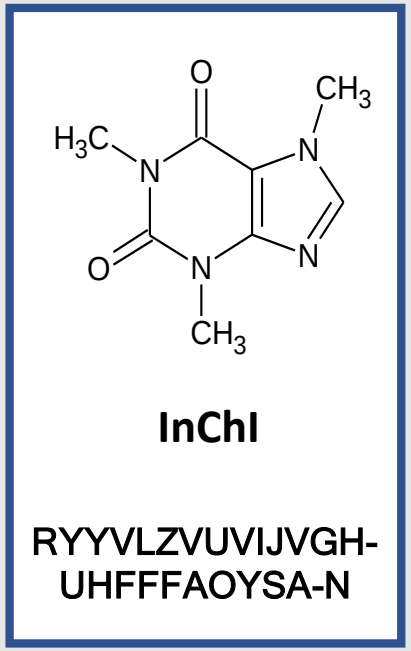
# The Many Forms of NIST 2020 MS Libraries

Steve Stein
NIST Mass Spectrometry Data Center
Biomolecular Measurement Division

# Key MS Library Qualities

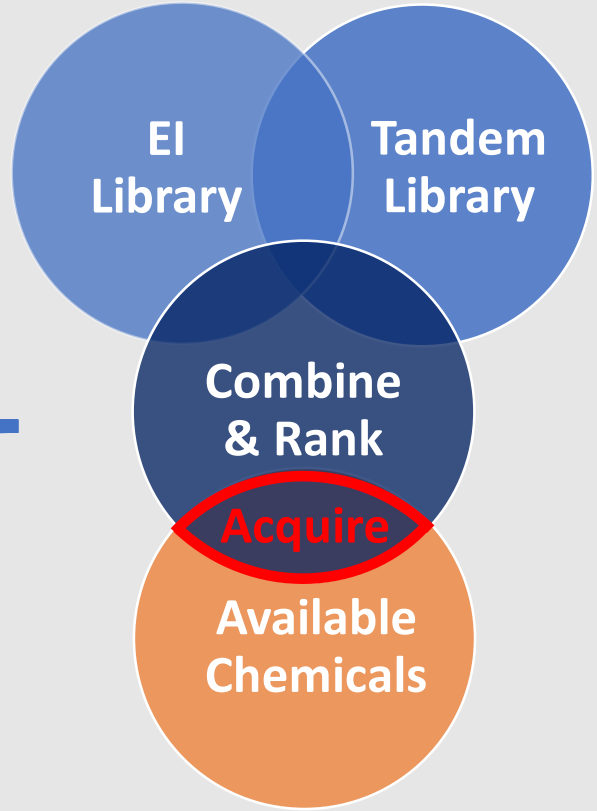
- Coverage
  - How big is it – NO!
  - Contains spectra of interest – YES!
    - at conditions of interest
- Quality
  - Can you rely on it?
  - All spectra curated, intercompared using best possible software
- Software
  - Find and confirm identification
  - Assists if compound not in library

