Version 2.0

Team MetumpX, 2017-2018
Bilal Wajid
Hasan Iqbal
Momina Jamil

MetumpX 2.0 User Guide

CONTENTS

I. Introduction

Metabolomics is the comprehensive (qualitative and quantitative) study of all the small molecules in an organism called **metabolites**. These molecules are smaller than 1500 daltons (Da).Some important metabolites are peptides, oligonucleotides, sugars organic acids ketones, aldehydes, amines and amino acids.

A range of analytical techniques are used to analyze metabolites in different organisms, tissues, or fluids. There are two approaches used in metabolomics:

1 Untargeted Metabolomics

It involves measurement of as many metabolites as possible from a biological sample to classify phenotypes based on metabolite pattern. It is also known as Metabolic fingerprinting.

2 Targeted Metabolomics

It involves measurement of metabolites of a focused group from a biological sample. It is also known as Metabolic profiling.

There are several techniques for extraction of metabolic information from samples such as Solid Phase Extraction, Chromatography, Mass Spectrometry (MS) and Nuclear Magnetic Resonance (NMR).

This paper presents **MetumpX**, a software package pertaining to **MS-based Untargeted Metabolomics**.

II. Metabolomics Workflow

Metabolomics Workflow can be divided into many steps as shown in Fig. 1. Details of each step is as follows:

1 Sample Acquisition

Metabolites are measured in different samples such as tissue, biofluids (blood, urine, feces, seminal fluid, saliva, bile, cerebrospinal fluid) and cell cultures. There are numerous laboratory methods of sample preparation.

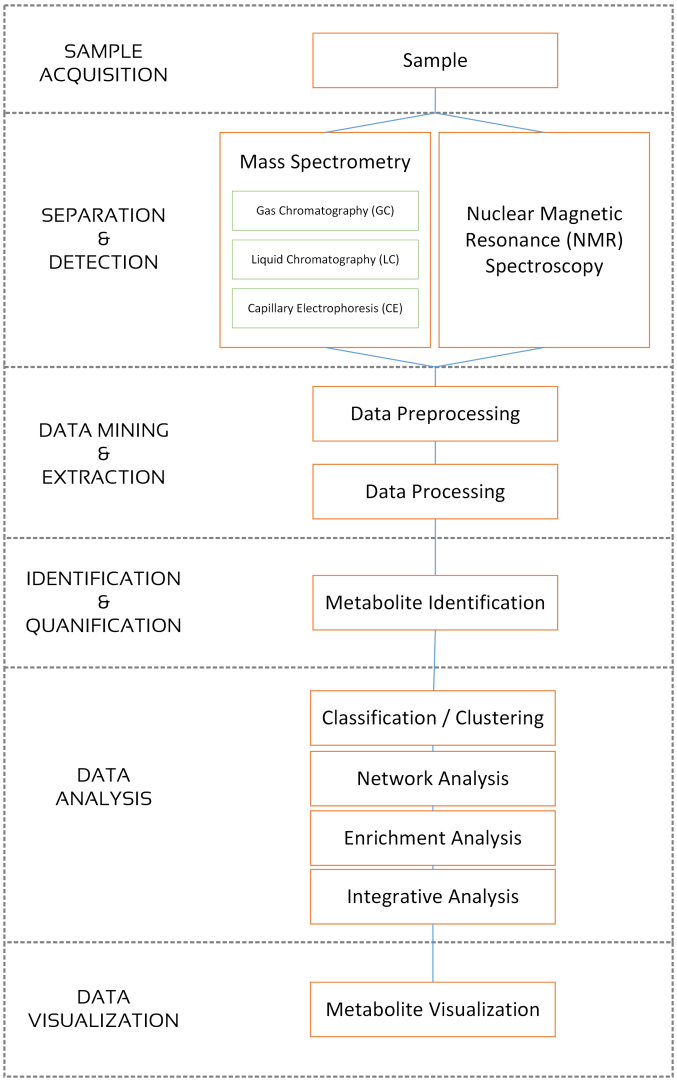

Fig. 1: Metabolomics Analytical Workflow

2 Separation and Detection of metabolites

The two main analytical methods for separation and detection are Mass Spectrometry (MS) and Nuclear Magnetic Resonance (NMR) Spectroscopy.

*1) **Nuclear Magnetic Resonance (NMR) Spectroscopy**:* In NMR spectroscopy, biological sample is placed in a magnetic field. Isotopes in the sample absorb the radiation and resonate at frequencies relative to the size of the molecules. The resultant spectrum shown in Fig. 2 is a collection of peaks at different positions and intensities and each sample has a unique pattern.

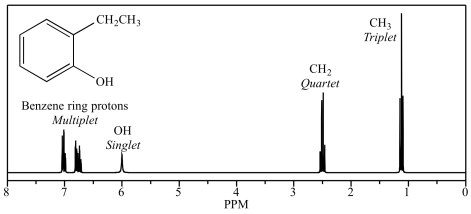

Fig. 2: NMR spectrum for 2-ethyl-phenol

*2) **Mass Spectrometry (MS)**:*

Mass Spectrometry can be used to analyze biological samples by ionizing the sample and then sorting the ions according to their mass-to-charge (m/z) ratio. Mass Spectrometer function is shown in Fig. 3. Significant methods for ionization of sample contain Chemical Ionization (CI), Electron Impact Ionization (EI) and Electrospray Ionization (ESI). A resulting MS Spectrum is shown in Fig. 4

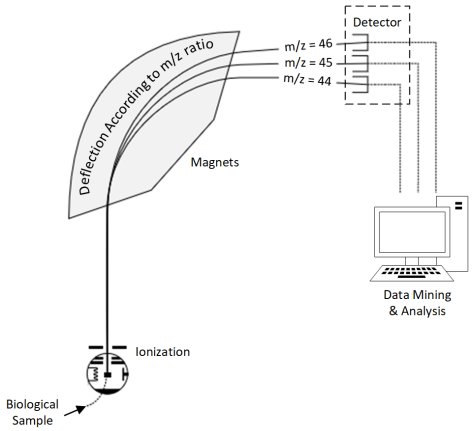

Fig. 3: Standard Mass Spectrometer

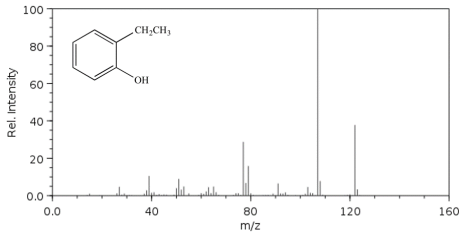

Fig. 4: MS spectrum for 2-ethyl-phenol

An important enhancement to the mass resolving and mass determining capabilities of mass spectrometry is using it in tandem with chromatographic and other separation techniques.

- **Gas Chromatography Mass Spectrometry (GC-MS)**:
  GC-MS is a commonly used in metabolomics for measurement of volatile compounds such as fatty acids and organic acids. Sample need to be volatile and thermally stable as separation in GC occurs in an oven at high temperatures. Metabolite loss is one of the major drawbacks of GC/MS. Characteristic spectral patterns and extensive libraries are available online for GC-MS. A Gas Chromatograph is shown in Fig. 5

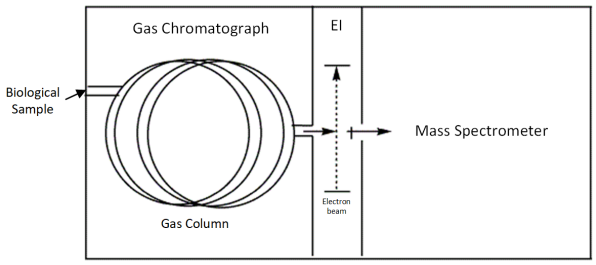

Fig. 5: Gas Chromatograph Mass Spectrometer

- **Liquid Chromatography Mass Spectrometry (LC-MS):**
  Liquid chromatography (LC) and High-Performance Liquid Chromatography (HPLC) is a technique that has high resolution and analytical flexibility. It can be used for the analysis of a specific metabolite or class of compounds. LCMS has one advantage over GCMS that there is no need for chemical derivatization of metabolites. A Liquid Chromatograph is shown in Fig. 6

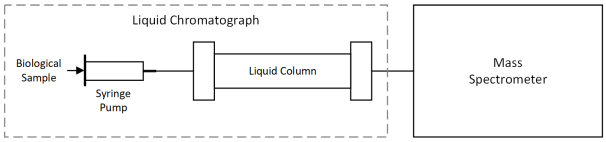

Fig. 6: Liquid Chromatograph Mass Spectrometer

- **Capillary Electrophoresis Mass Spectrometry (CE-MS):**
  CE-MS provides several advantages over other separation techniques such as high resolving power, very small sample requirement and short analysis time. One of the significant advantages of the CEMS is that it separates cations, anions and uncharged molecules in a single analytical run, and therefore CE can be used for simultaneous profiling of metabolites. Capillary Electrophoresis is shown in Fig. 7

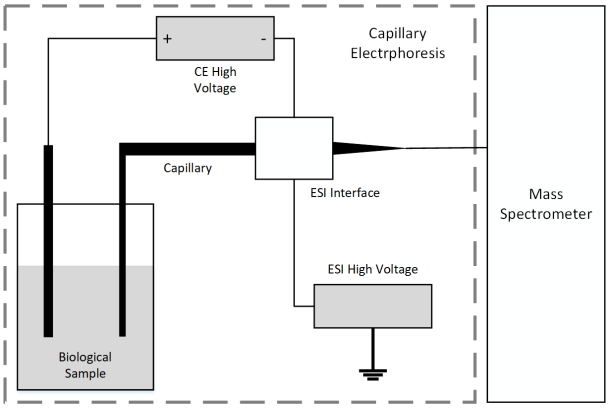

Fig. 7: Capillary Electrophoresis Mass Spectrometer

3 *Data Mining and Extraction*

Untargeted Metabolomic analysis generate large amount of complex data sets that require analysis by specialized software to properly interpret the data. Data is first preprocessed and then processed at this stage. There are several commercial and free software packages available to automate this process.

*1) **Data Preprocessing**:*

Data Preprocessing is divided into several sub-steps as follows:

- **Noise Filtering:**
  Noise filtering is the process of removing noise from a signal which facilitates further peak detection. It is an optional stage in data processing and can also be left out if the data is not noisy.
- **Chromatogram Alignment:**
  Chromatogram alignment is the process of aligning retention time for chromatographic methods with the mass spectrometers.
- **Peak Alignment:**
  Peak alignment is a process of correction of samples to point to the same metabolite or component. It is important in metabolomic studies as there is always a difference in the samples due to machine drift.
- **Spectral Deconvolution:**
  For GC-MS data, deconvolution is the process of computationally separating co-eluting components and creating a pure spectrum for each component. Deconvolution calculates the contribution of each component from an Extracted Ion Chromatogram that results from two or more components.
- **Retention Time Correction:**
  A process of removals of errors in the retention time due to temperature, polarity etc.

*2) **Data Processing**:*

Data Processing is divided into several substeps as follows:

- **Peak Detection:**
  Peak Detection is a process of identification and quantification of the features present in the spectra acquired from the spectrometer.
- **Data Normalization:**
  The total sample amount or concentration of metabolites in metabolomic workflow can be significantly different from one sample to another in each step. Data Normalization is the reduction or elimination of the effect of this variation.
- **Statistical Analysis:**
  After pre-processing, the LC-MS raw data is summarized by a peak list. The Statistical Analysis aims to detect those peaks whose intensity levels are significantly altered between distinct biological groups.
- **Quality Control:**
  Quality assurance and quality control provides a mechanism to ensure that a scientific process meets the predefined criteria. Quality control (QC) sample should qualitatively and quantitatively represent the entire collection of samples included in the study, providing an average of all of the metabolomes analysed in the study.
- **Metabolite Quantification:**
  Metabolite Quantification is method which evaluate changes in metabolic activity in response to disease, treatment, environmental and genetic perturbations.
- **Data Imputation:**
  Data Imputation is the process of handling the missing values in mass spectrometery. Typically, there are three types of missing values, missing not at random (MNAR), missing at random (MAR), and missing completely at random (MCAR).
- **In silico fragmentation:**
  in silico fragmentation in used to identify unknown compounds outside the database domain by comparing theoretical and experimental data.

