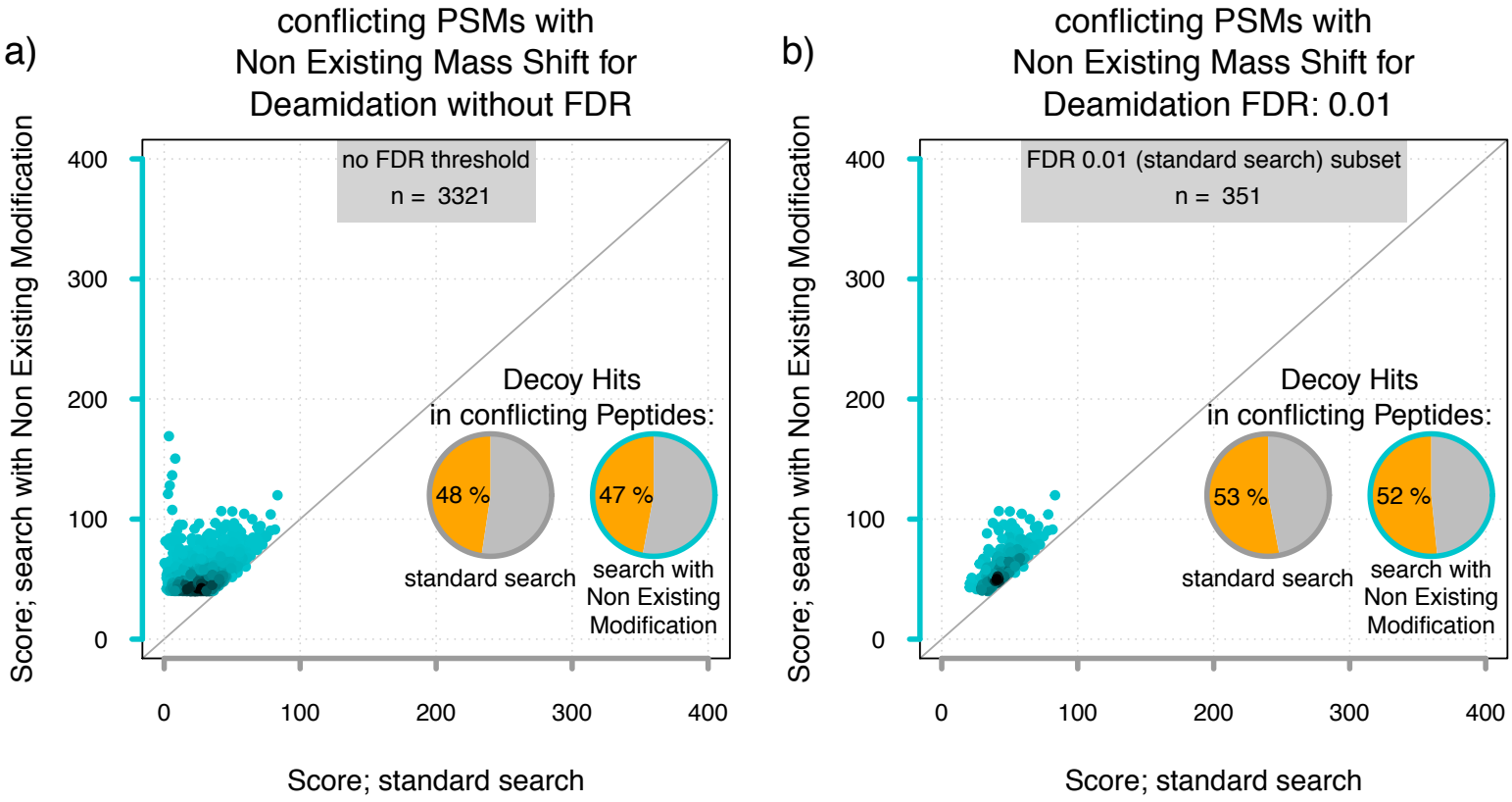
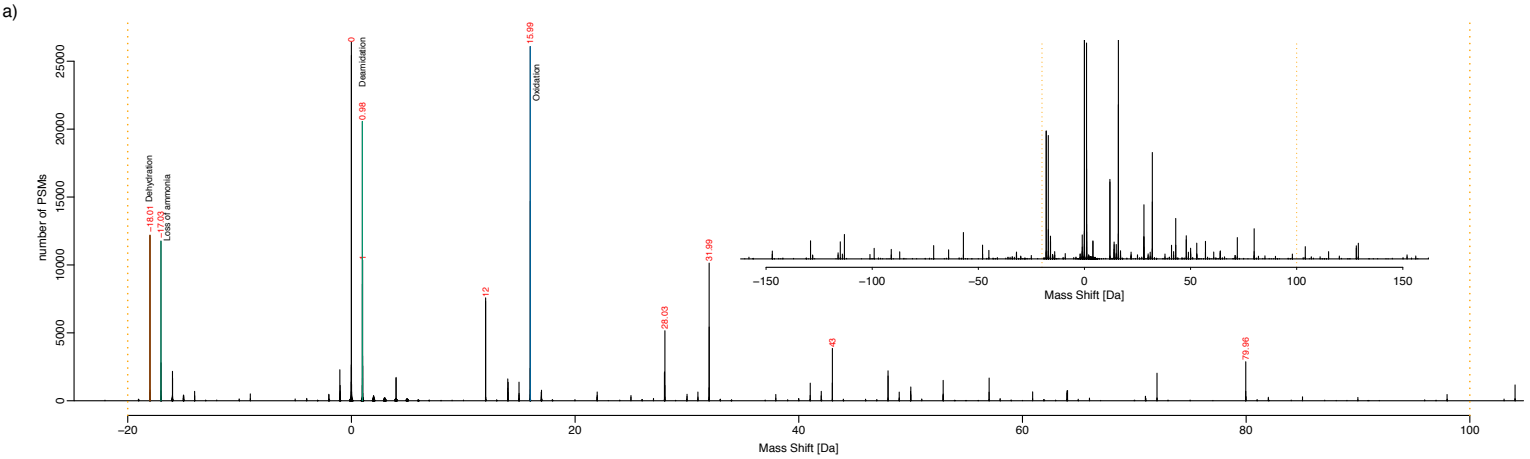
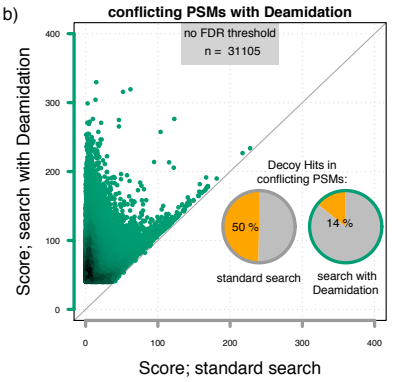
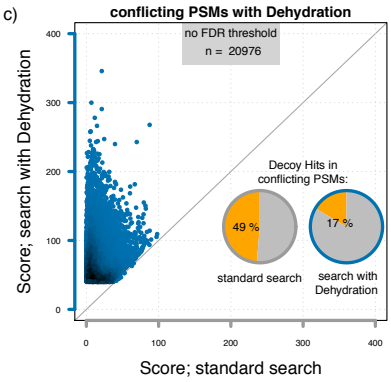
Supplemental Figure 1