# Coverage: Compound Selection

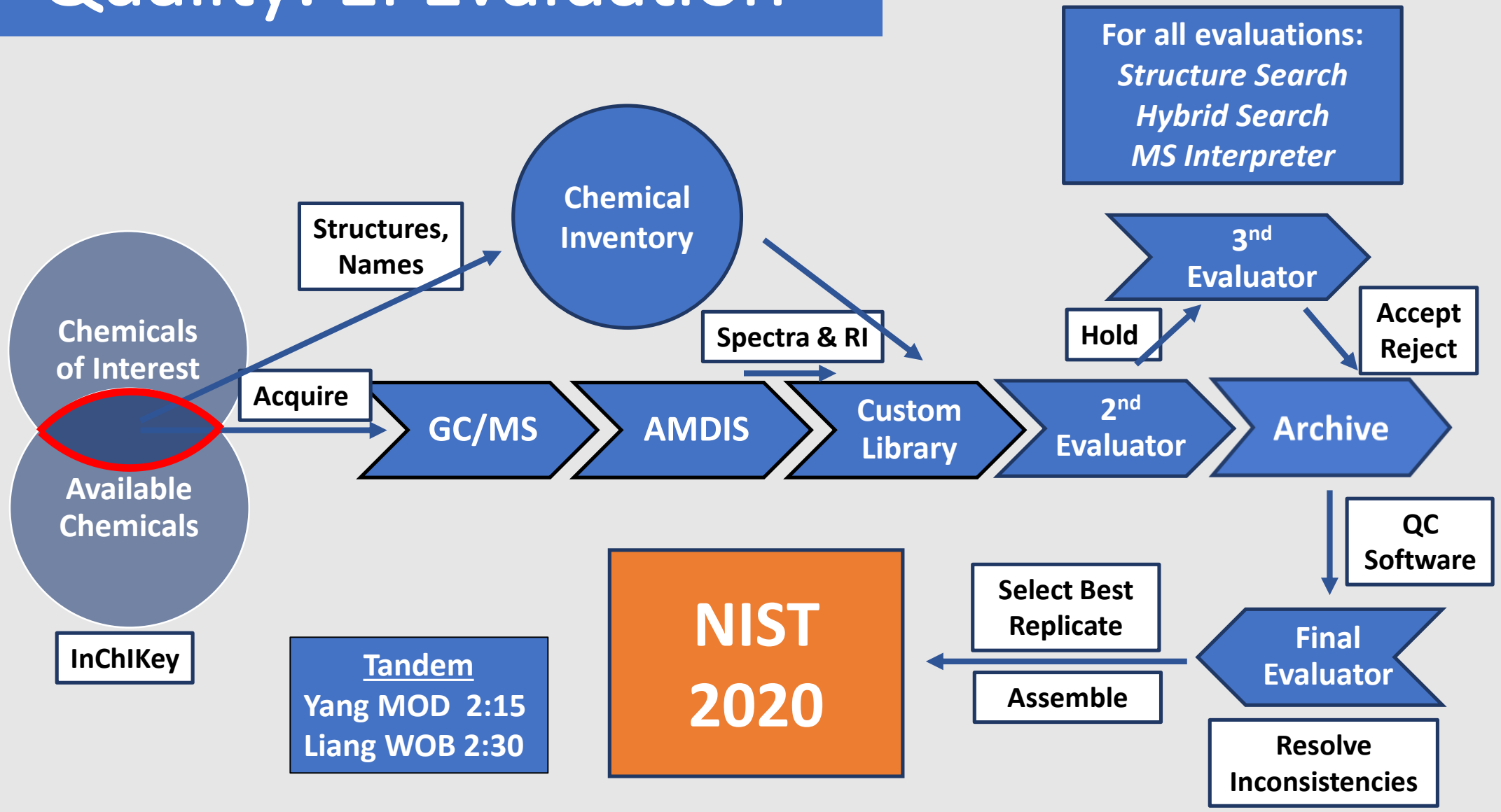



*30+ Collections*

- Wikipedia
- HMDB
- KEGG
- FoodDB
- EPA/FDA Lists
- Drug Bank
- ⋮



# Quality: EI Evaluation

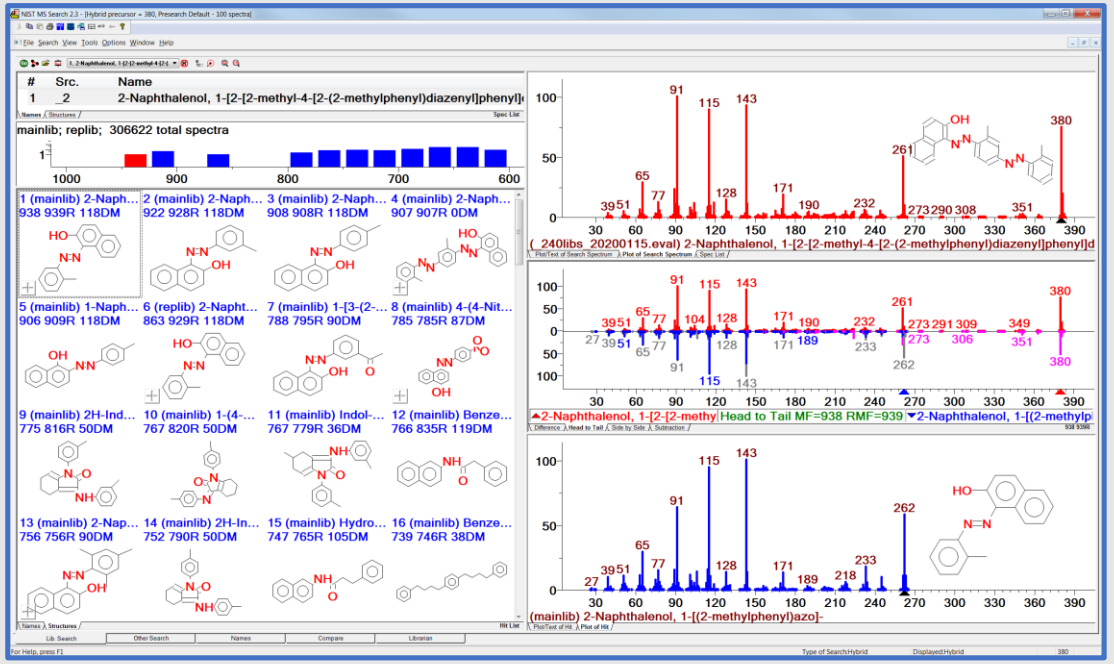


# SOFTWARE

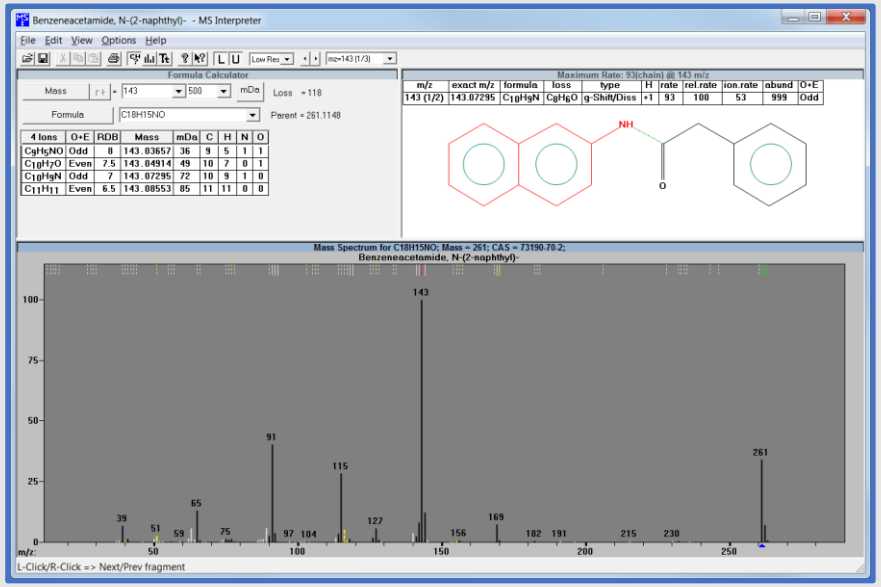
## NISTMS.EXE

### USER INTERFACE

### HYBRID SEARCH FOR COMPOUNDS NOT IN LIBRARY




### NEW AI RETENTION INDEX ESTIMATES




## MS INTERPRETER

### CONNECT PEAKS TO STRUCTURES

### MAJOR UPDATE

# Hybrid Search
## Identify Compounds Not in Library

analytical chemistry
Article
Cite This: *Anal. Chem.* 2019, 91, 2155−2162
pubs.acs.org/

