Using advanced options in msConvert

David L. Tabb, Ph.D.

dtabb1973@gmail.com

Overview

- Learning msConvert format support
- Dissecting filters for data processing
- Employing the CantWaiT peak lister

What can ProteoWizard read?

Vendor/Creator	Format
AB SCIEX	WIFF; T2D (with DataExplorer)
Agilent	MassHunter (.d directories)
Bruker	FID; .d directories; XMASS XML
Thermo	RAW
Waters	raw directories
HUPO PSI	mzML
ISB Seattle Proteome Center	mzXML
Matrix Science	MGF
Yates/MacCoss Laboratories	MS2/CMS2/BMS2
Steen & Steen Laboratory	mz5

Red indicates ability to write format

Holman et al, *Curr. Protoc. Bioinfo*. (2014) Supp46: 13.24

Filters: Spectral Subsetters

index	<index_value_set>	Selects spectra by index - an index value 0-based numerical order in which the spectrum appears in the input.
*msLevel	<mslevels>	Selects only spectra with the indicated MS levels.
*chargeState	<charge_states>	Keeps spectra that match the listed charge state. Use 0 to include spectra with no charge state at all.
*scanNumber	<scan_numbers>	Selects spectra by scan number.
scanEvent	<scan_event_set>	Selects spectra by scan event.
*scanTime	<scan_time_range>	Selects spectra within a given time range.
stripIT		Rejects ion trap data spectra with MS level 1.
mzPrecursors	<precursor_mz_list>	Retains spectra with specific precursor m/z values.
defaultArrayLength	<peak_count_range>	Keeps only spectra with peak counts within a given range.
*activation	<precursor_activation_type>	Keeps only spectra whose precursors have the specified activation type.
analyzer	<analyzer>	Keeps only spectra with the indicated mass analyzer type.
polarity	<polarity>	Keeps only spectra with scan of the selected polarity.

* indicates accessible in msConvertGUI

Subset examples

- I want to exclude MS/MS scans from data headed to XCMS in mzXML format:

```
msconvert –-filter “msLevel [1,1]” --mzXML *.raw
```

- I want to extract scans for peaks near 18 minutes (1080 sec) in these files:

```
msconvert –-filter “scanTime [1050,1110]” *.raw
```

- I want the negative ion mode scans:

```
msconvert –-filter “polarity negative” *.raw
```

Use standard double-quote on keyboard; do not copy text from the slide.

Filters: Processing

precursorRecalculation		Recalculates the precursor m/z and charge for MS2 spectra based on the MS1 data. Only works on orbitrap and FT data.
precursorRefine		Recalculates the precursor m/z and charge for MS2 spectra based on the MS1 data. Works on Orbitrap, FT, and TOF data.
*peakPicking	<prefer_vendor>, <ms_levels>	Performs centroiding on spectra with the selected MS levels. If used with other filters this must be set first.
sortByScanTime		Reorders spectra, sorting them by ascending scan start time.
metadataFixer		Add or replace a spectra's TIC/BPI metadata, usually after peak picking where the change from profile to centroided data may make the TIC and BPI values inconsistent with the revised scan data.
titleMaker	<format_string>	Adds or replaces spectrum titles according to format string given.
chargeStatePredictor	[<overrideExistingCharge>, [<maxMultipleCharge>, [<minMultipleCharge>, [<singleChargeFractionTIC>, [<algorithmMakeMS2>]]]]]	Predicts MSn spectrum precursors to be singly or multiply charged depending on the ratio of intensity above and below the precursor m/z.

Processing examples

- I need to get picked peaks from profile data‡:

```
msconvert --filter “peakPicking true 1-” *.raw
```

- I need to sort Waters scans for XCMS:

```
msconvert --filter “sortByScanTime” *.raw
```

- Combining both:

```
msconvert --filter “peakPicking true 1-” --filter
  “sortByScanTime” *.raw
```

‡ When stacking filters, this comes first!