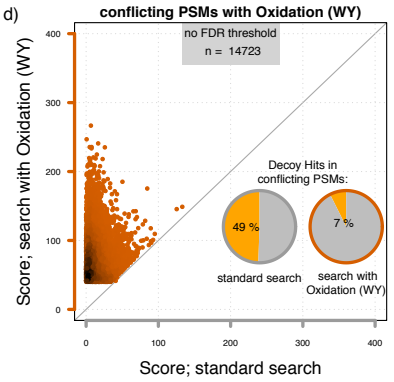


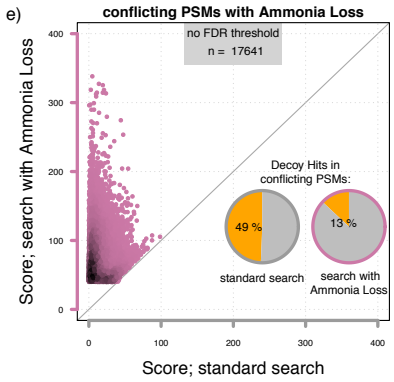
Supplemental Figure 2

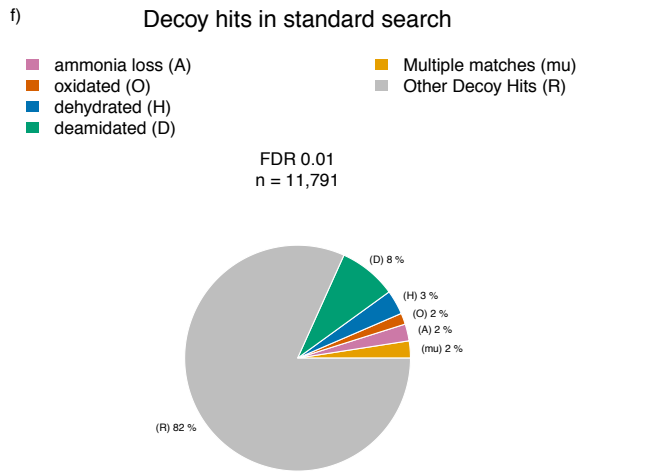


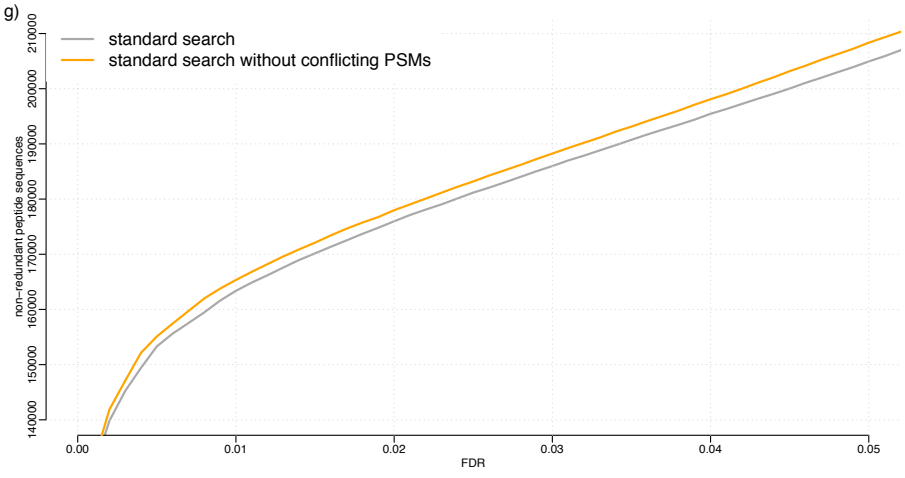




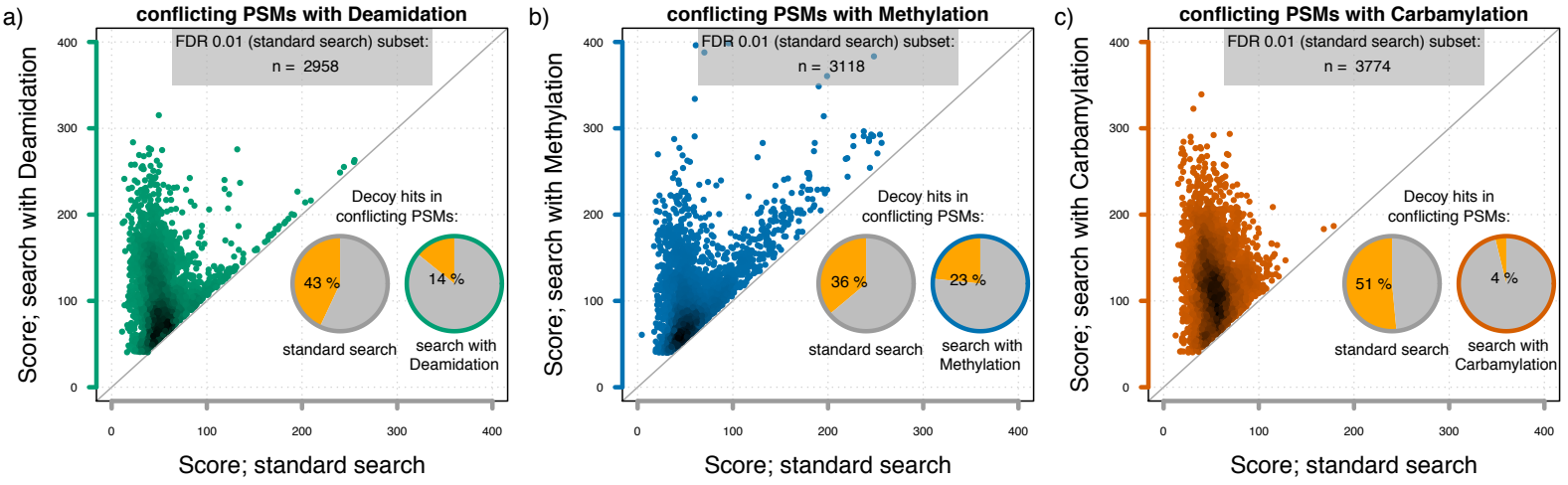