### Structure Annotation of All Mass Spectra in Untargeted Metabolomics

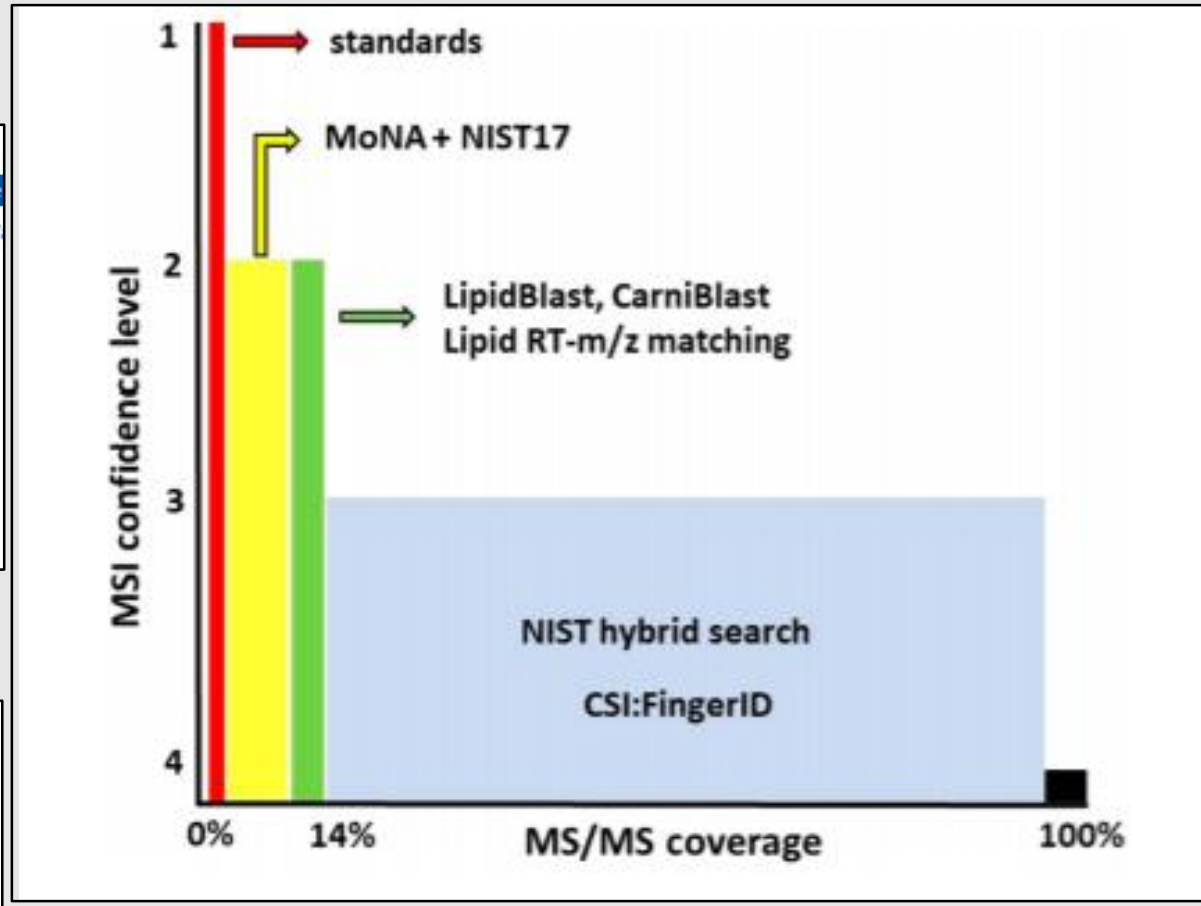
Ivana Blaženović,[†] Tobias Kind,[†] Michael R. Sa,[†] Jian Ji,[‡] Arpana Vaniya,[†] Benjamin Wancewicz,[†] Bryan S. Roberts,[†] Hrvoje Torbašinović,[§] Tack Lee,[∥] Sajjan S. Mehta,[†] Megan R. Showalter,[†] Hosook Song,[∥] Jessica Kwok,[†] Dieter Jahn,[⊥,#] Jayoung Kim,[▽,○,◆,¶] and Oliver Fiehn*[,†]

**MOD pm 3:50: Cooper *et al.***

analytical chemistry
Article
Cite This: *Anal. Chem.* 2019, 91, 13924−13932
pubs.acs.org/ac

### Hybrid Search: A Method for Identifying Metabolites Absent from Tandem Mass Spectrometry Libraries

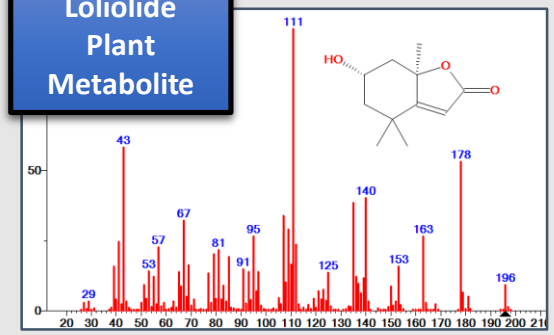
Brian T. Cooper,*[,†,‡] Xinjian Yan,[‡] Yamil Simón-Manso,[‡] Dmitrii V. Tchekhovskoi,[‡] Yuri A. Mirokhin,[‡] and Stephen E. Stein[‡]



