

The statistics of the No. of MS, PSMs, Peptides for original identification and Consensus MS identifications.

1. Why MARA-best and PRIDE-best has much less MS?
   1. Because best method only prepares consensus spectrum for clusters which have identified MSs. So Mara-clustering has (11218-4727) clusters which are unidentified; PRIDE-clustering has (9232-3755) clusters which are unidentified.
2. How to understand the No. MS between two clustering methods?
   1. Mara-cluster has 11218 clusters at p10 level;(???clusters at p20 level)
   2. PRIDE-cluster has 9232 clusters at 0.99 threshold.
3. What can we learn from this table?
   1. It looks like Mara-cluster has better identification Ratio on No.PSMs/No.MS: 0.844>0.819;
   2. It looks like Mara-cluster has better unique peptide identification ratio on No.Unique Peptide/No.Unique Peptide in Original Identification: 0.370>0.349;
   3. This trend also can be found at other RAW files too.
4. The description of columns.

|  |  |
| --- | --- |
| **#MS** | **The No. of total MS** |
| **Unique Peptide Sequences** | **The No. of Unique Peptides** |
| **One-time Peptide sequences** | **The No. of Unique Peptides which only appears one time in PSMs** |
| **Multi-times Peptide sequences** | **The No. of Unique Peptides which appears more than one times in PSMs** |
| **PSMs** | **The No. of PSMs** |
| **Proteins** | **The No. of referenced Proteins** |
| **No.PSMs/No.MS** | **The ratio of No.PSMs/No. MS** |
| **No.Unique Peptide/No.Unique Peptide of Original Identification** | **The ratio of Unique Peptide/ Unique Peptide in Original Identification** |
| **One-time Peptide /Unique Peptide** | **The ratio of One-time Peptide /Unique Peptide** |