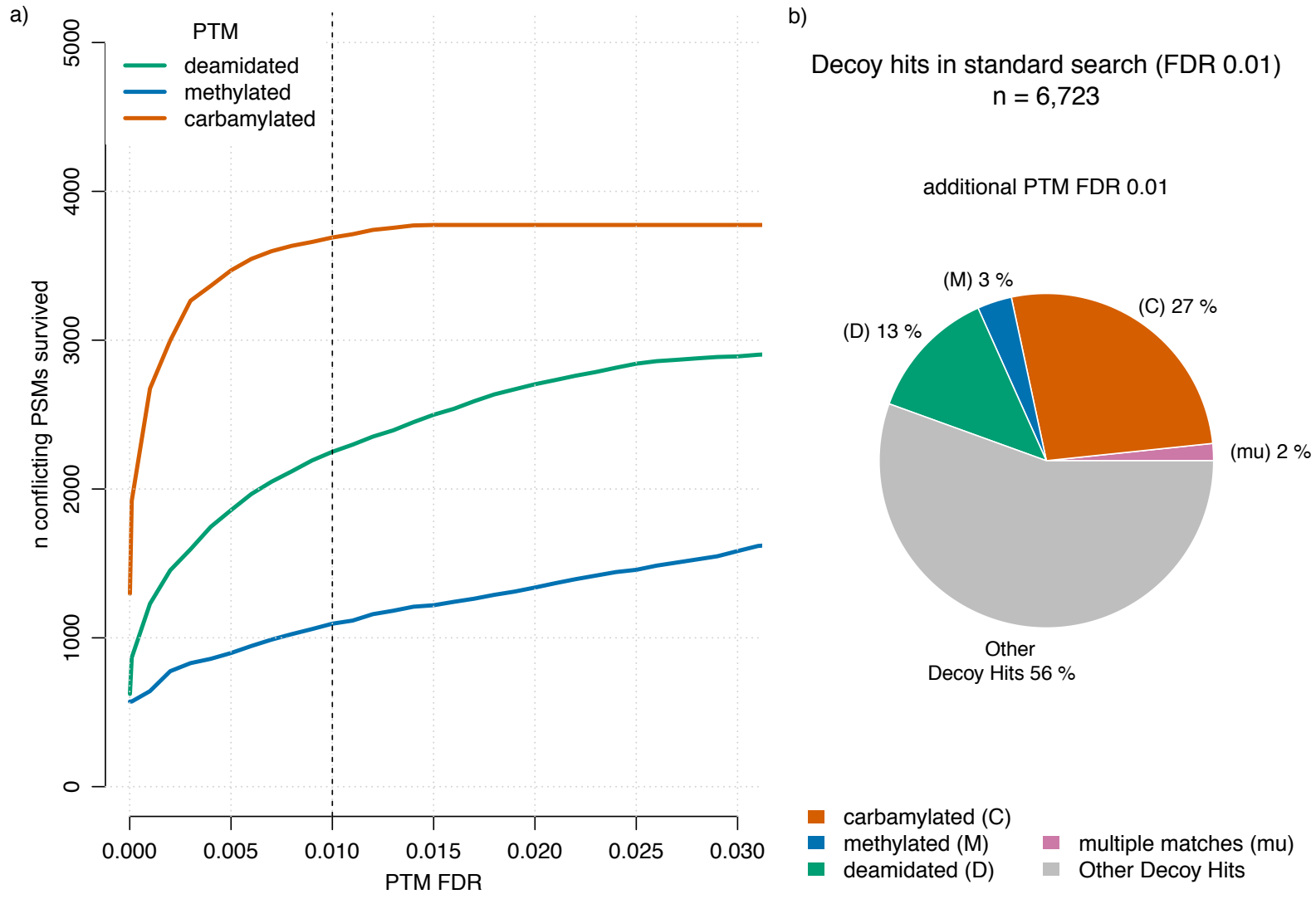




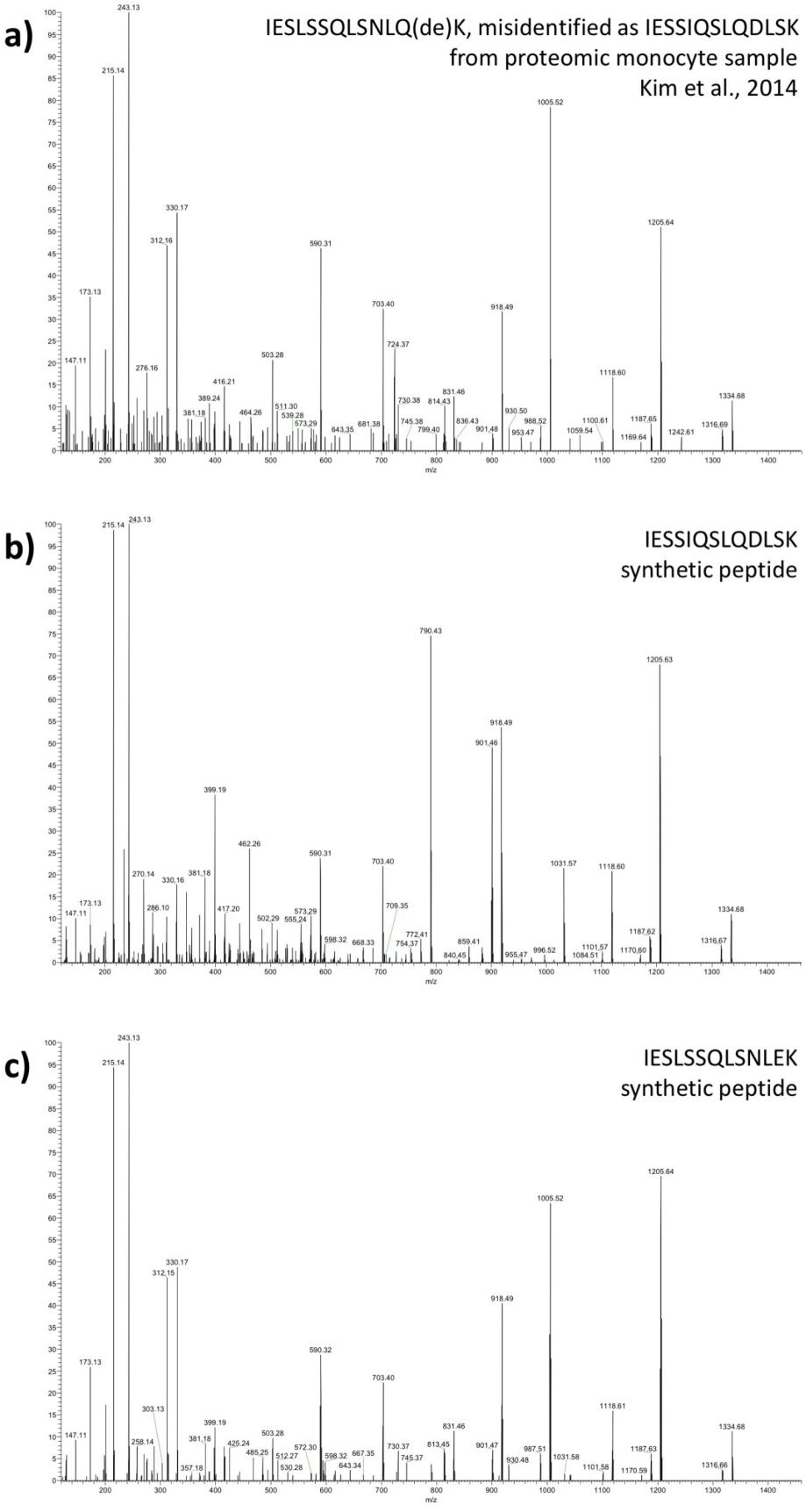