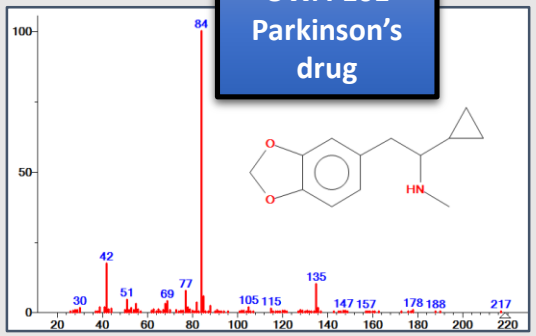
# NIST/EPA/NIH EI MS Library – NIST 20

**306,869 Compounds, 43,774 Replicate Spectra**
**40 K More Compounds than NIST 17**

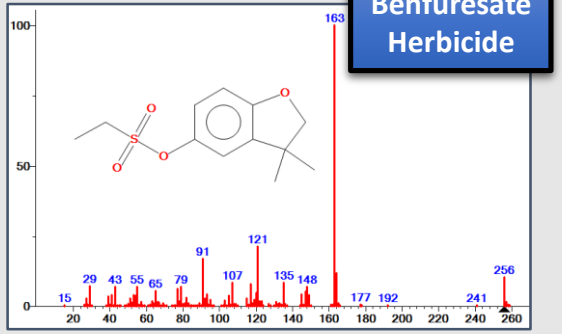
**Loliolide**
**Plant Metabolite**



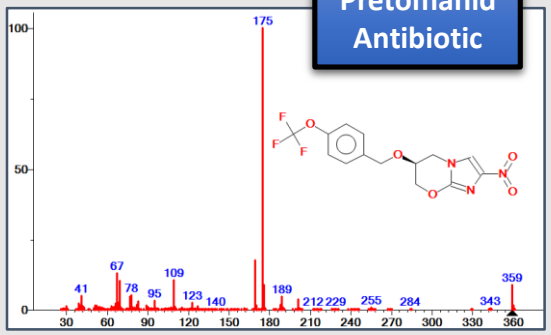
**UWA-101**
**Parkinson's drug**



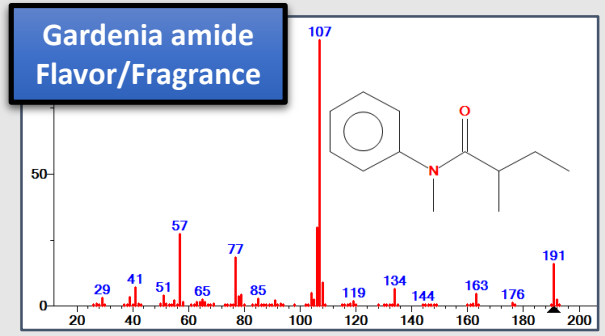
**Benfuresate**
**Herbicide**



**Pretomanid**
**Antibiotic**

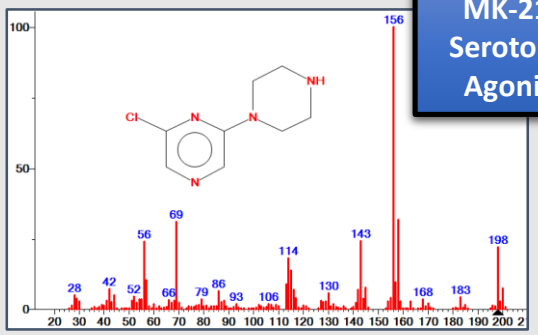


**Gardenia amide**
**Flavor/Fragrance**

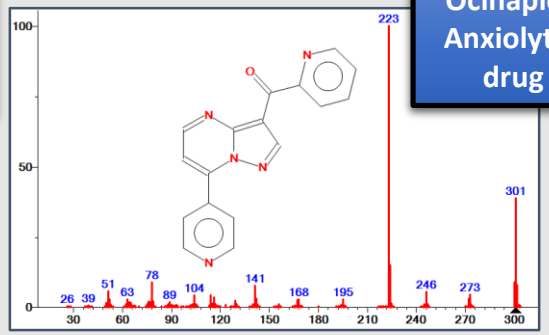


**Also TMS, Ac and Me Derivatives**

**MK-212**
**Serotonin Agonist**



**Ocinaplon**
**Anxiolytic drug**



**1,000s of New Compounds of Analytical Interest**

Human & Plant Metabolites
Flavor/Fragrance – Food
Drugs & their Metabolites
Forensics, Toxins
Pesticides –Contaminants
Industrial Chemicals
Petrochemicals,
Surfactants, Lipids, …

