I used the iPRG2012_nd.mgf file as the peak list for searching the data with Protein Prospector (v5.10). This file gave more peptide IDs than the alternative formats because Protein Prospector is able to determine charge-states for fragment ions if the software is told it is high resolution MSMS data, which is determined by the instrument setting (I used Q-TOF as the setting). I searched data with a precursor mass tolerance of 8 ppm and a fragment mass tolerance of 30 ppm, allowing for up to 2 missed cleavages.

Initially I searched the data with a minimal list of modifications. I then used the list of identified proteins to perform a mass modification search to assess the types of modifications that could be observed in the sample. The results from this search were not used for my submitted results, but helped me define a list of defined modifications I should consider.

To generate the submitted results I performed four searches, and then merged the results using SearchCompare to produce the final report.

Search 1: Biological and common modifications only. Up to 4 mods per peptide.

Search 2: Biological + chemical modifications. Up to 3 mods per peptide.

Search 3: Biological + chemical modifications. Up to 2 mods per peptide, but allowing for semi-tryptic peptides.

Search 4: Biological + chemical modifications. Up to 2 mods per peptide, but allowing for mass modifications of +/- 2 m/z, observed as a neutral loss (this will identify spectra where the fragments correspond to a different precursor ion of similar m/z and same charge as the intended precursor that may have been co-isolated for MSMS).

I list below the modifications I considered: Those with a * were considered in Search 1; other searches considered all of these modifications.

*Acetyl (Protein N-term); *Acetyl (Uncleaved K); *Acetyl+Oxidation (Protein N-term M); ACN (Neutral loss); *Carbamidomethyl (C); Cation:K (DEST); Cation:Na (DEST); *Deamidated (N); Deamidated (Q); *Dimethyl (Uncleaved K); *Dimethyl (Uncleaved R); Formyl (K); Formyl (N-term); Formyl (ST); *Gln->pyro-Glu (N-term Q); Glu->pyro-Glu (N-term E); *Met-loss (Protein N-term M); *Met-loss+Acetyl (Protein N-term M); *Methyl (K); *Methyl (R); *Nitro (Y); *Oxidation (M); Oxidation *(W); *Phospho (STY); *Pyro-carbamidomethyl (N-term C); *Sulfo (Neutral loss); *TriMethyl (Uncleaved K); Trp->Hydroxykynurenin (W); Trp->Kynurenin (W)

Results were filtered to an estimated 1 %FDR using target-decoy searching based on expectation value estimates. Site assignment scoring is automatically reported using SLIP scoring by Protein Prospector.