Supplemental Figure 3



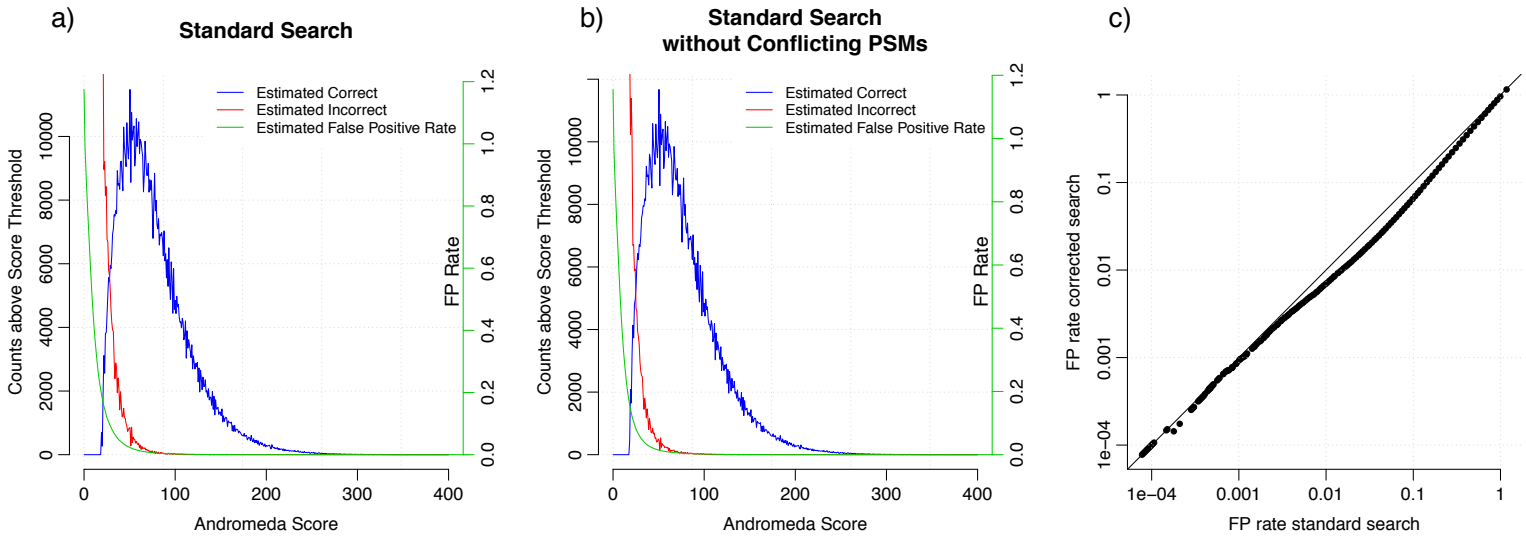
Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6