4 Metabolite Identification

Metabolite Identification is the main step in which the metabolites are identified, profiled and compared with the chemical database.

5 Data Analysis

Several types of Data Analysis is done as follows:

*1) **Classification and Clustering Analysis**:*

Clustering is a well-established technique in the which samples are grouped and visualized according to intrinsic similarities in their measurements, irrespective of sample groupings.

*2) **Network Analysis**:*

Network Analysis can be further divided into three types:

- **Metabolic Network Analysis:**
  Metabolite datasets are combined through clustering analysis and represented as connection networks between genes and metabolites. The process of analyzing these datasets in called Metabolic Network Analysis.
- **metabolite Genome-wide Association Study (mGWAS):**
  Genome-wide association studies with metabolic traits (mGWAS) investigate how genetic variation effect metabolic phenotypes specially metabolism and complex disease. It maps the loci, position of a genes on a chromosome, responsible for natural variations in a target phenotype.
- **Metabolite Mapping:**
  Metabolite mapping is a process of Integration of biochemical pathway and chemical relationships to map all detected metabolites in network graphs

*3) **Enrichment analysis**:*

Metabolite Enrichment Analysis is statistical analysis of metabolite annotations and/or associated quantitative data.

*4) **Integrative analysis**:*

Integrative Analysis is process of linking metabolite data with other types of data (e.g. transcriptomics, proteomics), and incorporating prior knowledge of pathways and molecular interactions.

6 Data Visualization

Visualization of the metabolic spectrum. These are visual graphs with peaks at certain points.

III. METHODS

The authors of MetumpX researched for all the software tools that are available for the MS-based Untargeted Metabolomics and applied a specific criteria is used for enclosure of tools in MetumpX. Softwares which are included are due to these specific reasons:

- Publication Date (later then 2012)
- Freeware License
- Linux based
- Offline

Table I shows the selection criterion.

TABLE I: MetumpX Selection Criteria

Sr. no	Software Tool Name	Pub. Date (2012-)	Offline	Linux based	Free	Installed
1	ProbMetab[1]	✓	✓	✓	✓	✓
2	intCor[2]	✓	✓	✓	✓	✓
3	CorrectOverloadedPeaks[3]	✓	✓	✓	✓	✓
4	iMet-Q[4]		✓	✓	✓	
5	AnalyzerPro[5]	✓	✓	✓		
6	ADAP-GC[6]		✓	✓	✓	
7	ChromGenius[7]	✓	✓	✓		
8	ChromA[8]	✓		✓	✓	
9	X13CMS[9]	✓	✓	✓	✓	✓
10	MET-COFEA[10]	✓	✓		✓	
11	MET-XAlign[11]	✓	✓		✓	
12	batchCorr[12]	✓	✓	✓	✓	✓
13	MZMine[13]	✓	✓	✓	✓	✓
14	MET-IDEA[14]	✓	✓		✓	
15	PyMS[15]		✓	✓	✓	
16	MassCascade[16]		✓	✓	✓	
17	Mnova-MS[17]	✓	✓	✓		
18	XCMS[18]	✓	✓	✓	✓	✓
19	flagme[19]	✓	✓	✓	✓	✓
20	Elm. Metabolomics[20]	✓	✓		✓	
21	MetCirc[21]	✓	✓	✓	✓	✓
22	mzAccess[22]	✓	✓	✓		
23	SpeckTackle[23]		✓	✓	✓	
24	COMSPARI[24]		✓	✓	✓	
25	DAVE[25]	✓		✓	✓	
26	TargetSearch[26]		✓	✓	✓	
27	HCor[27]	✓	✓	✓	✓	✓
28	MS-DIAL[28]	✓	✓		✓	
29	MaxEnt[29]		✓	✓	✓	
30	RANSY/RAMSY[30]	✓		✓	✓	
31	UniDec[31]	✓	✓		✓	
32	Metab[32]	✓	✓	✓	✓	✓
33	decoMS2[33]	✓	✓	✓	✓	✓
34	GAGdecon[34]	✓	✓	✓	✓	✓
35	TMTc+[35]	✓	✓	✓		
36	BernetAl2018[36]		✓	✓	✓	
37	BUPID-Top-Down[37]	✓		✓	✓	
38	PicaudEtAl2018[38]	✓		✓	✓	
39	msXpertSuite[39]	✓	✓	✓	✓	✓
40	proFIA[40]	✓	✓	✓	✓	✓
41	apLCMS[41]	✓	✓	✓	✓	✓
42	yamss[42]	✓	✓	✓	✓	✓
43	cosmiq[43]	✓	✓	✓	✓	✓
44	AMDIS[44]	✓	✓		✓	
45	mzMatch-ISO[45]	✓	✓	✓	✓	✓
46	Elgen-MS[46]		✓	✓	✓	
47	MetaQuant[47]		✓	✓	✓	
48	CAMERA[48]	✓	✓	✓	✓	✓
49	MS-FLO[49]	✓		✓	✓	
50	JumPm[50]	✓	✓	✓		
51	FastChrom[51]	✓	✓		✓	
52	GridMass[52]		✓	✓	✓	
53	KMMDA[53]	✓	✓	✓	✓	✓
54	HayStack[54]	✓		✓	✓	
55	msPeak[55]		✓	✓	✓	
56	GC-Aanlyzer[56]	✓	✓	✓		
57	MsXelerator[57]	✓	✓	✓		
58	MarkerLynx[58]	✓	✓	✓		
59	GCxGCAnalyzer[59]	✓	✓	✓		
60	MetNorm[60]	✓	✓	✓	✓	✓
61	MetTailor[61]	✓	✓	✓	✓	✓
62	MetaPre[62]	✓		✓	✓	

