

Table 1: Performance of different variants of IsoFormer for the prediction of transcript isoform expression using three different DNA encoders: Nucleotide Transformer, Enformer, and Borzoi (input sequence length in parentheses). R^2 and Spearman correlation across tissues for 5 different random seeds is reported. NT is used as RNA encoder while ESM is used to process protein sequences.

Base DNA Model	NT (12k)		Enformer (196k)		Borzoi (512k)	
	R^2	Spearman	R^2	Spearman	R^2	Spearman
DNA Only	0.13 ± 0.02	0.43 ± 0.01	0.21 ± 0.02	0.46 ± 0.00	0.12 ± 0.02	0.35 ± 0.01
RNA Only	0.36 ± 0.03	0.61 ± 0.01	0.36 ± 0.03	0.61 ± 0.01	0.36 ± 0.03	0.61 ± 0.01
Protein Only	0.20 ± 0.01	0.46 ± 0.01	0.20 ± 0.01	0.46 ± 0.01	0.20 ± 0.01	0.46 ± 0.01
RNA + Protein	0.40 ± 0.01	0.63 ± 0.01	0.40 ± 0.01	0.63 ± 0.01	0.40 ± 0.01	0.63 ± 0.01
DNA + Protein	0.28 ± 0.01	0.52 ± 0.01	0.39 ± 0.01	0.61 ± 0.01	0.30 ± 0.02	0.56 ± 0.01
DNA + RNA	0.39 ± 0.01	0.64 ± 0.01	0.52 ± 0.01	0.72 ± 0.01	0.47 ± 0.01	0.69 ± 0.00
DNA + RNA + Prot.	0.43 ± 0.01	0.65 ± 0.01	0.53 ± 0.01	0.71 ± 0.00	0.48 ± 0.01	0.69 ± 0.00

Table 2: Comparing the use of pre-trained and non-pre-trained encoders within IsoFormer. For this set of experiments the considered encoders are the Enformer for DNA, NT for RNA and ESM for proteins. ✓ indicates the use of a pre-trained encoder whereas ✗ indicates the encoder is trained from scratch (random initialization).

DNA	RNA	Protein	R^2	Spearman
✗	✗	✗	0.10 ± 0.03	0.31 ± 0.01
✓	✗	✗	0.45 ± 0.01	0.67 ± 0.00
✗	✓	✗	0.39 ± 0.01	0.61 ± 0.00
✗	✗	✓	0.34 ± 0.01	0.59 ± 0.01
✓	✓	✗	0.52 ± 0.01	0.71 ± 0.00
✓	✗	✓	0.48 ± 0.01	0.69 ± 0.00
✗	✓	✓	0.41 ± 0.01	0.64 ± 0.01
✓	✓	✓	0.53 ± 0.01	0.72 ± 0.00

Table 3: Comparing the performance of two Nucleotide Transformer (NT) models when finetuned on the BulkRNA downstream task. We compare the Base NT model (pre-trained) and IsoFormer’s DNA encoder after fine-tuning on transcript isoform expression prediction alongside RNA and protein.

Model	R^2	Spearman
Base NT	0.27	0.59
IsoFormer’s NT	0.39	0.67