



Output

12 Layers

Input

clone

10 layers

2 Layers

clone

2 Layers

10 layers

graft

graft

Linear

Crown Layers

Linear

Root Layers

training

Step 1. Depthwise Scalling

Step 2. Grafting

Residue embed

Position embed

Region embed

E	...	G	M	...	M	V	...	R	M	...	M	K	...	Y	M	M	...	M	M	W	...
1	...	25	26	..	33	34	...	50	51	...	58	59	...	95	96	97	...	107	108	109	...
f	...	f	c1	...	c1	f	...	f	c2	...	c2	f	f	f	c3	c3	...	c3	c3	f	...