# **Generating Identification Files**

Peptide identifications reported from many proteomic pipelines have the option of being exported in a tabular format with columns delimited by some sort of separation character, commonly tabs or commas. Many of these files, however, contain some measure of metadata which is not necessary for spectral annotation. As such, IPSA has been configured to extract peptide identifications in a simple comma-separated value format (\*.csv). This file has 4 columns:

**Scan:** MS2 scan number

**Sequence:** Proposed peptide sequence

**Charge:** Precursor charge

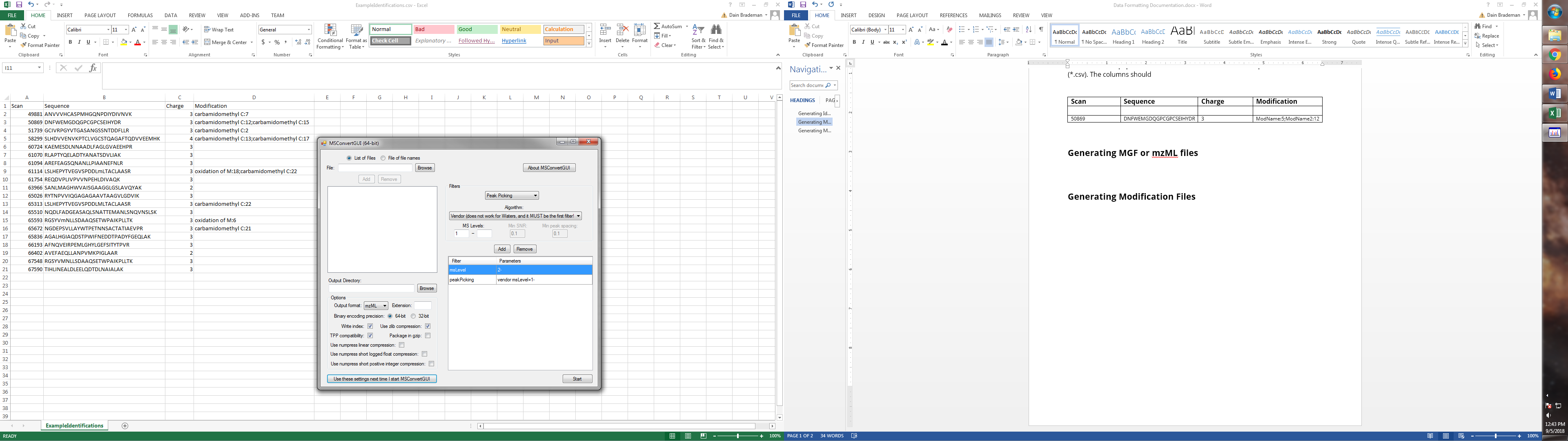
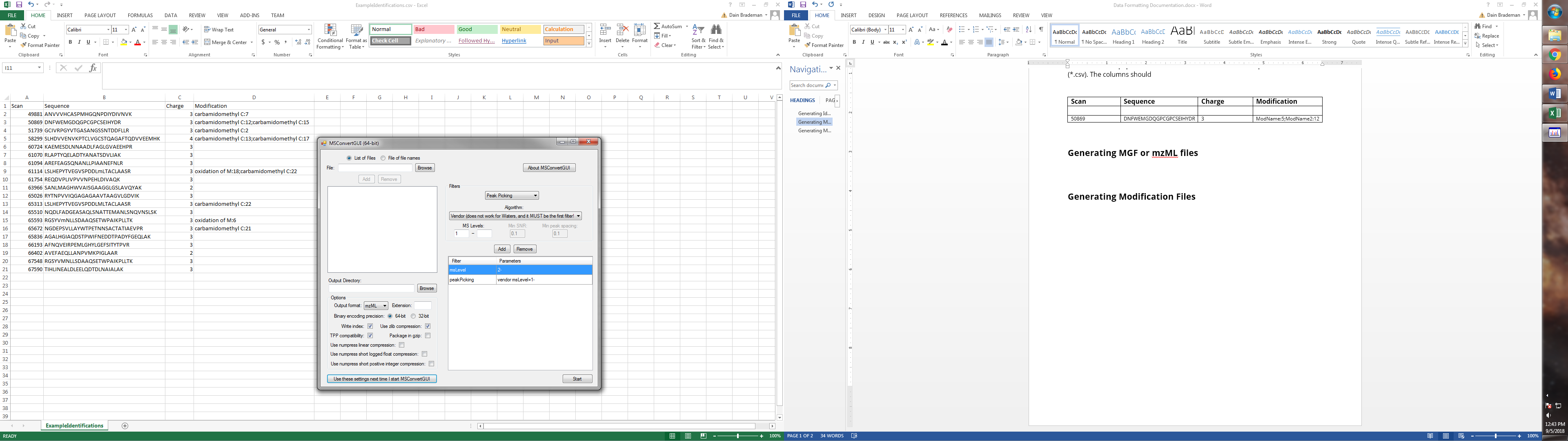
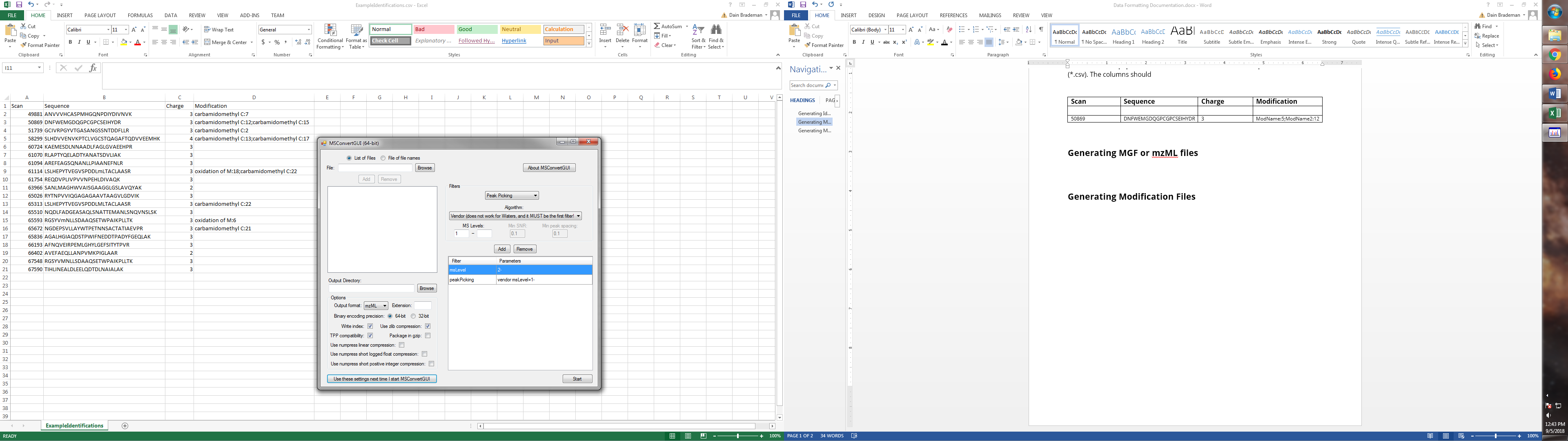
**Modification:** PTM names and locations. Name and location separated by a colon. Multiple modifications separated by semicolons

This identification file is parsed server-side to extract an MS2 scan number, the peptide sequence, precursor charge, and post-translational modifications. Each row in the identifications file should correspond to a unique scan number resulting in a peptide-spectral match, though the rows don’t necessarily need to be sorted.

Shown in the table below are several examples of peptides which IPSA is capable of processing

|  |  |  |  |
| --- | --- | --- | --- |
| **Scan** | **Sequence** | **Charge** | **Modification** |
|  |  |  |  |
| 50869 | DNFWEMGDQGPCGPCSEIHYDR | 3 | carbamidomethyl C:12;carbamidomethyl C:15 |
| 8203 | VNDVVCEQIANGENAITGVMIESNINEGNQGIPAEGK | -6 | carbamidomethyl C:6 |
|  |  |  |  |

# **Generating MGF or mzML files**



# **Generating Modification Files**