## Optimal clustering Feature similarity: ABIDE DSM (IV) Cat. WISC (PCA) **ADOS total** Shared-feature space Vineland (PCA) Subject similarity Cluster fit DSM (IV) Cat. Model fit (r) Model fit (r) 24000 ADOS (PCA) atent Dim 2 (UMAP, A.U.) 23500 ADOS Social 23000 ADOS Total 22500 Shared ASD VAE Shared ASD FIQ 🎝 22000 (n.s.) (\*\*\*) Gender 21500 Age ADI (PCA) 21000 Gender .15 20500 -2 ó 1 2 3 4 5 6 7 8 9 10 Model fit (r) Latent Dim 1 (UMAP, A.U.) Number of clusters Model fit (r) Scanner Type ASD-specific feature space **Scanning Site** Subject similarity Cluster fit 30000 VAE Shared ASD Shared ASD Model fit difference (ASD-specific - Shared) Latent Dim 2 (UMAP, A.U.) 29000 <u> </u> 28000 Feature similarity: SFARI 표 27000 Genotype DSM (IV) Cat. 26000 Genotype 25000 .12 Model fit (r) 24000 Model fit (r) .08 SFARI CNV DSM (IV) Cat. -2 -1 0 2 3 4 5 6 7 8 9 10 0.04 Number of clusters Latent Dim 1 (UMAP, A.U.) VAE space -.02 Cluster fit Subject similarity Shared ASD VAE Shared ASD 75000 Gender Gender atent Dim 2 (UMAP, A.U.) 70000 65000 Model fit (r) Model fit (r) 60000 55000

VAE Shared

(\*\*\*)

50000

2 3 4 5 6 7 8 9 10

**Number of clusters** 

-2 -1 0

Latent Dim 1 (UMAP, A.U.)

Scanner Type

Shared ASD

VAE

0.05

-0.15 -0.10 -0.05 0.00

Model fit difference (ASD-specific - Shared)