**Table S1.** The summary of deep learning models

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Optimizer | Learning Rate | Epochs | Batch Number | Loss Function |
| Fitness – Prediction | Adam | 10-4 | 2000 | 4 | MSE |
| Fitness – VAE | Adam | 10-4 | 5000 | 4 | MSE+KL |
| Sequence – Prediction | Adam | 10-5 | 20000 | 100 | MSE+KL |
| MD – Prediction | Adam | 10-5 | 20000 | 100 | MSE+KL |
| MD – Prediction with Attention | Adam | 10-5 | 20000 | 4 | MSE |
| MD – VAE | Adam | 10-4 | 5000 | 4 | MSE+KL |