Scrubbing MS/MS scans

*threshold	<type>, <threshold>, <orientation>, [<mslevels>]	Keeps fragments whose values meet various threshold criteria.
mzPresent	<tolerance>, <type>, <threshold>, <orientation>, <mz_list>, [<include_or_exclude>]	Keeps data whose values meet various threshold criteria. Contains a wider array of options than the "threshold" filter.
*mzWindow	<mzrange>	Keeps mz/intensity pairs whose m/z values fall within the specified range.
*zeroSamples	<mode>, [<MS_levels>]	Deals with zero values in spectra - either removing them, or adding them where they are missing.
MS2Denoise	[<peaks_in_window>, [<window_width_Da>, [multicharge_fragment_relaxation]]]	Noise peak removal for spectra with precursor ions.
MS2Deisotope	[<hi_res>, [<mz_tolerance>]]	Deisotopes MS level 2 spectra using Markey method.
*ETDFilter	[<removePrecursor>, [<removeChargeReduced>, [<removeNeutralLoss>, [<blanketRemoval>, [<matchingTolerance>]]]]]	Filters ETD MSn spectrum data points, removing unreacted precursors, charge-reduced precursors, and neutral losses.

Scrubbing examples

- Export MS/MS containing a major six-carbon sugar fragment within 0.5 of 163 m/z:

```
msconvert --filter “mzPresent 0.5mz count 10 most-intense [163]” *.raw
```

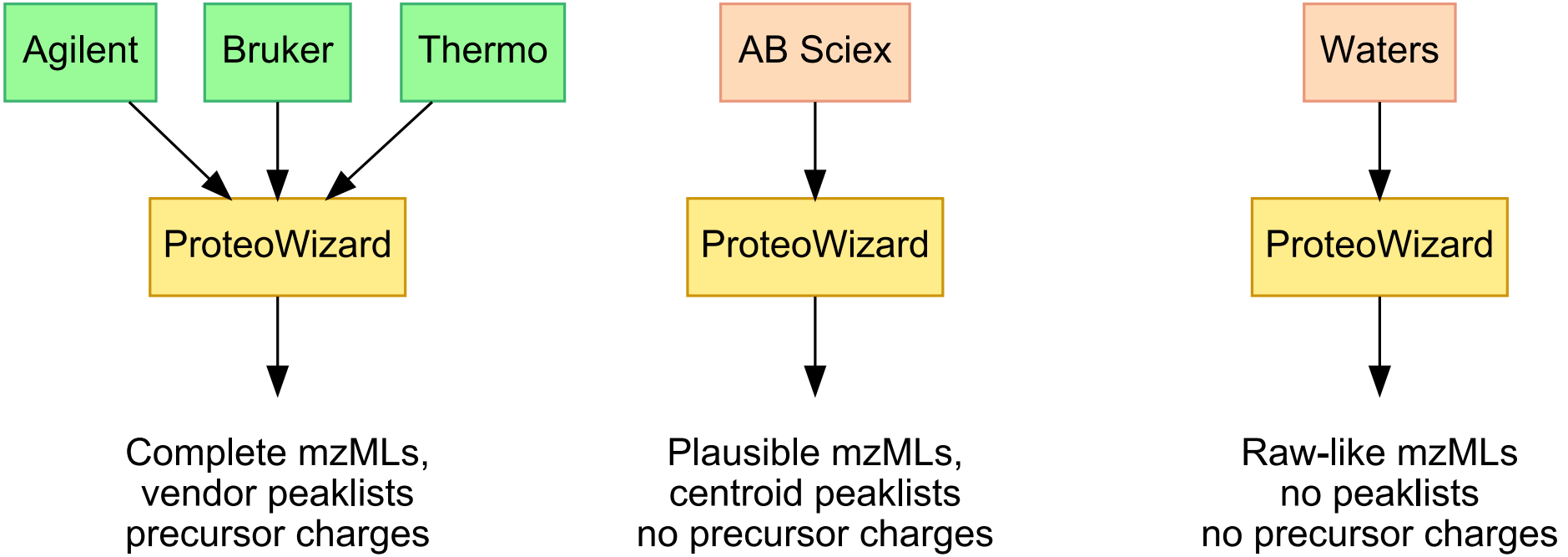
- Remove all zero-intensity fragments from MS/MS scans:

```
msconvert --filter “zeroSamples removeExtra 2” *.raw
```

- Deisotope tandem mass spectra:

```
msconvert --filter “MS2Deisotope Poisson” *.raw
```

ProteoWizard msConvert
strengths and weaknesses



- AB Sciex Converter serves as alternative
- Waters PLGS export to mzML seems broken
- Mascot Distiller yields MGF for MS/MS only