63	NOREVA[63]	✓		✓	✓	
64	crmn[64]		✓	✓	✓	
65	Normalizer[65]	✓		✓	✓	
66	metaX[66]	✓		✓	✓	
67	MetabR[67]		✓	✓	✓	
68	LowessNormalization[68]	✓	✓		✓	
69	MSPrep[69]	✓	✓	✓	✓	✓
70	AIoutput [70]	✓	✓		✓	
71	Ionwinze[71]	✓	✓		✓	
72	MPP [72]	✓	✓	✓		
73	R2DGC[73]	✓	✓	✓	✓	✓
74	mSPA[74]	✓	✓	✓	✓	✓
75	Maui-VIA[75]	✓	✓		✓	
76	SECIMTools[76]	✓	✓	✓	✓	✓
77	BatMass[77]	✓	✓	✓	✓	✓
78	MetaboQC[78]	✓	✓	✓		
79	QCScreen[79]	✓	✓	✓	✓	✓
80	QC-RFSC[80]		✓	✓	✓	
81	mscompare[81]	=	✓	✓	✓	
82	PYQUAN[82]	✓	✓		✓	
83	MINMA[83]	✓	✓	✓	✓	✓
84	MetaboloDerivatizer[84]	✓	✓		✓	
85	SIRIUS[85]	✓	✓	✓	✓	✓
86	HAMMER[86]	✓	✓		✓	
87	ISDB-MN[87]		✓	✓	✓	
88	SweetSubstitute [88]	✓	✓		✓	
89	MetFrag[89]	✓		✓	✓	
90	MetFusion[90]	✓		✓	✓	
91	iontree[91]	✓	✓	✓	✓	✓
92	ACD/MS Fragmenter[92]		✓	✓	✓	
93	MassFrontier[93]	✓	✓	✓		
94	MWASTools[94]	✓	✓	✓	✓	✓
95	RegScan[95]		✓	✓	✓	
96	InCroMAP[96]	✓	✓	✓	✓	✓
97	PathVisio[97]	✓	✓	✓	✓	✓
98	CHem-SMP[98]	✓	✓	✓		
99	cPath[99]		✓	✓	✓	
100	MetaMapp[100]		✓	✓	✓	
101	Mapping Tool[101]	✓	✓	✓	✓	✓
102	BLASTX [102]	✓		✓	✓	
103	PSSAlib [103]		✓	✓	✓	
104	iPath [104]	✓		✓	✓	
105	MetExplore [105]	✓		✓	✓	
106	CATABOL[106]		✓	✓	✓	
107	Paintomics[107]	✓		✓	✓	
108	ProtienLounge[108]	✓		✓	✓	
109	CellMLTools[109]		✓	✓	✓	
110	FCF[110]		✓	✓	✓	
111	PathPred[111]	✓		✓	✓	
112	SABIO-RK[112]	✓		✓	✓	
113	OptCom[113]		✓	✓	✓	
114	Subpathway-GM[114]	✓		✓	✓	
115	IPAVS[115]	✓		✓	✓	
116	GAM[116]	✓		✓	✓	
117	GLAMM[117]	✓		✓	✓	
118	PASMet[118]	✓		✓	✓	
119	NICElips[119]		✓	✓	✓	
120	MetaMapR[120]	✓		✓	✓	
121	PAPi[121]	✓	✓	✓	✓	✓
122	ReactPRED[122]	✓	✓	✓	✓	✓
123	JigCell[123]		✓	✓	✓	
124	Cell++[124]		✓	✓	✓	
125	MetNetter[125]		✓	✓	✓	
126	MEMOSys[126]	✓		✓	✓	
127	MonaLisa[127]	✓	✓	✓	✓	✓
128	QSSPN[128]	✓	✓	✓	✓	✓
129	NetSeed[129]	✓		✓	✓	
130	tEFMA[130]		✓	✓	✓	
131	NetCmpt[131]	✓		✓	✓	
132	SED-ED [132]	✓	✓	✓	✓	✓
133	MetNetMaker [133]	✓	✓		✓	
134	RxnSim [134]	✓	✓	✓	✓	✓
135	MetExploreViz[135]	✓		✓	✓	
136	CorrelationCalculator[136]		✓	✓	✓	
137	WebMetabase [137]	✓		✓	✓	
138	MoDitify[138]	✓	✓	✓	✓	✓
139	MetaboSignal[139]	✓	✓	✓	✓	✓
140	JMassBalance[140]	✓	✓	✓	✓	✓
141	MetaDiff[141]	✓	✓	✓	✓	✓
142	FTA [142]		✓	✓	✓	
143	PySCeSToolbox[143]	✓	✓	✓	✓	✓
144	MEBS[144]	✓	✓	✓	✓	✓
145	Subpathway-GMir[145]	✓	✓	✓	✓	✓
146	Prol[146]	✓		✓	✓	
147	MetaNetSam[147]	✓	✓	✓	✓	✓
148	lumpGEM [148]		✓	✓	✓	
149	redGEM[149]		✓	✓	✓	
150	CCC[150]	✓	✓	✓		
151	HuEtAl2018[151]	✓	✓	✓		
152	Kamneva[152]	✓	✓	✓	✓	✓
153	XeDetect[153]		✓	✓	✓	
154	ReactomePA[154]	✓	✓	✓	✓	✓
155	IPPAD[155]	✓		✓	✓	
156	MMinte[156]	✓	✓	✓	✓	✓
157	SED-ML[157]		✓	✓	✓	
158	NetCooperate[158]	✓		✓	✓	
159	PathRings[159]	✓		✓	✓	
160	IntPath [160]	✓		✓	✓	
161	EvoMS[161]	✓		✓	✓	
162	phraSED-ML[162]	✓	✓	✓	✓	✓
163	CARMEN[163]	✓		✓	✓	
164	ScrumPy[164]	✓	✓	✓	✓	✓
165	Pybrn[165]	✓	✓	✓	✓	✓
166	Pyabolism[166]	✓	✓	✓	✓	✓
167	KEGGREST[167]		✓	✓	✓	
168	VIENNA-RNL[168]		✓	✓	✓	
169	Fbar[169]	✓	✓	✓	✓	✓
170	MetaCore[170]	✓		✓	✓	
171	PathwayLab[171]	✓	✓	✓		
172	MetScape[172]	✓	✓	✓	✓	✓
173	MPEA[173]	✓		✓	✓	
174	IMPaLA[174]	✓		✓	✓	
175	MBRole[175]	✓		✓	✓	
176	zeroSum[176]	✓	✓	✓	✓	✓
177	ChemRICH[177]	✓		✓	✓	
178	FELLA[178]	✓	✓	✓	✓	✓
179	BinChE[179]	✓		✓	✓	
180	MetaboliteIDConvertor[180]	✓	✓	✓	✓	✓
181	MetaboAnalyst[181]	✓	✓	✓	✓	✓
182	MapMan[182]		✓	✓	✓	
183	3Omics[183]	✓		✓	✓	
184	integrOmics[184]	✓	✓	✓	✓	✓
185	MetDisease[185]	✓	✓	✓	✓	✓
186	MetaBridge[186]	✓		✓	✓	
187	MetMask[187]		✓	✓	✓	
188	ProMeTra [188]		✓	✓	✓	
189	MAGI[189]	✓		✓	✓	
190	KPIC2[190]	✓	✓	✓	✓	✓
191	MarVis-Suite[191]	✓	✓		✓	
192	MSClust[192]		✓	✓	✓	
193	MetMSLine[193]	✓	✓	✓	✓	✓
194	SimExTargid [194]	✓	✓	✓	✓	✓
195	MetaboliteDetector [195]		✓	✓	✓	
196	specmine[196]	✓	✓	✓	✓	✓
197	W4M [197]	✓		✓	✓	
198	MeltDB [198]	✓	✓		✓	
199	xMSAnalyzer [199]	✓	✓	✓	✓	✓
200	ChromaTOF[200]	✓	✓	✓		
201	MetabolmeExpress[201]	✓		✓	✓	
202	Metabox [202]	✓	✓	✓	✓	✓
203	PiMP[203]	✓		✓	✓	
204	MET-COFEI[204]	✓	✓		✓	
205	MAIT [205]	✓	✓	✓	✓	✓
206	BinVestigate[206]	✓		✓	✓	
207	CEU Mass Mediator [207]	✓		✓	✓	
208	MAGMa[208]	✓		✓	✓	
209	CSI:FingerID [209]	✓		✓	✓	
210	MS2LDA SUPPORT [210]	✓		✓	✓	
211	MetExtract [211]		✓	✓	✓	
212	T-BioInfo [212]	✓		✓	✓	
213	MetAlign[213]	✓	✓		✓	
214	CFM-ID [214]	✓		✓	✓	
215	Ideom [215]		✓	✓	✓	
216	AStream [216]		✓	✓	✓	
217	PUTMEDID-LCMS[217]		✓	✓	✓	
218	DECOMP [218]		✓	✓	✓	
219	MetiTree[219]	✓		✓	✓	
220	MIA [220]	✓	✓	✓	✓	✓
221	MFSearcher[221]	✓	✓		✓	
222	ChemDistiller[222]		✓	✓	✓	
223	MSeasy[223]	✓	✓	✓	✓	✓
224	SIMPLE[224]		✓	✓	✓	
225	MAVEN[225]		✓	✓	✓	
226	SpectConnect [226]	✓		✓	✓	
227	RAMClustR[227]	✓	✓	✓	✓	✓
228	Molfind[228]	✓	✓	✓	✓	✓
229	MS2Analyzer [229]	✓	✓	✓	✓	✓
230	MS-FINDER [230]	✓	✓		✓	
231	geoRge[231]	✓	✓	✓	✓	✓
232	MetFamily[232]	✓		✓	✓	
233	eRAH[233]	✓	✓	✓	✓	✓
234	IsoMS [234]	✓	✓		✓	
235	MetDIA [235]		✓	✓	✓	
236	iMET[236]	✓		✓	✓	
237	MIDAS[237]		✓	✓	✓	
238	InterpretMSSpectrum [238]	✓	✓	✓	✓	✓
239	AssayR[239]	✓	✓	✓	✓	✓
240	MCID[240]	✓		✓	✓	
241	compMS2Miner[241]	✓		✓	✓	
242	MetShot [242]	✓	✓	✓	✓	✓
243	MINE[243]	✓		✓	✓	
244	NP-StructurePredictor [244]	✓	✓	✓		
245	MetaboSearch[245]	✓	✓	✓	✓	✓
246	ALLocator[246]	✓		✓	✓	
247	PROFANCY[247]	✓	✓		✓	
248	SpiderMass[248]	✓	✓		✓	
249	MZedDB[249]	✓		✓	✓	
250	BinBase [250]		✓	✓	✓	
251	PowerGet[251]	✓	✓		✓	
252	AMDORAP [252]		✓	✓	✓	
253	MBIdent [253]		✓	✓	✓	
254	peakANOVA[254]		✓	✓	✓	
255	MI-Pack[255]		✓	✓	✓	
256	SetupX [256]		✓	✓	✓	

257	FeatureFinderMetabo [257]		✓	✓	✓	
258	MassMetaSite[258]		✓	✓	✓	
259	MetDNA[259]	✓		✓	✓	
260	DASI[260]	✓	✓	✓		
261	MetaboList[261]	✓	✓	✓	✓	✓
262	MetaMS[262]	✓	✓	✓	✓	✓
263	SIEVE[263]	✓	✓	✓		
264	SimMet[264]	✓		✓	✓	
265	Apex[265]	✓	✓	✓		
266	Nontarget[266]	✓	✓	✓	✓	✓
267	NIST MS Search[267]	✓	✓		✓	
268	El-Maven [268]	✓	✓	✓	✓	✓

The authors employed ”Zenity” as the front-end GUI for users so that they have an easy and user friendly interface for installation of eighty-eight tools. However, linux bash scripting was employed to carefully handle each installation including necessary dependencies to ensure the users are able to develop workstations quickly and easily. The dependencies which are necessary for the installation of other tools are mentioned in Table II along with their sizes. These dependencies are automatically installed during the execution and Installation of MetumpX.

TABLE II: MetumpX Selection Criteria

Sr. no	Dependencies	Size (MB)
1	Oracle Java JDK	177
2	Libraries	13.8
3	bwidget	167
4	itcl	48.5
5	tk	1.3
6	tdom	0.2
7	tcllib	4
8	Tclx	0.05
9	tkimg	0.1
10	rxvt-uicode	3.4
11	sdcc	3.5
12	doxygen	15.9
13	hunspell	0.04
14	pandas	8.5
15	scipy	31.2
16	numpy	13.9
17	pyomo	1.8
18	decorator	0.01
19	pytz	0.5
20	six	0.01
21	PyUtilib	0.2
22	appdirs	0.01
23	ply	0.05
24	nose	0.1
25	cython	3.4
26	h5py	2.8
27	jupyter client	0.1
28	qtconsole	0.1
29	ipywidgets	0.1
30	pygments	0.8
31	nbconvert	0.2
32	notebook	8.9
33	ipykernel	0.1
34	traitlets	0.1
35	jupyter core	0.1
36	ipython	0.7
37	widgetsnbextension	2.2
38	nbformat	0.2
39	prompt toolkit	0.2
40	jinja2	0.1
41	testpath	0.1
42	pyzmq	1.0
43	tornado	0.5
44	pexpect	0.1
45	jedi	0.2
46	html5lib	0.1
47	parso	0.1
48	psutil	0.5
49	matplotlib	0.01
50	pyparsing	0.06
51	kiwisolver	1.0
52	binutils	2.3
53	gcc-avr	14.1
54	avr-libc	4.3
55	uisp	0.5
56	avrdude	0.3
57	flex	0.3
58	byacc	0.5
59	bison	0.5

A Software Categorization Scheme is shown in Table III which shows the the category of each software along with the sizes and versions of each tool included in MetumpX.

One software can be multi-functional and is listed in more then one category but authors have recommended a specific category in which the tools works most efficiently. This categorization is further shown in Fig.8

TABLE III: MetumpX Software Categorization Scheme

Sr. No.	Software	Noise Filtering	Chromatogram Alignment	Peak Alignment	Retention Time Correction	Spectral Deconvolution	Peak Detection	Data Normalization	Statistical Analysis	Quality Control	Metabolite Quantification	Data Imputation	In-silico Fragmentation	Metabolite Identification	Spectral Visualization	Clustering Analysis	mGWAS	Mapping	Network Analysis	Enrichment Analysis	Integrative Analysis	Size (MB)	Version
1	**CorrectOverloadedPeaks**[3]	✓																				4.8	
2	specmine[196]	✓					✓	✓	✓			✓		✓								15.2	
3	intCor[2]	✓	✓	✓																		3.0	
4	**batchCorr**[12]		✓																			17.2	
5	**mSPA**[74]			✓																		0.1	
6	**Metab**[32]					✓																3.2	
7	decoMS2[28]					✓								✓								5.2	
8	GAGdecon[34]					✓																0.1	
9	msXpertSuite[39]					✓																0.7	
10	**HCor**[27]				✓																	0.1	
11	MetMSLine[193]	✓	✓				✓								✓							1.6	
12	**X13CMS**X13CMS[9]		✓				✓															0.1	
13	proFIA[40]	✓					✓								✓							2.0	
14	cosmiq[43]		✓				✓				✓											17.5	
15	mzMatch-ISO[45]						✓							✓								0.1	
16	Metabox							✓	✓										✓	✓	✓	53.0	
17	**MetNorm**[60]	✓						✓														0.5	
18	**KMMDA**[53]						✓		✓													0.5	
19	MSPrep[69]							✓	✓			✓										1.0	
20	flagme[19]		✓				✓		✓		✓				✓							22.0	
21	SECIMTools[76]							✓	✓	✓												0.6	
22	**QCScreen**[79]									✓												12.1	
23	**MetTailor**[61]		✓								✓											0.9	
24	MetaQuant[47]						✓	✓			✓			✓								12.0	
25	apLCMS[41]		✓				✓				✓			✓								14.1	
26	**MINMA**[83]											✓										2.6	
27	**SIRIUS**[85]												✓									37.0	
28	iontree[91]												✓									0.9	
29	MetShot[242]													✓								0.9	
30	MolFind[228]													✓								36.2	
31	MIA[220]													✓								2.7	
32	MetaMS[262]													✓								3.7	
33	MSeasy[223]													✓								5.8	
34	RAMClustR[227]													✓								51.6	
35	MetaboSearch[245]													✓								45.7	
36	El-Maven[268]													✓								91.4	
37	**geoRge**[231]													✓								13.7	
38	eRAH[233]													✓								3.7	
39	MetaboLis[261]t													✓								0.3	
40	InterpretMSSpectrum[238]													✓								0.2	
41	AssayR[239]													✓								71.6	
42	MS2Analyzer[229]													✓								3.0	
43	Nontarget[266]													✓								3.5	
44	SimExTargid[194]	✓	✓	✓			✓	✓			✓			✓	✓	✓						49.8	
45	MetaboAnalyst[181]							✓	✓					✓		✓			✓		✓	49.3	
46	CAMERA[48]						✓							✓		✓						2.1	
47	KPIC2		✓				✓				✓					✓						15.7	
48	**MWASTools**[94]																✓					56.3	
49	**InCroMAP**[96]																	✓				54.6	
50	PathVisio[97]																	✓				17.6	
51	Mapping Tool[101]																	✓				15.0	
52	PycesToolbox[143]																		✓			5.1	
53	MetaDiff[141]																		✓			6.0	
54	ReactomePA																		✓			11.6	
55	MEBS[144]																		✓			73.2	
56	RxnSim [134]																		✓			21.6	
57	phraSED-ML[162]																		✓			1.8	
58	ScrumPy[164]																		✓			1093.9	
59	Subpathway-GMir[145]																		✓			1.7	
60	Kamneva[152]																		✓			101.8	
61	MetaboSignal[139]																		✓			169.1	
62	ReactPRED[122]																		✓			52.4	
63	Mminte																		✓			52.2	
64	**PAPi**[121]																		✓			0.6	
65	MetaNetSam[147]																		✓			3.8	
66	Fbar[169]																		✓			2.1	
67	SED-ED [132]																		✓			6.4	
68	QSSPN[128]																		✓			2.1	
69	JMassBalance[140]																		✓			3.7	
70	Pyabolism[166]																		✓			0.4	
71	Pybrn[165]																		✓			0.6	
72	MonaLisa[127]																		✓			17.8	
73	MoDitify[138]																		✓			0.6	
74	integrOmics[184]																				✓	0.1	
75	MetScape[172]																		✓	✓	✓	17.2	
76	MetDisease[185]																		✓		✓	15.2	
77	zeroSum[176]																			✓		1.0	
78	FELLA[178]																			✓		3.1	
79	MetaboliteIDConvertor[180]																			✓		0.2	
80	MZMine[13]	✓	✓			✓	✓	✓			✓			✓	✓							148.7	
81	XCMS[18]		✓	✓			✓	✓			✓			✓	✓							3.5	
82	yamss[42]		✓				✓															15.5	
83	R2DGC[73]		✓	✓			✓							✓	✓							0.6	
84	**BatMass**[77]									✓					✓							0.1	
85	MAIT		✓				✓		✓			✓		✓	✓				✓			36.2	
86	ProbMetab[1]	✓	✓				✓		✓					✓	✓							0.2	
87	MetCirc[21]			✓										✓	✓							5.0	
88	xMSAnalyzer[199]		✓			✓	✓			✓				✓	✓							38.9	