# NIST Tandem Mass Spectral Library

## 2020 Release

**31K Compounds, 2X More than 2017**
**186K Precursor Ions - 1.3M Spectra**

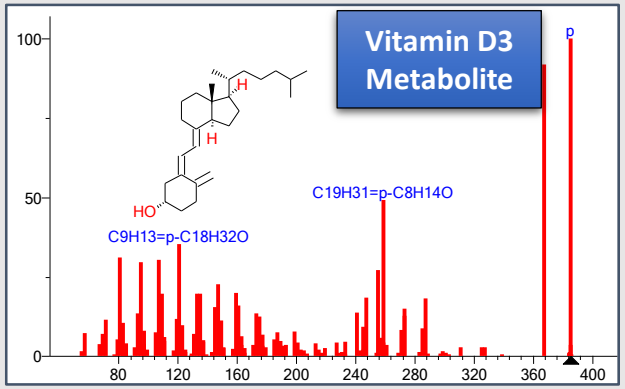
### Compounds: Fragmentation Methods

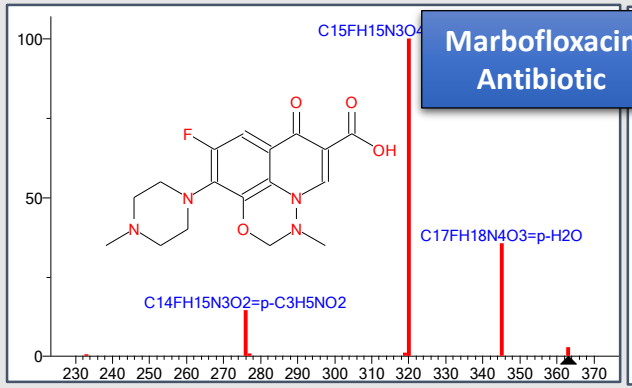
27,840 HRAM (High Res Accurate Mass)
29,890 QTOF, HCD, IT-HRAM, QqQ
29,444 Ion Trap (Low Res., up to $MS^4$)
246 APCI HRAM 'Extractables and Leachables'

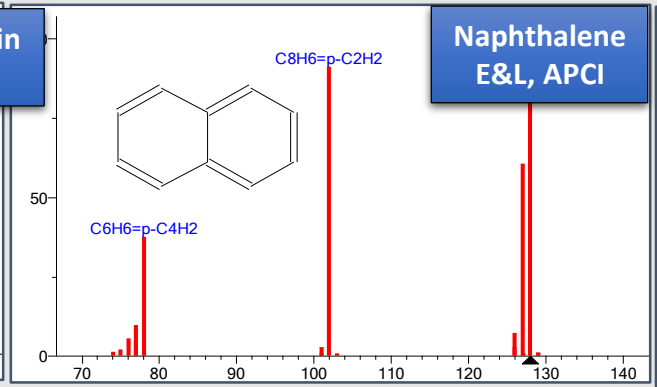
### Precursor Ion Types

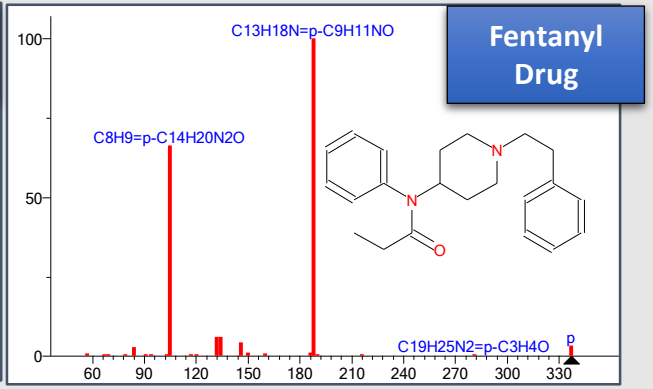
26,575 Protonated
12,589 Deprotonated
10,032 Water/Ammonia Loss
24,167 Other In-Source Generated