Enabling use of Waters data sets

Implementing functions other vendors provide

Needed functions

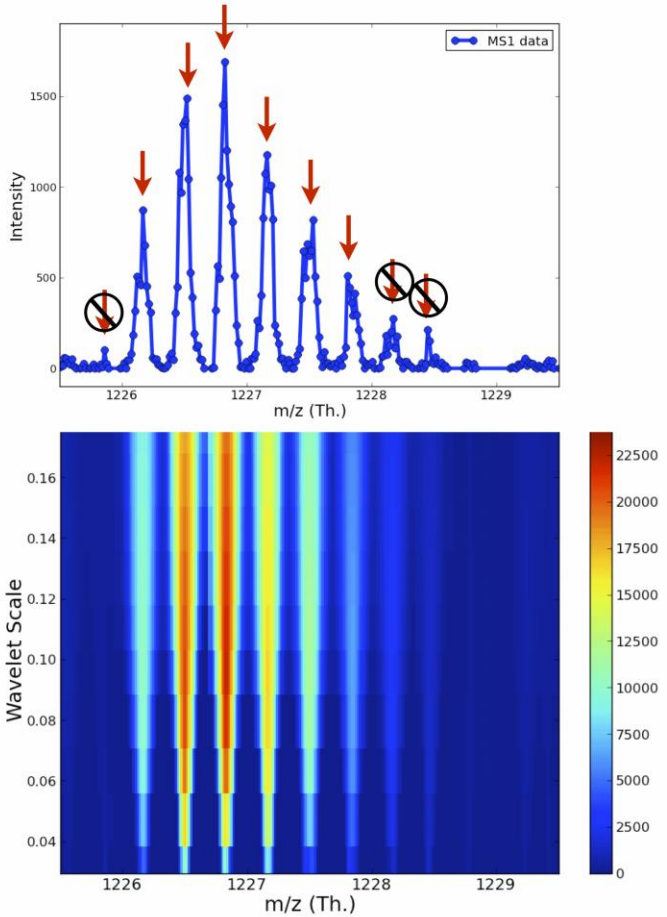
- Enumerate Peaks: reduce peak profiles to observed m/z values and intensities
- Infer Precursor Charge: determine spacing of peaks in isotopic packet
- Sum Sub-Scans: adding together profiles in advance of peak enumeration

Enabled applications

- Produce mzMLs for identification easily from raw data via msConvert
- Compare vendor peak-picking techniques against a quality open-source option
- Achieve improved mass accuracy through use of successive MS scans

Incorporating a high-end peaklister

- Scan summing combines signal from subscans.
- Continuous Wavelet Transform (CWT) divides profile scans into “Ricker wavelets.”
- Turbocharger infers z from isotope spacing and intensity patterns.



French et al. *J. Proteome Res*. (2015) 14:1299

open-source peak picking

turbocharger	[minCharge=<minCharge>] [maxCharge=<maxCharge>] [precursorsBefore=<before>] [precursorsAfter=<after>] [halfIsoWidth=<half-width of isolation window>] [defaultMinCharge=<defaultMinCharge>] [defaultMaxCharge=<defaultMaxCharge>] [useVendorPeaks=<useVendorPeaks>]	Determine charge states from precursor ion isotopic packet
scanSumming	[precursorTol=<precursor tolerance>] [scanTimeTol=<scan time tolerance>]	Accumulates signal across multiple subscans
*peakPicking	<prefer_vendor>, <ms_levels>	Performs centroiding on spectra with the selected MS levels. If used with other filters this must be set first.

Using open source MS/MS processing

- Get CantWaiT peak lists for MS/MS scans

```
msconvert --filter “peakPicking cwt msLevel=2-” *.raw
```

- Infer charges by Turbocharger, using MS isolation window of +/- 1.5 m/z

```
msconvert --filter “turbocharger halfIsoWidth=1.5”
  *.raw
```

- Do both and add deisotoping

```
msconvert --filter “peakPicking cwt msLevel=2-” --
  filter “turbocharger halfIsoWidth=1.5” --filter
  “MS2Deisotope Poisson” *.raw
```

Takeaway messages

- File format conversion is only the tip of the iceberg for what msConvert can perform.
- If you are concerned about the way vendor software performs, you can use our open source alternative.