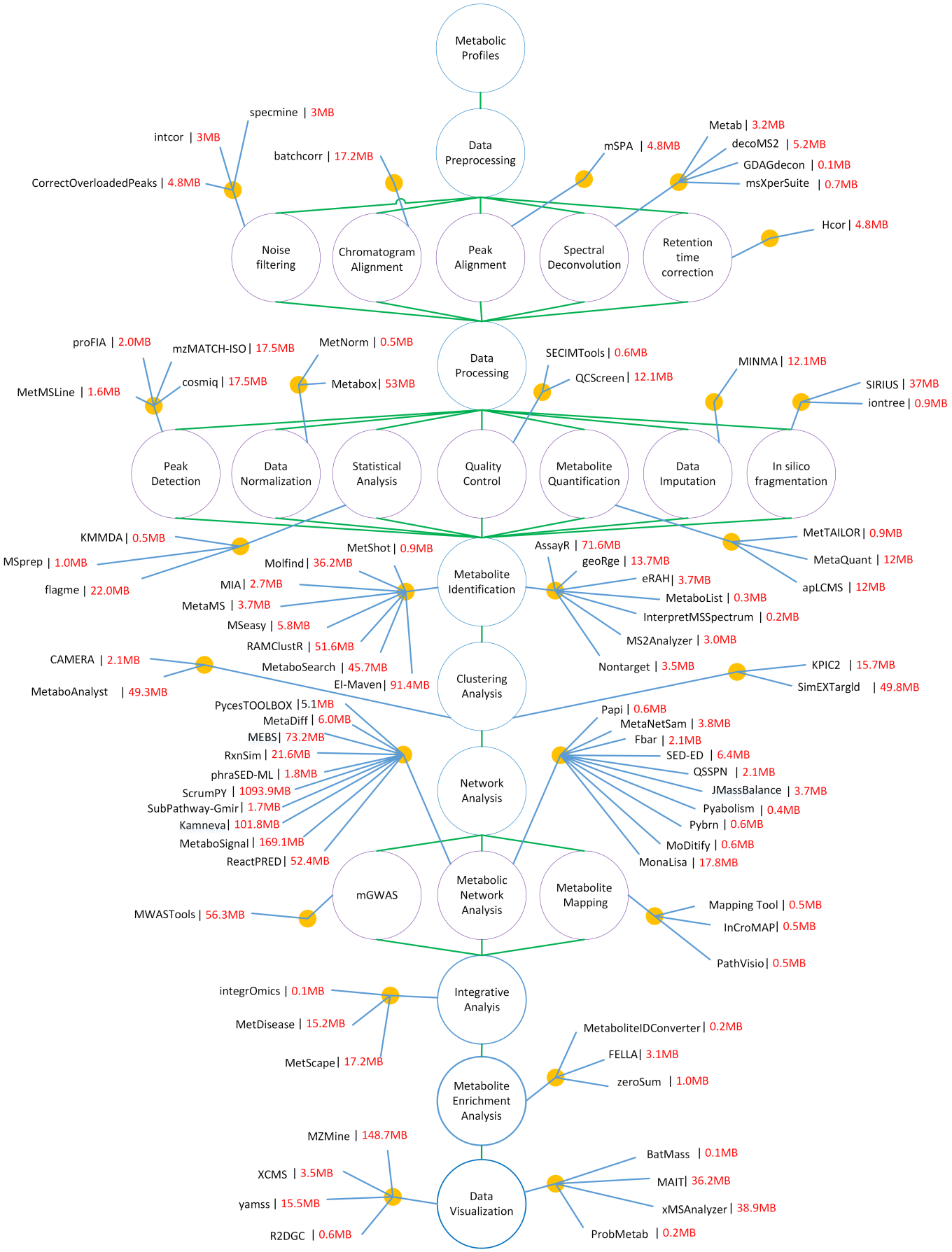

Fig. 8: The figure presents an overview of the Metabolomics software pipeline. The software tools packaged in MetumpX, along with their size are mentioned along each recommended category in the figure.

IV. Software Installation Process

For downloading of MetumpX, visit its website: https://github.com/hasaniqbal777/MetumpX

1) MetumpX can also be downloaded from the following command through git:

```
git clone <url>
```

2) Now run the following command on terminal:

```
chmod +x MetumpX_setup_enUS.sh
sudo bash MetumpX_setup_enUS.sh
```

3) Installation wizard of MetumpX will start
Note: Installation of MetumpX require a proper internet connection to proceed, otherwise the installation terminates.

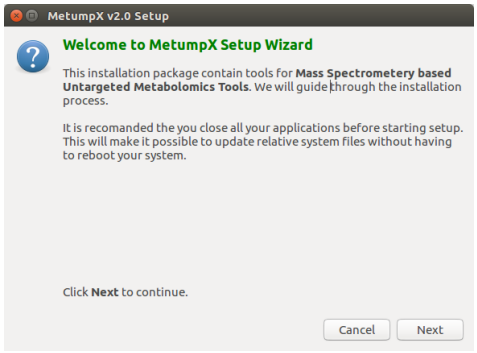

Fig. 9: Installation Welcome Screen

4) Click **Next** to proceed and confirm the internet connection.

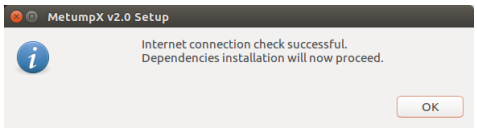

Fig. 10: Internet Check Successful Screen

5) Installation is **terminated** if there is no internet.

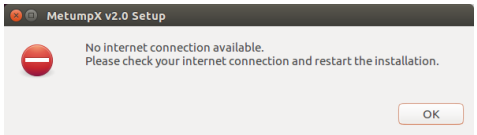

Fig. 11: Internet Check failed screen

6) **Dependencies** related to software will start installing.

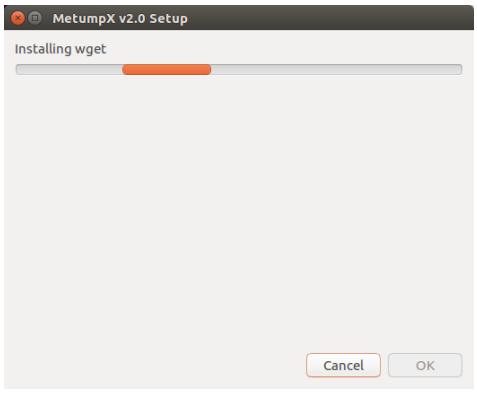

Fig. 12: Installation Screen

7) All the dependencies which are installed are displayed at the end of the installation. Click **Ok**.

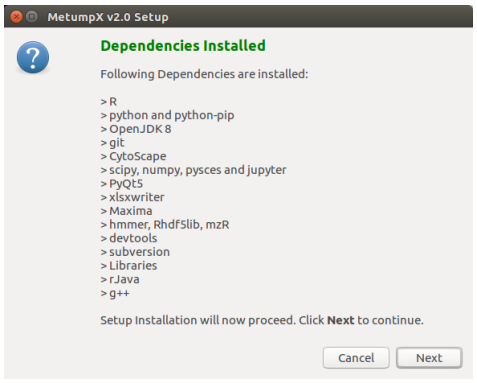

Fig. 13: Dependencies installation finished screen

8) Selection screen for Pre-processing Tools is displayed. Select the required tools and Click **Next**.

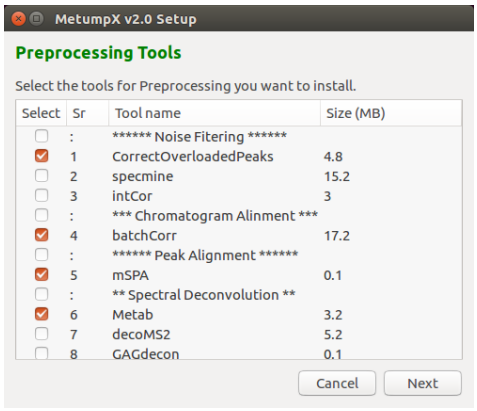

Fig. 14: Preprocessing tools selection screen

9) Selection screen for Processing Tools is displayed. Select the required tools and Click **Next**.

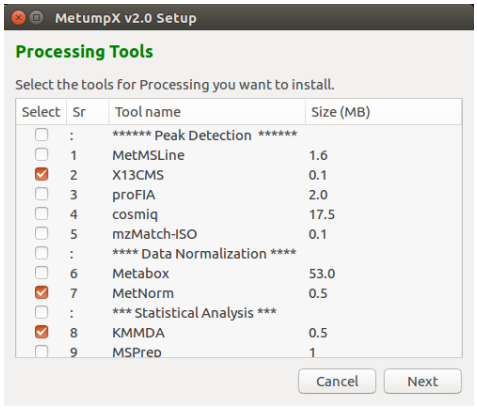

Fig. 15: Processing tools selection screen

10) Selection screen for Metabolite Identification tools is displayed. Select the required tools and Click **Next**.

