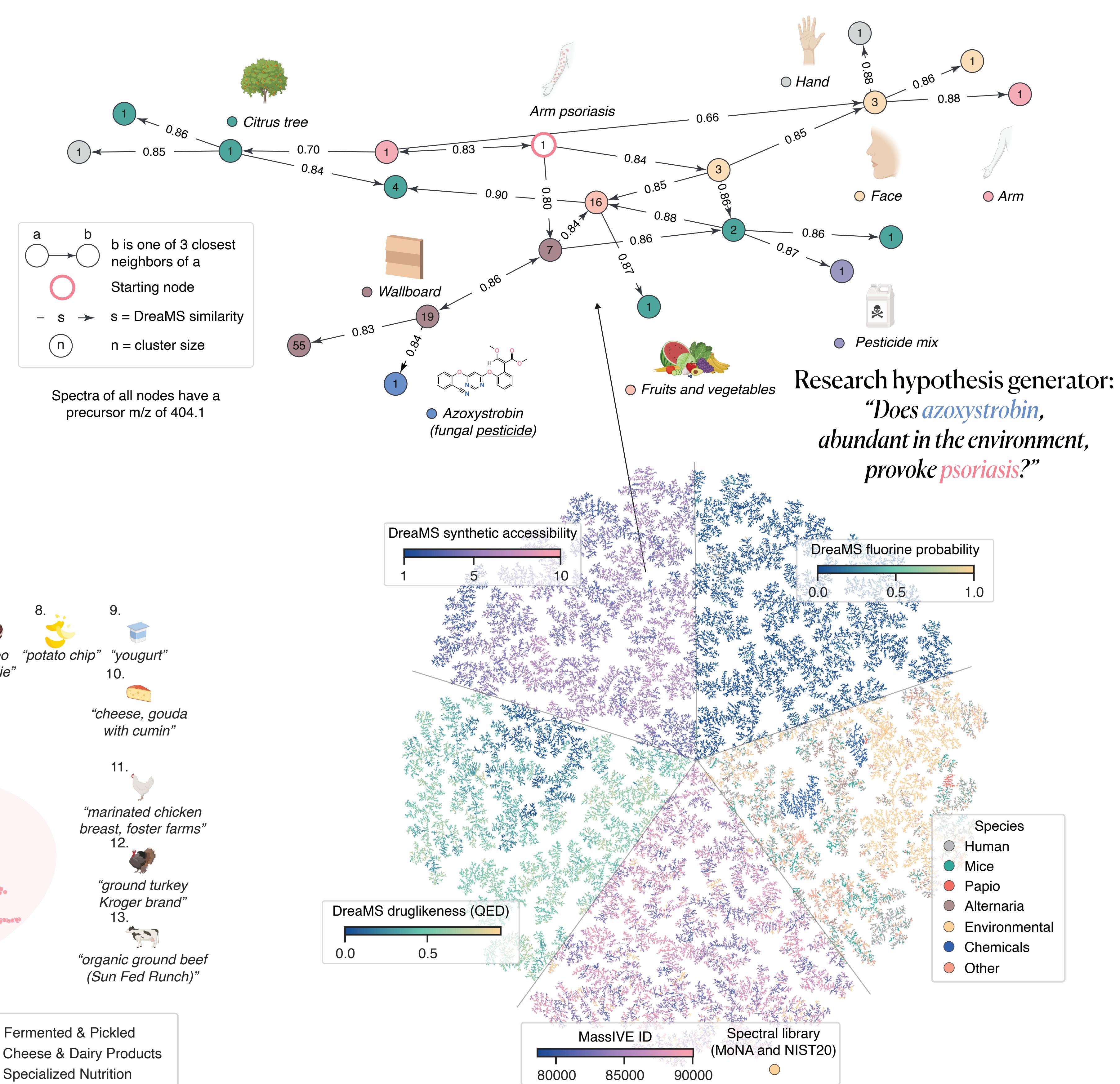
DreaMS embeddings of MS/MS spectra



b is one of 3 closest neighbors of a
Starting node
s = DreaMS similarity
n = cluster size

Spectra of all nodes have a
precursor m/z of 404.1

Molecular network of 201 million MS/MS spectra from MassIVE GNPS



DreaMS embeddings of LC-MS/MS experiments

