

Supplementary material:

Table 1: Result set obtained by Nearest Neighbor classifier

Families	Globin	Homeobox	HSP	Kinase	Efficiency
Globin	299	0	1	0	99.67%
Homeobox	0	277	2	21	92.33%
Heat Shock protein	10	26	203	61	67.67%
Kinase	3	2	55	240	80.00%
Efficiency	84.92 %				

After excluding kinase protein family we have the following results.

Table 2: Result set obtained by Nearest Neighbor classifier excluding kinase protein family

Families	Globin	Homeobox	HSP	Efficiency
Globin	297	2	1	99.00%
Homeobox	1	299	0	99.67%
Heat Shock protein	8	9	283	94.34%
Efficiency	97.67%			