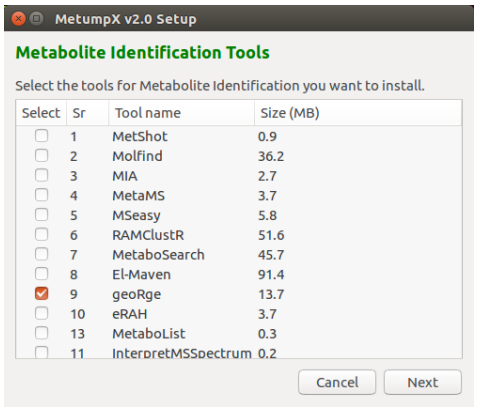

Fig. 16: Metabolite identification tools selection screen

11) Selection screen for Clustering Analysis tools is displayed. Select the required tools and Click **Next**.

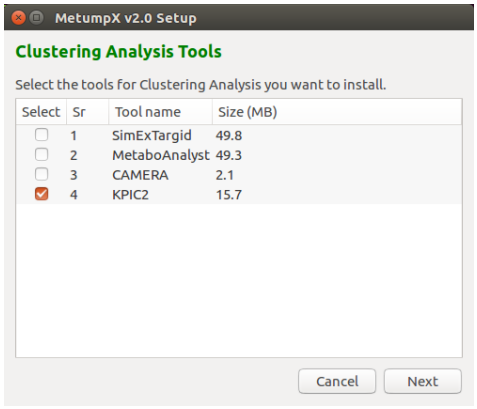

Fig. 17: Clustering analysis tools selection screen

12) Selection screen for Data Network Analysis tools is displayed. Select the required tools and Click **Next**.

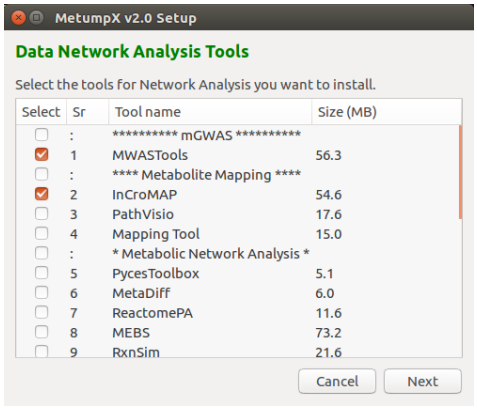

Fig. 18: Data network analysis tools selection screen

13) Selection screen for Enrichment analysis tools is displayed. Select the required tools and Click **Next**.

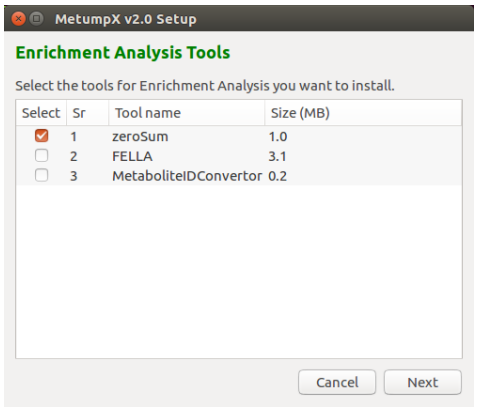

Fig. 19: Enrichment analysis tools selection screen

14) Selection screen for Integrative analysis tools is displayed. Select the required tools and Click **Next**.

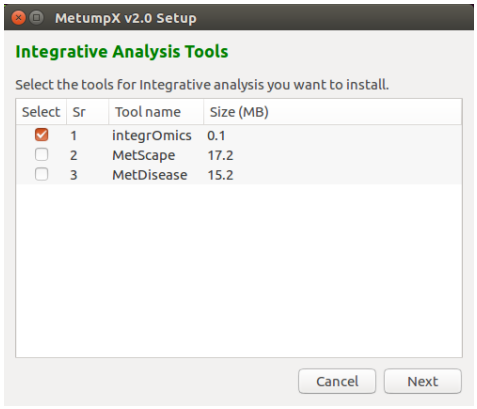

Fig. 20: Integrative analysis tools selection screen

15) Selection screen for Data visualization tools is displayed. Select the required tools and Click **Next**.

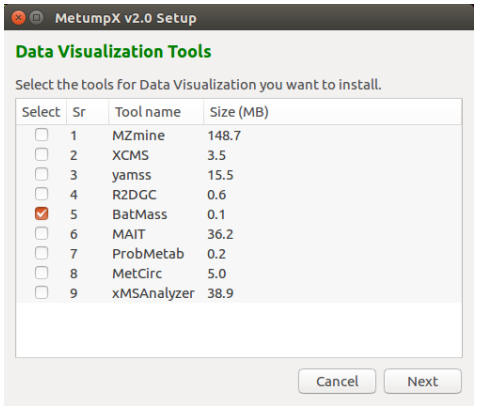

Fig. 21: Data visualization tools selection screen

16) Click **Next** to proceed with the installation.

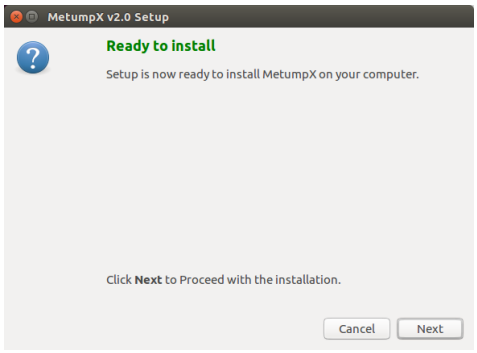

Fig. 22: Ready to Install Screen

17) Tool installation will **continue**. Some tools **install** as standalone installation and will be called automatically.

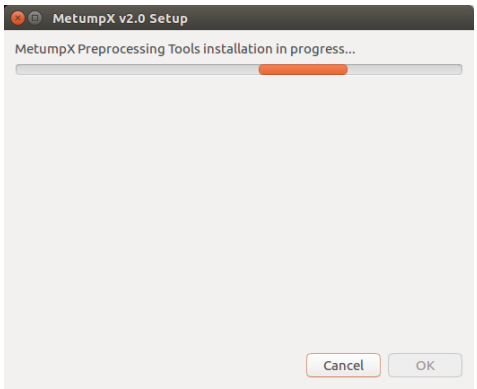

Fig. 23: Tools installation progress screen

18) Installation of MetumpX is now finished. Click **Finish** to use the tools.

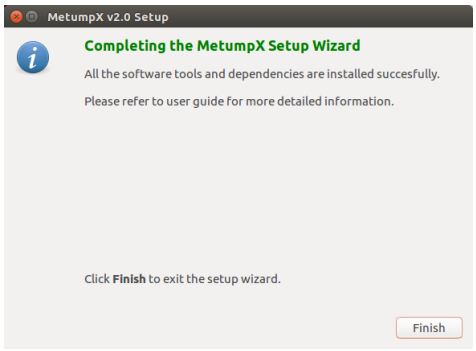

Fig. 24: Finishing Installation Screen

References

[1] R. R. Silva, F. Jourdan, D. M. Salvanha, F. Letisse, E. L. Jamin, S. Guidetti-Gonzalez, C. A. Labate, and R. Z. Vêncio, "Probmetab: an r package for bayesian probabilistic annotation of lc–ms-based metabolomics," *Bioinformatics*, vol. 30, no. 9, pp. 1336–1337, 2014.

[2] D. Ghosh and H. Choi, "metaarray package for meta-analysis of microarray data," *R package. http://www. r-project. org*, 2006.

[3] H. Poppe, "Overloading and interaction phenomena in electrophoretic separations," *Analytical Chemistry*, vol. 64, no. 17, pp. 1908–1919, 1992.

[4] H.-Y. Chang, C.-T. Chen, T. M. Lih, K.-S. Lynn, C.-G. Juo, W.-L. Hsu, and T.-Y. Sung, "imet-q: a user-friendly tool for label-free metabolomics quantitation using dynamic peak-width determination," *PloS one*, vol. 11, no. 1, p. e0146112, 2016.

[5] C. Baldari, V. Bonavolontà, G. P. Emerenziani, M. C. Gallotta, A. J. Silva, and L. Guidetti, "Accuracy, reliability, linearity of accutrend and lactate pro versus ebio plus analyzer," *European journal of applied physiology*, vol. 107, no. 1, pp. 105–111, 2009.

[6] Y. Ni, Y. Qiu, W. Jiang, K. Suttlemyre, M. Su, W. Zhang, W. Jia, and X. Du, "Adap-gc 2.0: deconvolution of coeluting metabolites from gc/tof-ms data for metabolomics studies," *Analytical chemistry*, vol. 84, no. 15, pp. 6619–6629, 2012.

[7] E. Tyteca, M. Talebi, R. Amos, S. H. Park, M. Taraji, Y. Wen, R. Szucs, C. A. Pohl, J. W. Dolan, and P. R. Haddad, "Towards a chromatographic similarity index to establish localized quantitative structure-retention models for retention prediction: use of retention factor ratio," *Journal of Chromatography A*, vol. 1486, pp. 50–58, 2017.

[8] R. G. Edwards and B. Joo, "The chroma software system for lattice qcd," *arXiv preprint hep-lat/0409003*, 2004.

[9] X. Huang, Y.-J. Chen, K. Cho, I. Nikolskiy, P. A. Crawford, and G. J. Patti, "X13cms: global tracking of isotopic labels in untargeted metabolomics," *Analytical chemistry*, vol. 86, no. 3, pp. 1632–1639, 2014.

[10] W. Zhang, J. Chang, Z. Lei, D. Huhman, L. W. Sumner, and P. X. Zhao, "Met-cofea: a liquid chromatography/mass spectrometry data processing platform for metabolite compound feature extraction and annotation," *Analytical chemistry*, vol. 86, no. 13, pp. 6245–6253, 2014.

[11] W. Zhang, Z. Lei, D. Huhman, L. W. Sumner, and P. X. Zhao, "Met-xalign: a metabolite cross-alignment tool for lc/ms-based comparative metabolomics," *Analytical chemistry*, vol. 87, no. 18, pp. 9114–9119, 2015.

[12] E. Valentini, M. Zampieri, M. Malavolta, M. G. Bacalini, R. Calabrese, T. Guastafierro, A. Reale, C. Franceschi, A. Hervonen, B. Koller *et al.*, "Analysis of the machinery and intermediates of the 5hmc-mediated dna demethylation pathway in aging on samples from the mark-age study," *Aging (Albany NY)*, vol. 8, no. 9, p. 1896, 2016.

[13] M. Katajamaa, J. Miettinen, and M. Orešič, "Mzmine: toolbox for processing and visualization of mass spectrometry based molecular profile data," *Bioinformatics*, vol. 22, no. 5, pp. 634–636, 2006.

[14] C. D. Broeckling, I. R. Reddy, A. L. Duran, X. Zhao, and L. W. Sumner, "Met-idea: data extraction tool for mass spectrometry-based metabolomics," *Analytical chemistry*, vol. 78, no. 13, pp. 4334–4341, 2006.

[15] J. J. Boon, "Analytical pyrolysis mass spectrometry: new vistas opened by temperature-resolved in-source pyms," *International Journal of Mass Spectrometry and Ion Processes*, vol. 118, pp. 755–787, 1992.

[16] C.-A. Darveau, R. K. Suarez, R. D. Andrews, and P. W. Hochachka, "Allometric cascade as a unifying principle of body mass effects on metabolism," *Nature*, vol. 417, no. 6885, p. 166, 2002.

[17] D. Reddy, C. Sekar, C. Gaur, D. B. J. Causon, and M. Pillai, "Identification & confirmation of structurally related degradation products of simvastatin."

[18] C. A. Smith, E. J. Want, G. O'Maille, R. Abagyan, and G. Siuzdak, "Xcms: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification," *Analytical chemistry*, vol. 78, no. 3, pp. 779–787, 2006.

[19] P. M. Ayton, E. H. Chen, and M. L. Cleary, "Binding to nonmethylated cpg dna is essential for target recognition, transactivation, and myeloid transformation by an mll oncoprotein," *Molecular and cellular biology*, vol. 24, no. 23, pp. 10 470–10 478, 2004.

