

Decoy database search : A primer

Mass spectral matching for multiple spectra is generally carried out on a global scale. Since manual confirmation of all the identifications in such large datasets is not possible, statistical methods are used to estimate the quality of the protein identification process. In order to measure the extent of error, methods which involve searches against decoy databases have been employed. One such decoy database is reverse database, wherein amino acid sequences are reversed so as to estimate error rate in a spectral search. Any spectral match to a reverse amino acid sequence represents a “false hit” which can be used to false positive rate in the dataset. Recent reviews have extensively covered the methodology used and advantages of false positive rate analysis in mass spectral search using decoy databases. Readers are recommended to read research articles by Elias and Gygi (2007), Choi and Nesvizhskii (2008) and Tabb (2007) for further details.

- 1) Elias JE, Gygi SP. *Nat Methods*. 2007, 4, 207-214.
- 2) Choi H, Nesvizhskii AI. *J Proteome Res*. 2008, 7, 47-50.
- 3) Tabb DL. *J Proteome Res*. 2007, 7, 45-46.