

	Baseline CNN	ATTN	ATTN	HyenaDNA	HyenaDNA	HyenaDNA K-mer	Bidirectional Hyena	DNABERT
	scratch	scratch	pretrain	scratch	pretrain	scratch	scratch	pretrain
Mouse Enhancers	69	79.3	79.3	84.7	85.1	81.8	80.6	66.9
Coding vs Intergenic	87.6	89.3	91.2	90.9	91.3	86.7	90.3	92.5
Human vs Worm	93	94.8	96.6	96.4	96.6	92.9	95.9	96.5
Human Enhancer Cohn	69.5	67.7	72.9	72.9	74.2	69.8	72.1	74
Human Enhancer Ensembl	68.9	79.0	88.3	85.7	89.2	88.0	85.9	85.7
Human Regulatory	93.3	90.2	91.8	90.4	93.8	90.2	89.1	88.1
Human Nontata Promoter	84.6	85.2	90.1	93.3	96.6	83.5	88.5	85.6
Human OCR Ensembl	68	68.3	79.9	78.8	80.9	70.2	75.3	75.1

Table R1. GenomicBenchmarks ablations.

	NT 2.5B	AttnDNA	HyenaDNA	HyenaDNA
	pretrained	pretrained	pretrained	scratch
enhancer	58	59.3	59.7	58.6
enhancer_types	47.4	51.9	51.9	48.4
H3	81.4	75.8	82.3	79.9
H3K4me1	55.9	38.7	56.7	43.4
H3K4me2	32.6	28.8	51.8	34.5
H3K4me3	42.1	28.3	61.2	40.2
H3K9ac	57.5	49.2	63.6	52.6
H3K14ac	55	41.1	65.5	48.0
H3K36me3	63.2	47.8	65.7	53.4
H3K79me3	64.2	58.9	71.4	59.7
H4	82.2	77.7	79.8	79.1
H4ac	50.1	36.4	63.3	43.5
promoter_all	97.4	96.3	96.5	96.1
promoter_non_tata	97.7	96.6	96.7	96.5
promoter_tata	96.4	96.6	96.4	96.1
splice_sites_acceptor	99	99	96.6	96.6
splice_sites_donor	98.4	98.1	96.8	96.5
splice_sites_all	98.3	98	97.9	97.3

Table R2. Nucleotide Transformer benchmarks.

Model	HyenaDNA	DNABERT	NT 2.5B	
Params	436K	110M	2.5B (0.5% trainable)	
GPUs	1-A10	1-A10	1-A10	
Mouse Enhancers	2s	32s	9m 8s	
Human Enhancer Cohn	14s	9m 25s	41m 55s	
Human Nontata Promoter	10s	5m 44s	35m 36s	
Coding vs Intergenic	25s	12m 32s	1h 40m	
Human vs Worm	25s	12m 31s	1h 46m	
Human Enhancer Ensembl	1m 20s	55m 30s	2h 56m	
Human OCR Ensembl	1m 32s	1h 3m	3h 30m	
Human Regulatory	2m 37s	1h 46m	6h 39m	

Table R5. Finetune runtime comparison per epoch on GenomicBenchmarks

Length	Scratch	Pretrained
1k	53.9	61.1
32k	70.7	93.4
250K	65.7	97.9
450K	64.5	99.4

Table R3. Species classification (acc)

	DNABERT	Nucleotide Trx	HyenaDNA
Params	110M	2.5B	436K
GPUs	8-2080 TI	128-A100-80GB	1-A100-40GB
Wall clock	25 days	28 days	80 mins
GPU-hrs	12,000	215,000	1.3

Table R4. Pretraining compute comparison.