[20] K. Jobu, C. Sun, S. Yoshioka, J. Yokota, M. Onogawa, C. Kawada, K. Inoue, T. Shuin, T. Sendo, and M. Miyamura, "Metabolomics study on the biochemical profiles of odor elements in urine of human with bladder cancer," *Biological and Pharmaceutical Bulletin*, vol. 35, no. 4, pp. 639–642, 2012.

[21] T. Naake and E. Gaquerel, "Metcirc: navigating mass spectral similarity in high-resolution ms/ms metabolomics data," *Bioinformatics*, vol. 33, no. 15, pp. 2419–2420, 2017.

[22] D. N. Cole, "Controlling the spread of campsites at popular wilderness destinations," *Journal of Soil and Water Conservation*, vol. 37, no. 5, pp. 291–295, 1982.

[23] S. Beisken, P. Conesa, K. Haug, R. M. Salek, and C. Steinbeck, "Specktackle: Javascript charts for spectroscopy," *Journal of cheminformatics*, vol. 7, no. 1, p. 17, 2015.

[24] E. Want and P. Masson, "Processing and analysis of gc/lc-ms-based metabolomics data," in *Metabolic profiling*. Springer, 2011, pp. 277–298.

[25] D. A. Shirley, "High-resolution x-ray photoemission spectrum of the valence bands of gold," *Physical Review B*, vol. 5, no. 12, p. 4709, 1972.

[26] M. C. Potter, M. M. Chun, B. S. Banks, and M. Muckenhoupt, "Two attentional deficits in serial target search: the visual attentional blink and an amodal task-switch deficit." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 24, no. 4, p. 979, 1998.

[27] D. M. Small, "The effects of glyceride structure on absorption and metabolism," *Annual review of nutrition*, vol. 11, no. 1, pp. 413–434, 1991.

[28] H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn, and M. Arita, "Ms-dial: data-independent ms/ms deconvolution for comprehensive metabolome analysis," *Nature methods*, vol. 12, no. 6, p. 523, 2015.

[29] J. Elith, S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates, "A statistical explanation of maxent for ecologists," *Diversity and distributions*, vol. 17, no. 1, pp. 43–57, 2011.

[30] C. Szabo, C. Ransy, K. Módis, M. Andriamihaja, B. Murghes, C. Coletta, G. Olah, K. Yanagi, and F. Bouillaud, "Regulation of mitochondrial bioenergetic function by hydrogen sulfide. part i. biochemical and physiological mechanisms," *British journal of pharmacology*, vol. 171, no. 8, pp. 2099–2122, 2014.

[31] Y. Inoue, T. Kubokawa, and K. Sato, "Photocatalytic activity of alkali-metal titanates combined with ruthenium in the decomposition of water," *The Journal of Physical Chemistry*, vol. 95, no. 10, pp. 4059–4063, 1991.

[32] R. Aggio, S. G. Villas-Bôas, and K. Ruggiero, "Metab: an r package for high-throughput analysis of metabolomics data generated by gc-ms," *Bioinformatics*, vol. 27, no. 16, pp. 2316–2318, 2011.

[33] I. Nikolskiy, N. G. Mahieu, Y.-J. Chen, R. Tautenhahn, and G. J. Patti, "An untargeted metabolomic workflow to improve structural characterization of metabolites," *Anal. Chem*, vol. 85, no. 16, pp. 7713–7719, 2013.

[34] J. Zaia, K. Khatri, J. Klein, C. Shao, Y. Sheng, and R. Viner, "Complete molecular weight profiling of low-molecular weight heparins using size exclusion chromatography-ion suppressor-high-resolution mass spectrometry," *Analytical chemistry*, vol. 88, no. 21, pp. 10 654–10 660, 2016.

[35] M. Sonnett, E. Yeung, and M. Wuhr, "Accurate, sensitive, and precise multiplexed proteomics using the complement reporter ion cluster," *Analytical chemistry*, vol. 90, no. 8, pp. 5032–5039, 2018.

[36] M.-F. Bernet, D. Brassart, J.-R. Neeser, and A. Servin, "Lactobacillus acidophilus la 1 binds to cultured human intestinal cell lines and inhibits cell attachment and cell invasion by enterovirulent bacteria." *Gut*, vol. 35, no. 4, pp. 483–489, 1994.

[37] W. Cui, H. W. Rohrs, and M. L. Gross, "Top-down mass spectrometry: recent developments, applications and perspectives," *Analyst*, vol. 136, no. 19, pp. 3854–3864, 2011.

[38] V. Picaud, J.-F. Giovannelli, C. Truntzer, J.-P. Charrier, A. Giremus, P. Grangeat, and C. Mercier, "Linear maldi-tof simultaneous spectrum deconvolution and baseline removal," *BMC bioinformatics*, vol. 19, no. 1, p. 123, 2018.

[39] M. Bartholomae, T. Baumann, J. H. Nickling, D. Peterhoff, R. Wagner, N. Budisa, and O. P. Kuipers, "Expanding the genetic code of lactococcus lactis and escherichia coli to incorporate non-canonical amino acids for production of modified lantibiotics," *Frontiers in microbiology*, vol. 9, p. 657, 2018.

[40] A. Delabrière, U. M. Hohenester, B. Colsch, C. Junot, F. Fenaille, and E. A. Thévenot, "profia: a data preprocessing workflow for flow injection analysis coupled to high-resolution mass spectrometry," *Bioinformatics*, vol. 33, no. 23, pp. 3767–3775, 2017.

[41] T. Yu, Y. Park, J. M. Johnson, and D. P. Jones, "aplcmsadaptive processing of high-resolution lc/ms data," *Bioinformatics*, vol. 25, no. 15, pp. 1930–1936, 2009.

[42] U. Kolitsch, H. Seifert, T. Ludwig, and F. Aldinger, "Phase equilibria and crystal chemistry in the y 2 o 3–al 2 o 3–sio 2 system," *Journal of materials research*, vol. 14, no. 2, pp. 447–455, 1999.

[43] V. Santini, A. Sanna, A. Bosi, G. Alimena, G. Loglisci, A. Levis, F. Salvi, C. Finelli, C. Clissa, G. Specchia *et al.*, "An observational multicenter study to assess the cost of illness and quality of life in patients with myelodysplastic syndromes in italy," 2011.

[44] S. Vey and A. Voigt, "Amdis: adaptive multidimensional simulations," *Computing and Visualization in Science*, vol. 10, no. 1, pp. 57–67, 2007.

[45] A. Chokkathukalam, A. Jankevics, D. J. Creek, F. Achcar, M. P. Barrett, and R. Breitling, "mzmatch–iso: an r tool for the annotation and relative quantification of isotope-labelled mass spectrometry data," *Bioinformatics*, vol. 29, no. 2, pp. 281–283, 2012.

[46] P. C. das Neves, P. H. N. Rangel, and V. d. A. Cutrim, "Selección recurrente para la producción de arroz híbrido." Centro Internacional de Agricultura Tropical (CIAT), 1997.

[47] B. Bunk, M. Kucklick, R. Jonas, R. Münch, M. Schobert, D. Jahn, and K. Hiller, "Metaquant: a tool for the automatic quantification of gc/ms-based metabolome data," *Bioinformatics*, vol. 22, no. 23, pp. 2962–2965, 2006.

[48] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, 2000.

[49] B. C. DeFelice, S. S. Mehta, S. Samra, T. Cajka, B. Wancewicz, J. F. Fahrmann, and O. Fiehn, "Mass spectral feature list optimizer (ms-flo): A tool to minimize false positive peak reports in untargeted liquid chromatography–mass spectroscopy (lc-ms) data processing," *Analytical chemistry*, vol. 89, no. 6, pp. 3250–3255, 2017.

[50] T. Jansen and I. Wegener, "On the analysis of evolutionary algorithmsa proof that crossover really can help," in *European Symposium on Algorithms*. Springer, 1999, pp. 184–193.

[51] L. G. Johnsen, T. Skov, U. Houlberg, and R. Bro, "An automated method for baseline correction, peak finding and peak grouping in chromatographic data," *Analyst*, vol. 138, no. 12, pp. 3502–3511, 2013.

[52] V. Treviño, I.-L. Yañez-Garza, C. E. Rodriguez-López, R. Urrea-López, M.-L. Garza-Rodriguez, H.-A. Barrera-Saldaña, J. G. Tamez-Peña, R. Winkler, and R.-I. Díaz de-la Garza, "Gridmass: a fast two-dimensional feature detection method for lc/ms," *Journal of Mass Spectrometry*, vol. 50, no. 1, pp. 165–174, 2015.

[53] W. Peng and D. J. Edwards, "K-means like minimum mean distance algorithm for wireless sensor networks," in *2010 2nd International Conference on Computer Engineering and Technology*, vol. 1. IEEE, 2010, pp. V1–120.

[54] L. K. Grover, "Quantum mechanics helps in searching for a needle in a haystack," *Physical review letters*, vol. 79, no. 2, p. 325, 1997.

[55] P. Tasiopoulou and N. Gerolymos, "Constitutive modeling for sand with emphasis on the evolution of bounding and phase transformation lines," in *Proc. of the 8th European Conference on Numerical Methods in Geotechnical Engineering, Delft, the Netherlands*, 2014, pp. 109–114.

[56] N. H. Beltrán, M. A. Duarte-Mermoud, V. A. S. Vicencio, S. A. Salah, and M. A. Bustos, "Chilean wine classification using volatile organic compounds data obtained with a fast gc analyzer," *IEEE Transactions*

on Instrumentation and Measurement, vol. 57, no. 11, pp. 2421–2436, 2008.

[57] M. G. Kok, M. M. Ruijken, J. R. Swann, I. D. Wilson, G. W. Somsen, and G. J. de Jong, "Anionic metabolic profiling of urine from antibiotic-treated rats by capillary electrophoresis–mass spectrometry," *Analytical and bioanalytical chemistry*, vol. 405, no. 8, pp. 2585–2594, 2013.

[58] W. Zhou, S.-L. Su, J.-A. Duan, J.-M. Guo, D.-W. Qian, E.-X. Shang, and J. Zhang, "Characterization of the active constituents in shixiao san using bioactivity evaluation followed by uplc-qtof and markerlynx analysis," *Molecules*, vol. 15, no. 9, pp. 6217–6230, 2010.

[59] D. Mao, R. Lookman, H. Van De Weghe, R. Weltens, G. Vanermen, N. De Brucker, and L. Diels, "Estimation of ecotoxicity of petroleum hydrocarbon mixtures in soil based on hplc–gcxgc analysis," *Chemosphere*, vol. 77, no. 11, pp. 1508–1513, 2009.

[60] J. Suhy, W. Rooney, D. Goodkin, A. Capizzano, B. Soher, A. A. Maudsley, E. Waubant, P. Andersson, and M. Weiner, "1h mrsi comparison of white matter and lesions in primary progressive and relapsing-remitting ms," *Multiple Sclerosis Journal*, vol. 6, no. 3, pp. 148–155, 2000.

[61] B. B. Misra, J. F. Fahrmann, and D. Grapov, "Review of emerging metabolomic tools and resources: 2015–2016," *Electrophoresis*, vol. 38, no. 18, pp. 2257–2274, 2017.

[62] N. Zhong, J.-Z. Dong, Y. Yao, and S. Ohsuga, "Gastric cancer data mining with ordered information," in *International Conference on Rough Sets and Current Trends in Computing*. Springer, 2002, pp. 467–478.

[63] B. Li, J. Tang, Q. Yang, S. Li, X. Cui, Y. Li, Y. Chen, W. Xue, X. Li, and F. Zhu, "Noreva: normalization and evaluation of ms-based metabolomics data," *Nucleic acids research*, vol. 45, no. W1, pp. W162–W170, 2017.