Supplemental Figure 1 Conflicting PSMs when a non-existing modification was considered. Pairwise comparison of Andromeda scores from conflicting PSMs with and without considering a non-existing modification. For this control search an arbitrary, but non-existent mass-shift (according to the ModifiComb estimation in Fig 1a), of -11.01 was selected and allowed to occur on N and Q residues. The fraction of decoy database hits for the conflicting subsets in either search variant is given as pie chart. In **(a)** without filtering the standard search and in **(b)** with filtering the standard search to a PSM FDR of 0.01. Note, for this control experiment none of the searches correctly identified this subset of scans.

Supplemental Figure 2 Modified peptides as a source of false spectra identification in a deep proteomic HeLa cell dataset from Nagaraj et al. (a) Distribution of most frequent mass shifts of base to dependent peptides in the region from −20 to 100 Da. A view on the wider mass region is shown in the insert. **(b-e)** Pairwise comparison of Andromeda scores from conflicting PSMs without FDR filtering in searches with and without deamidation (D), dehydration (H), oxidation (O) or ammonia loss (A), as indicated. The fraction of decoy database hits for the subset of conflicting PSMs in either search variant is given as a pie chart. **(f)** Fraction of conflicting PSMs among all decoy database hits from a standard search at PSM FDR 0.01. Decoy database hits that were explained by modified peptides in more than one of the searches are denoted as "mu" (multiple matches). **(g)** Spectra that gave rise to conflicting sequence information were removed (standard search without D, H, A and O) and compared to the uncorrected standard search. The number of identified non-redundant peptide sequences is plotted as a function of the PSM FDR.

Supplemental Figure 3 Modified peptides as a source of false spectra identification after filtering the standard search to a PSM FDR of 0.01. (a-c) Pairwise comparison of Andromeda scores from conflicting PSMs with and without deamidation, methylation or carbamylation, as indicated. The fraction of decoy database hits for the conflicting subsets in either search variant is given as a pie chart.

Supplemental Fig. 4 False positives due to modified peptides as a function of the PTM FDR. (a) The standard search was filtered to a PSM FDR of 0.01. PSMs from the filtered standard search were compared to the three individual modified searches. The number of conflicting PSMs that survive a subgroup filtering on modified PSMs is depicted as a function of the PTM FDR (= that is FDR filtering applied on PSMs reported to be modified). **(b)** Decoy hits from the standard search that survived a stringent 0.01 PSM FDR cut-off are shown. The fraction of conflicting PSMs that survived an additional 0.01 PTM FDR cut-off in the individual modified searches is depicted.

Supplemental Figure 5 Confirmation of peptide identity by synthetic peptides. (a) Tandem MS spectrum of a peptide reported as GRID2-derived with the sequence IESSIQSLQDLSK from a proteomic monocyte sample, as reported by Kim et al., 2014. Tandem MS spectra of **(b)** a synthetic GRID2-derived peptide with the sequence IESSIQSLQDLSK and **(c)** a synthetic deamidated LMNB1 peptide with the sequence IESLSSQLSNLEK.

Supplemental Figure 6 Influence of the targeted removal of conflicting PSMs on the estimated false positive rate. Populations of estimated incorrect (that is, number of decoy matches *2) and estimated correct (that is, total matches – 2* number of decoy matches) were simulated based on the data transformations proposed by Elias and Gygi, Nat. Methods, 2007 for **(a)** the standard search and **(b)** the search without conflicting PSMs. **(c)** Pairwise comparison of the false positive rate (cumulative) for both strategies in each score bin.