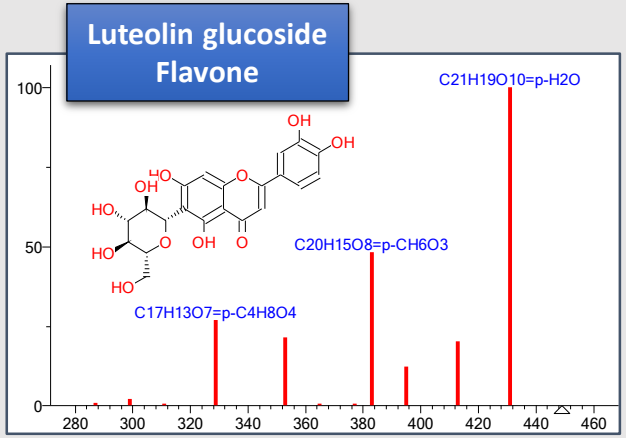
# 2020 Tandem Library
## Wide Range of Compounds

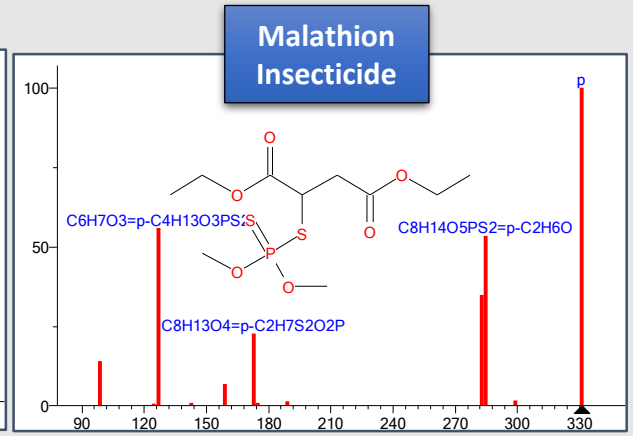


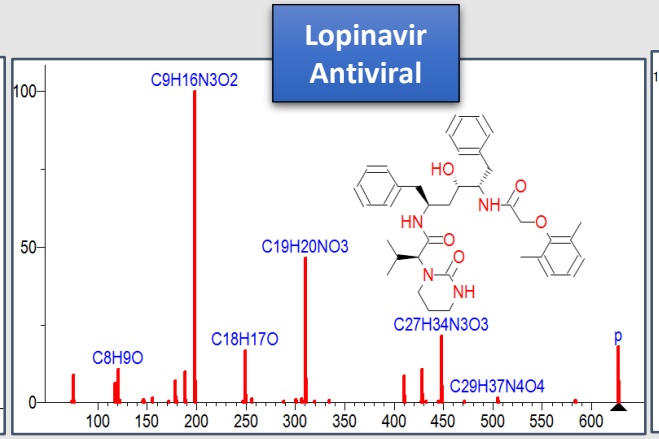



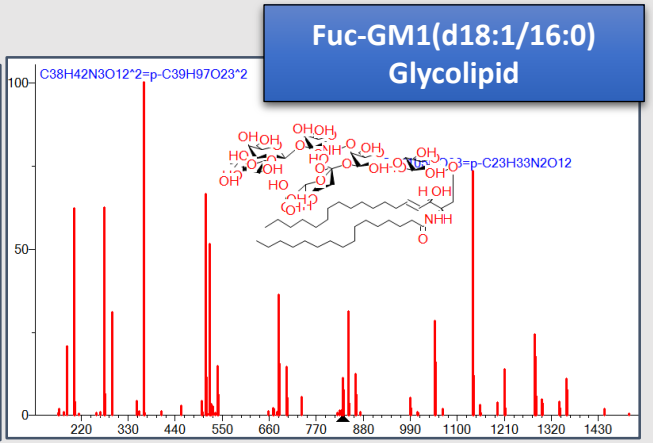



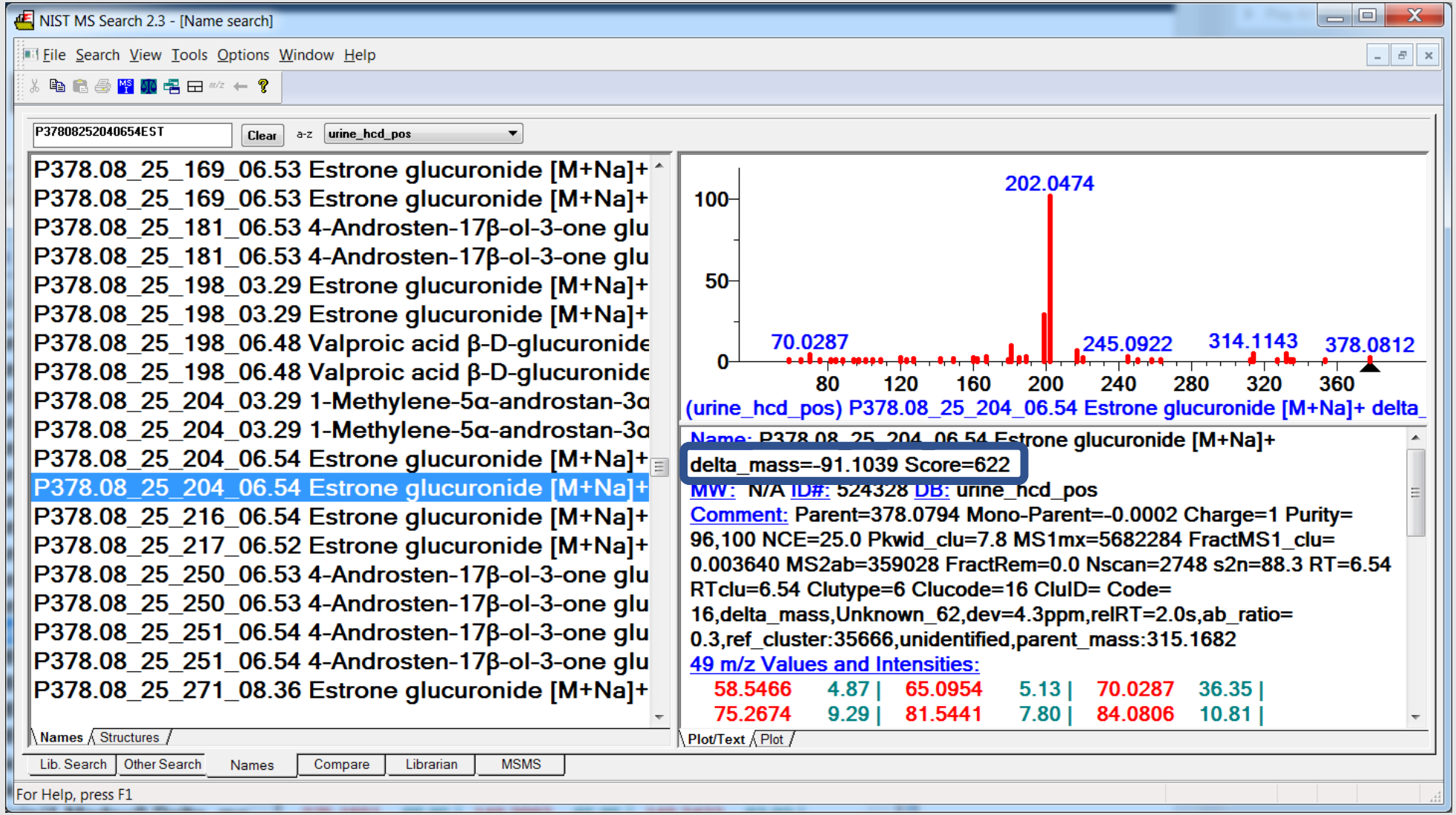



# Coverage:
# Recurring, Unidentified Spectra

- Annotated Recurrent Unidentified Spectra (ARUS)
  - Good quality, RUS converted to ‘consensus’ spectra
  - Annotated by Hybrid Search, then by Evaluator
  - LC/MS – Milk Oligos, Urine and Plasma/Serum, Acylcarnitines
  - GC/MS – RI in Essential Oils/Foods
- Available: *http://chemdata.NIST.gov/*
  - Documented in Multiple Papers
  - Urine/Plasma Updated for NIST 2020 Tandem Library

# LC/MS ARUS

## Hybrid Identified HRAM Spectra




From Multiple NIST Standard Reference Materials

Plasma/Serum, Urine, Milk Pos/Neg, *Chemdata.NIST.gov*

*Simon et al. TP 392*

# Peptide Library – 2020 Version

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Hs | Human | Ion Trap | H. sapiens | 05-29-2014 | 340,357 | 207,910 | Download |
| | | Orbitrap - HCD | H. sapiens | 05-19-2020 | 911,783 3 libraries | 605,677 | Download |

**85% increase in peptide ions**
**489,921 (2016) -> 911,783**

**25% to 36% proteome coverage**

**Quality – 7 quality filters**
**1 spectrum/peptide ion**

**Peptide.NIST.gov**

***ThP 291: Sheetlin et al.***

# Thank You!