[64] N. Desmarais, F. Reuse, and S. Khanna, "Magnetic coupling in neutral and charged cr 2, mn 2, and crmn dimers," *The Journal of Chemical Physics*, vol. 112, no. 13, pp. 5576–5584, 2000.

[65] B. M. Sullivan and S.-P. Hsu, "Biometric score normalizer," Jul. 8 2003, uS Patent 6,591,224.

[66] B. Wen, Z. Mei, C. Zeng, and S. Liu, "metax: a flexible and comprehensive software for processing metabolomics data," *BMC bioinformatics*, vol. 18, no. 1, p. 183, 2017.

[67] B. Ernest, J. R. Gooding, S. R. Campagna, A. M. Saxton, and B. H. Voy, "Metabr: an r script for linear model analysis of quantitative metabolomic data," *BMC research notes*, vol. 5, no. 1, p. 596, 2012.

[68] J. A. Berger, S. Hautaniemi, A.-K. Järvinen, H. Edgren, S. K. Mitra, and J. Astola, "Optimized lowess normalization parameter selection for dna microarray data," *BMC bioinformatics*, vol. 5, no. 1, p. 194, 2004.

[69] G. Hughes, C. Cruickshank-Quinn, R. Reisdorph, S. Lutz, I. Petrache, N. Reisdorph, R. Bowler, and K. Kechris, "Msprepsummarization, normalization and diagnostics for processing of mass spectrometry–based metabolomic data," *Bioinformatics*, vol. 30, no. 1, pp. 133–134, 2013.

[70] H. Tsugawa, T. Bamba, M. Shinohara, S. Nishiumi, M. Yoshida, and E. Fukusaki, "Practical non-targeted gas chromatography/mass spectrometry-based metabolomics platform for metabolic phenotype analysis," *Journal of bioscience and bioengineering*, vol. 112, no. 3, pp. 292–298, 2011.

[71] T. Kokubun and L. DCosta, "Direct and unbiased information recovery from liquid chromatography–mass spectrometry raw data for phenotype-differentiating metabolites based on screening window coefficient of ion currents," *Analytical chemistry*, vol. 85, no. 18, pp. 8684–8691, 2013.

[72] A. Geremia, C. V. Arancibia-Cárcamo, M. P. Fleming, N. Rust, B. Singh, N. J. Mortensen, S. P. Travis, and F. Powrie, "Il-23–responsive innate lymphoid cells are increased in inflammatory bowel disease," *Journal of Experimental Medicine*, vol. 208, no. 6, pp. 1127–1133, 2011.

[73] R. C. Ramaker, E. R. Gordon, and S. J. Cooper, "R2dgc: threshold-free peak alignment and identification for 2d gas chromatography-mass spectrometry in r," *Bioinformatics*, vol. 34, no. 10, pp. 1789–1791, 2017.

[74] E. A. Manrao, I. M. Derrington, A. H. Laszlo, K. W. Langford, M. K. Hopper, N. Gillgren, M. Pavlenok, M. Niederweis, and J. H. Gundlach, "Reading dna at single-nucleotide resolution with a mutant mspa nanopore and phi29 dna polymerase," *Nature biotechnology*, vol. 30, no. 4, p. 349, 2012.

[75] P. Kuich, N. Hoffmann, and S. Kempa, "Maui-via: a user-friendly software for visual identification, alignment, correction, and quantification of gas chromatography–mass spectrometry data," *Frontiers in bioengineering and biotechnology*, vol. 2, p. 84, 2015.

[76] A. S. Kirpich, M. Ibarra, O. Moskalenko, J. M. Fear, J. Gerken, X. Mi, A. Ashrafi, A. M. Morse, and L. M. McIntyre, "Secimtools: a suite of metabolomics data analysis tools," *BMC bioinformatics*, vol. 19, no. 1, p. 151, 2018.

[77] D. M. Avtonomov, A. Raskind, and A. I. Nesvizhskii, "Batmass: a java software platform for lc–ms data visualization in proteomics and metabolomics," *Journal of proteome research*, vol. 15, no. 8, pp. 2500–2509, 2016.

[78] M. Calderón-Santiago, M. A. López-Bascón, Á. Peralbo-Molina, and F. Priego-Capote, "Metaboqc: A tool for correcting untargeted metabolomics data with mass spectrometry detection using quality controls," *Talanta*, vol. 174, pp. 29–37, 2017.

[79] A. M. Simader, B. Kluger, N. K. N. Neumann, C. Bueschl, M. Lemmens, G. Lirk, R. Krska, and R. Schuhmacher, "Qcscreen: a software tool for data quality control in lc-hrms based metabolomics," *BMC bioinformatics*, vol. 16, no. 1, p. 341, 2015.

[80] H. Luan, F. Ji, Y. Chen, and Z. Cai, "stattarget: A streamlined tool for signal drift correction and interpretations of quantitative mass spectrometry-based omics data," *Analytica chimica acta*, vol. 1036, pp. 66–72, 2018.

[81] B. Hoekman, R. Breitling, F. Suits, R. Bischoff, and P. Horvatovich, "mscompare: a framework for quantitative analysis of label-free lc-ms data for comparative candidate biomarker studies," *Molecular & Cellular Proteomics*, vol. 11, no. 6, pp. M111–015 974, 2012.

[82] M. M. Smits, R. Carleer, and J. V. Colpaert, "Pyquan: A rapid workflow around the amdis deconvolution software for high throughput analysis of pyrolysis gc/ms data," *Journal of analytical and applied pyrolysis*, vol. 118, pp. 335–342, 2016.

[83] Z. Jin, J. Kang, and T. Yu, "Missing value imputation for lc-ms metabolomics data by incorporating metabolic network and adduct ion relations," *Bioinformatics*, vol. 34, no. 9, pp. 1555–1561, 2017.

[84] T. Matsuo, H. Tsugawa, H. Miyagawa, and E. Fukusaki, "Integrated strategy for unknown ei–ms identification using quality control calibration curve, multivariate analysis, ei–ms spectral database, and retention index prediction," *Analytical chemistry*, vol. 89, no. 12, pp. 6766–6773, 2017.

[85] A. A. Gonzalez Ibarra, K. Wrobel, E. Yanez Barrientos, A. R. Corrales Escobosa, J. F. Gutierrez Corona, I. Enciso Donis, and K. Wrobel, "Changes of metabolomic profile in helianthus annuus under exposure to chromium (vi) studied by caphplc-esi-qtof-ms and ms/ms," *Journal of analytical methods in chemistry*, vol. 2017, 2017.

[86] D. Shen and C. Davatzikos, "Hammer: hierarchical attribute matching mechanism for elastic registration," in *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2001)*. IEEE, 2001, pp. 29–36.

[87] P.-M. Allard, G. Genta-Jouve, and J.-L. Wolfender, "Deep metabolome annotation in natural products research: towards a virtuous cycle in metabolite identification," *Current opinion in chemical biology*, vol. 36, pp. 40–49, 2017.

[88] S. Clerens, W. Van den Ende, P. Verhaert, L. Geenen, and L. Arckens, "Sweet substitute: A software tool for in silico fragmentation of peptide-linked n-glycans," *Proteomics*, vol. 4, no. 3, pp. 629–632, 2004.

[89] S. Wolf, S. Schmidt, M. Müller-Hannemann, and S. Neumann, "In silico fragmentation for computer assisted identification of metabolite mass spectra," *BMC bioinformatics*, vol. 11, no. 1, p. 148, 2010.

[90] M. Gerlich and S. Neumann, "Metfusion: integration of compound identification strategies," *Journal of Mass Spectrometry*, vol. 48, no. 3, pp. 291–298, 2013.

[91] M. Cao, K. Fraser, and S. Rasmussen, "Computational analyses of spectral trees from electrospray multi-stage mass spectrometry to aid metabolite identification," *Metabolites*, vol. 3, no. 4, pp. 1036–1050, 2013.

[92] A. Pelander, E. Tyrkkö, and I. Ojanperä, "In silico methods for predicting metabolism and mass fragmentation applied to quetiapine in liquid chromatography/time-of-flight mass spectrometry urine drug screening," *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*, vol. 23, no. 4, pp. 506–514, 2009.

[93] G. G. Raffelt, J. Redondo, and N. V. Maira, "The mev mass frontier of axion physics," *Physical Review D*, vol. 84, no. 10, p. 103008, 2011.

[94] A. Rodriguez-Martinez, R. Ayala, J. M. Posma, N. Harvey, B. Jiménez, K. Sonomura, T.-A. Sato, F. Matsuda, P. Zalloua, D. Gauguier *et al.*, "pjres binning algorithm (jba): a new method to facilitate the recovery of metabolic information from pjres 1h nmr spectra," *Bioinformatics*, 2018.

[95] T. Haller, M. Kals, T. Esko, R. Mägi, and K. Fischer, "Regscan: a gwas tool for quick estimation of allele effects on continuous traits and their

combinations," *Briefings in bioinformatics*, vol. 16, no. 1, pp. 39–44, 2013.

[96] Y. Wen, P. Li, J. Hao, C. Duan, J. Han, A. He, Y. Du, L. Liu, X. Liang, F. Zhang *et al.*, "Integrating genome-wide dna methylation and mrna expression profiles identified different molecular features between kashin-beck disease and primary osteoarthritis," *Arthritis research & therapy*, vol. 20, no. 1, p. 41, 2018.

[97] M. Kutmon, M. P. van Iersel, A. Bohler, T. Kelder, N. Nunes, A. R. Pico, and C. T. Evelo, "Pathvisio 3: an extendable pathway analysis toolbox," *PLoS computational biology*, vol. 11, no. 2, p. e1004085, 2015.

[98] R. M. Schueler, "Self-deploying trusses containing shape-memory polymers," 2008.

[99] E. G. Cerami, G. D. Bader, B. E. Gross, and C. Sander, "cpath: open source software for collecting, storing, and querying biological pathways," *BMC bioinformatics*, vol. 7, no. 1, p. 497, 2006.

[100] D. K. Barupal, P. K. Haldiya, G. Wohlgemuth, T. Kind, S. L. Kothari, K. E. Pinkerton, and O. Fiehn, "Metamapp: mapping and visualizing metabolomic data by integrating information from biochemical pathways and chemical and mass spectral similarity," *BMC bioinformatics*, vol. 13, no. 1, p. 99, 2012.

[101] L. B. Jorde, "Linkage disequilibrium as a gene-mapping tool." *American journal of human genetics*, vol. 56, no. 1, p. 11, 1995.

[102] D. H. Huson and C. Xie, "A poor mans blastxhigh-throughput metagenomic protein database search using pauda," *Bioinformatics*, vol. 30, no. 1, pp. 38–39, 2013.

[103] O. Ostrenko, P. Incardona, R. Ramaswamy, L. Brusch, and I. F. Sbalzarini, "pssalib: The partial-propensity stochastic chemical network simulator," *PLoS computational biology*, vol. 13, no. 12, p. e1005865, 2017.

[104] A. Joyce, U. Z. Ijaz, C. Nzeteu, A. Vaughan, S. L. Shirran, C. H. Botting, C. Quince, V. OFlaherty, and F. Abram, "Linking microbial community structure and function during the acidified anaerobic digestion of grass," *Frontiers in microbiology*, vol. 9, p. 540, 2018.

[105] A. Marco-Ramell, M. Palau-Rodriguez, A. Alay, S. Tulipani, M. Urpi-Sarda, A. Sanchez-Pla, and C. Andres-Lacueva, "Evaluation and comparison of bioinformatic tools for the enrichment analysis of metabolomics data," *BMC bioinformatics*, vol. 19, no. 1, p. 1, 2018.

