

**Supplementary Fig. 1.** Data Distribution After Processing (D represents drop, N represents NaN value)

SUPPLEMENTARY TABLE I

Train and Validation Loss of all AEs

|  |  |
| --- | --- |
|  |  |
| Initial Architecture, 468 Protein Expressions Encode to Dimension of 6 | Initial Architecture, 456 Protein Expressions Encode to Dimension of 6 |
|  |  |
| Initial Architecture, 468 Protein Expressions Encode to Dimension of 2 | Initial Architecture, 456 Protein Expressions Encode to Dimension of 2 |
|  |  |
| Initial Architecture, 468 Protein Expressions Encode to Dimension of 12 | DAE, 468 Protein Expressions Encode to Dimension of 12 |
|  |  |
| DAE, 468 Protein Expressions Encode to Dimension of 48 | DAE, 468 Protein Expressions Encode to Dimension of 36 |
|  |  |
| DAE, 468 Protein Expressions Encode to Dimension of 3 | DAE, 468 Protein Expressions Encode to Dimension of 32 |
|  |  |
| DAE, 468 Protein Expressions Encode to Dimension of 28 | DAE, 468 Protein Expressions Encode to Dimension of 24 |
|  |  |
| **DAE, 468 Protein Expressions Encode to Dimension of 18** | SAE, 468 Protein Expressions Encode to Dimension of 12 |
|  |  |
| DSAE, 468 Protein Expressions Encode to Dimension of 18 |