

# **CryoDRGN: a tool for reconstructing highly heterogeneous structural ensembles from cryo-electron micrographs**

**J Davis<sup>1</sup>, E Zhong<sup>2</sup>, B Berger<sup>2</sup>, T Bepler<sup>3</sup>**

**<sup>1</sup>Massachusetts Institute of Technology, Department of Biology, Program in Computational and Systems Biology, Cambridge, MA, USA, <sup>2</sup>Massachusetts Institute of Technology, Department of Mathematics, Program in Computational and Systems Biology, Cambridge, MA, USA,**

**<sup>3</sup>Massachusetts Institute of Technology, Program in Computational and Systems Biology, Cambridge, MA, USA  
[jhdavis@mit.edu](mailto:jhdavis@mit.edu)**

Massive macromolecular machines such as the ribosome and spliceosome perform essential cellular functions by undergoing dramatic structural changes. Some such complexes exhibit seemingly continuous conformational changes, whereas others transition between disparate states in a highly cooperative manner such that conformational intermediates are sparsely populated. To understand the extent of complex heterogeneity, the degree of cooperativity in their conformational changes, and to estimate ensembles of such structures, we have developed a neural-network based single-particle analysis framework. This approach, named cryoDRGN, maps individual particles to a low-dimensional latent space, which we find arranges structurally related particles in close proximity. Direct inspection of such a structurally ordered latent space provides insights into the degree of structural heterogeneity in the dataset, provides estimates of particle abundance in each state, and relates the observed states to one another. Additionally, we demonstrate that three-dimensional structures can be sampled from this latent space, allowing users to visualize conformational trajectories sampled along the data manifold. In this talk, I detail our method and apply it to exemplar simulated and experimental datasets to illustrate its utility in analyzing both continuously and discretely heterogeneous complexes. We find that these neural networks accurately estimate the conformational landscape of simulated datasets, reveal significant discrete and continuous conformational heterogeneity in experimentally derived particle stacks, and accurately produce ensembles of medium-resolution density maps, which can be used to understand how the complex's structural changes drive its function.