[106] S. Dimitrov, V. Kamenska, J. Walker, W. Windle, R. Purdy, M. Lewis, and O. Mekenyan, "Predicting the biodegradation products of perfluorinated chemicals using catabol," *SAR and QSAR in Environmental Research*, vol. 15, no. 1, pp. 69–82, 2004.

[107] R. Hernández-de Diego, S. Tarazona, C. Martinez-Mira, L. Balzano-Nogueira, P. Furio-Tari, G. J. Pappas Jr, and A. Conesa, "Paintomics 3: a web resource for the pathway analysis and visualization of multi-omics data," *Nucleic acids research*, vol. 46, no. W1, pp. W503–W509, 2018.

[108] N. Lavender, D. W. Hein, G. Brock, and L. C. R. Kidd, "Evaluation of oxidative stress response related genetic variants, pro-oxidants, antioxidants and prostate cancer," *AIMS medical science*, vol. 2, no. 4, p. 271, 2015.

[109] A. A. Cuellar, C. M. Lloyd, P. F. Nielsen, D. P. Bullivant, D. P. Nickerson, and P. J. Hunter, "An overview of cellml 1.1, a biological model description language," *Simulation*, vol. 79, no. 12, pp. 740–747, 2003.

[110] S. Wigby, L. K. Sirot, J. R. Linklater, N. Buehner, F. C. Calboli, A. Bretman, M. F. Wolfner, and T. Chapman, "Seminal fluid protein allocation and male reproductive success," *Current Biology*, vol. 19, no. 9, pp. 751–757, 2009.

[111] Y. Moriya, D. Shigemizu, M. Hattori, T. Tokimatsu, M. Kotera, S. Goto, and M. Kanehisa, "Pathpred: an enzyme-catalyzed metabolic pathway prediction server," *Nucleic acids research*, vol. 38, no. suppl_2, pp. W138–W143, 2010.

[112] U. Wittig, R. Kania, M. Golebiewski, M. Rey, L. Shi, L. Jong, E. Algaa, A. Weidemann, H. Sauer-Danzwith, S. Mir *et al.*, "Sabio-rkdatabase for biochemical reaction kinetics," *Nucleic acids research*, vol. 40, no. D1, pp. D790–D796, 2011.

[113] A. R. Zomorrodi, M. M. Islam, and C. D. Maranas, "d-optcom: dynamic multi-level and multi-objective metabolic modeling of microbial communities," *ACS synthetic biology*, vol. 3, no. 4, pp. 247–257, 2014.

[114] C. Li, J. Han, Q. Yao, C. Zou, Y. Xu, C. Zhang, D. Shang, L. Zhou, C. Zou, Z. Sun *et al.*, "Subpathway-gm: identification of metabolic subpathways via joint power of interesting genes and metabolites and their topologies within pathways," *Nucleic acids research*, vol. 41, no. 9, pp. e101–e101, 2013.

[115] P. K. Sreenivasaiah, S. Rani, J. Cayetano, N. Arul, and D. H. Kim, "Ipavs: Integrated pathway resources, analysis and visualization system," *Nucleic acids research*, vol. 40, no. D1, pp. D803–D808, 2011.

[116] D. B. Warheit, B. R. Laurence, K. L. Reed, D. H. Roach, G. A. Reynolds, and T. R. Webb, "Comparative pulmonary toxicity assessment of single-wall carbon nanotubes in rats," *Toxicological sciences*, vol. 77, no. 1, pp. 117–125, 2004.

[117] J. T. Bates, D. Chivian, and A. P. Arkin, "Glamm: genome-linked application for metabolic maps," *Nucleic acids research*, vol. 39, no. suppl_2, pp. W400–W405, 2011.

[118] K. Sriyudthsak, R. F. Mejia, M. Arita, and M. Y. Hirai, "Pasmet: a web-based platform for prediction, modelling and analyses of metabolic systems," *Nucleic acids research*, vol. 44, no. W1, pp. W205–W211, 2016.

[119] N. Hadadi, K. C. Soh, M. Seijo, A. Zisaki, X. Guan, M. R. Wenk, and V. Hatzimanikatis, "A computational framework for integration of lipidomics data into metabolic pathways," *Metabolic engineering*, vol. 23, pp. 1–8, 2014.

[120] D. Grapov, K. Wanichthanarak, and O. Fiehn, "Metamapr: pathway independent metabolomic network analysis incorporating unknowns," *Bioinformatics*, vol. 31, no. 16, pp. 2757–2760, 2015.

[121] L. M. Mulligan, J. B. Kwok, C. S. Healey, M. J. Elsdon, C. Eng, E. Gardner, D. R. Love, S. E. Mole, J. K. Moore, L. Papi *et al.*, "Germ-line mutations of the ret proto-oncogene in multiple endocrine neoplasia type 2a," *Nature*, vol. 363, no. 6428, p. 458, 1993.

[122] T. V. Sivakumar, V. Giri, J. H. Park, T. Y. Kim, and A. Bhaduri, "Reactpred: a tool to predict and analyze biochemical reactions," *Bioinformatics*, vol. 32, no. 22, pp. 3522–3524, 2016.

[123] M. Vass, N. Allen, C. A. Shaffer, N. Ramakrishnan, L. T. Watson, and J. J. Tyson, "The jigcell model builder and run manager," *Bioinformatics*, vol. 20, no. 18, pp. 3680–3681, 2004.

[124] C. Sanford, M. L. Yip, C. White, and J. Parkinson, "Cell++simulating biochemical pathways," *Bioinformatics*, vol. 22, no. 23, pp. 2918–2925, 2006.

[125] F. Jourdan, R. Breitling, M. P. Barrett, and D. Gilbert, "Metanetter: inference and visualization of high-resolution metabolomic networks," *Bioinformatics*, vol. 24, no. 1, pp. 143–145, 2007.

[126] S. Pabinger, R. Snajder, T. Hardiman, M. Willi, A. Dander, and Z. Trajanoski, "Memosys 2.0: an update of the bioinformatics database for genome-scale models and genomic data," *Database*, vol. 2014, 2014.

[127] J. Einloft, J. Ackermann, J. Nöthen, and I. Koch, "Monalisavisualization and analysis of functional modules in biochemical networks," *Bioinformatics*, vol. 29, no. 11, pp. 1469–1470, 2013.

[128] C. P. Fisher, N. J. Plant, J. B. Moore, and A. M. Kierzek, "Qsspn: dynamic simulation of molecular interaction networks describing gene regulation, signalling and whole-cell metabolism in human cells," *Bioinformatics*, vol. 29, no. 24, pp. 3181–3190, 2013.

[129] R. Carr and E. Borenstein, "Netseed: a network-based reverse-ecology tool for calculating the metabolic interface of an organism with its environment," *Bioinformatics*, vol. 28, no. 5, pp. 734–735, 2012.

[130] M. P. Gerstl, C. Jungreuthmayer, and J. Zanghellini, "tefma: computing thermodynamically feasible elementary flux modes in metabolic networks," *Bioinformatics*, vol. 31, no. 13, pp. 2232–2234, 2015.

[131] A. Kreimer, A. Doron-Faigenboim, E. Borenstein, and S. Freilich, "Netcmpt: a network-based tool for calculating the metabolic competition between bacterial species," *Bioinformatics*, vol. 28, no. 16, pp. 2195–2197, 2012.

[132] R. R. Adams, "Sed-ed, a workflow editor for computational biology experiments written in sed-ml," *Bioinformatics*, vol. 28, no. 8, pp. 1180–1181, 2012.

[133] T. Forth, G. A. McConkey, and D. R. Westhead, "Metnetmaker: a free and open-source tool for the creation of novel metabolic networks in sbml format," *Bioinformatics*, vol. 26, no. 18, pp. 2352–2353, 2010.

[134] V. Giri, T. V. Sivakumar, K. M. Cho, T. Y. Kim, and A. Bhaduri, "Rxnsim: a tool to compare biochemical reactions," *Bioinformatics*, vol. 31, no. 22, pp. 3712–3714, 2015.

[135] M. Chazalviel, C. Frainay, N. Poupin, F. Vinson, B. Merlet, Y. Gloaguen, L. Cottret, and F. Jourdan, "Metexploreviz: web component for interactive metabolic network visualization," *Bioinformatics*, vol. 34, no. 2, pp. 312–313, 2017.

[136] S. Basu, W. Duren, C. R. Evans, C. F. Burant, G. Michailidis, and A. Karnovsky, "Sparse network modeling and metscape-based visualization methods for the analysis of large-scale metabolomics data," *Bioinformatics*, vol. 33, no. 10, pp. 1545–1553, 2017.

[137] T. Radchenko, F. Fontaine, L. Morettoni, and I. Zamora, "Web-metabase: cleavage sites analysis tool for natural and unnatural substrates from diverse data source," *Bioinformatics*, 2018.

[138] S. J. Sheridan and H. L. Springfield, "Age appropriate games in the teaching of leisure skills to persons with mental retardation." 1987.

[139] A. Rodriguez-Martinez, R. Ayala, J. M. Posma, A. L. Neves, D. Gauguier, J. K. Nicholson, and M.-E. Dumas, "Metabosignal: a network-based approach for topological analysis of metabotype regulation via metabolic and signaling pathways," *Bioinformatics*, vol. 33, no. 5, pp. 773–775, 2017.

[140] G. Basler and Z. Nikoloski, "Jmassbalance: Mass-balanced randomization and analysis of metabolic networks," *Bioinformatics*, vol. 27, no. 19, pp. 2761–2762, 2011.

[141] C. Jia, W. Guan, A. Yang, R. Xiao, W. W. Tang, C. S. Moravec, K. B. Margulies, T. P. Cappola, C. Li, and M. Li, "Metadiff: differential isoform expression analysis using random-effects meta-regression," *BMC bioinformatics*, vol. 16, no. 1, p. 208, 2015.

[142] E. C. Samuelsson, F. A. Victor, G. Tibblin, and K. F. Svärdsudd, "Signs of genital prolapse in a swedish population of women 20 to 59 years of age and possible related factors," *American journal of obstetrics and gynecology*, vol. 180, no. 2, pp. 299–305, 1999.

[143] C. D. Christensen, J.-H. S. Hofmeyr, and J. M. Rohwer, "Pyscestoolbox: a collection of metabolic pathway analysis tools," *Bioinformatics*, vol. 34, no. 1, pp. 124–125, 2017.

[144] V. De Anda, I. Zapata-Peñasco, A. C. Poot-Hernandez, L. E. Eguiarte, B. Contreras-Moreira, and V. Souza, "Mebs, a software platform to evaluate large (meta) genomic collections according to their metabolic machinery: unraveling the sulfur cycle," *GigaScience*, 2017.

[145] L. Feng, Y. Xu, Y. Zhang, Z. Sun, J. Han, C. Zhang, H. Yang, D. Shang, F. Su, X. Shi *et al.*, "Subpathway-gmir: identifying mirna-mediated metabolic subpathways by integrating condition-specific genes, micrornas, and pathway topologies," *Oncotarget*, vol. 6, no. 36, p. 39151, 2015.

[146] J. Ruiz-Armenta and R. Prol-Ledesma, "Techniques for enhancing the spectral response of hydrothermal alteration minerals in thematic mapper images of central mexico," *International Journal of Remote Sensing*, vol. 19, no. 10, pp. 1981–2000, 1998.

[147] D. Jiao, Y. Ye, and H. Tang, "Probabilistic inference of biochemical reactions in microbial communities from metagenomic sequences," *PLoS computational biology*, vol. 9, no. 3, p. e1002981, 2013.

[148] M. Ataman and V. