

A - Generalised profile construction

homeodomain alignment (eg PG9)



BLOSUM substitution matrix

	A	R	N	D	C	Q	E
A	8	-3	-4	-5	-2	-2	-3
R	-3	10	-2	-5	-8	0	-2
N	-4	-2	11	1	-5	-1	-2
D	-5	-5	1	10	-8	-2	2
C	-2	-8	-5	-8	14	-7	-9
Q	-2	0	-1	-2	-7	11	2

pftools



B - HoxPred classification principle



score

score

score

score

score



Prediction + posterior probability

PG4 (1)

>sequence_to_classify
NPSYNGGEPKRSRTAYTRQQVL
ELEKEFHFNRYLTRRRRIEIAH
TLCLSERQVKIWFQNRRLMKWKK

C - Linear discriminant classifier training



trains to classify from these data

dataset	sequence	scores for each profile					known class
HOX	D9_HETFR	2910	5037	4185	3131	2826	PG9
	D9_LATME	2927	4959	4095	3085	2746	PG9
	D9_MOUSE	2926	5019	4113	3148	2849	PG9
	A10_LATME	2772	5756	3972	2995	2719	PG10
	A10_MOUSE	2932	5892	4132	3155	2879	PG10
	A10_XENTR	2949	5832	4072	3225	2921	PG10
RANDOM	rand_1	251	312	488	246	305	CTL
	rand_10	176	316	189	134	153	CTL
HOMEO	ALX3_HUMAN	1952	1227	1634	1459	1584	CTL
	ALX3_MOUSE	1915	1227	1634	1459	1584	CTL
	ALX4_BOVIN	1949	1364	1710	1468	1614	CTL