



L: protein length
 M: number of PTM sites on the whole protein
 D: maximum distance between neighbor sites in the cluster
 N: number of PTM sites in the cluster
 P: probability of the cluster

$$P = \left(1 - \left(1 - \frac{M}{L}\right)^D\right)^N$$

$$P1 = \left(1 - \left(1 - \frac{15}{300}\right)^3\right)^4 = 0.00041$$

$$P2 = \left(1 - \left(1 - \frac{15}{300}\right)^{11}\right)^7 = 0.0028$$



Cluster1

Cluster2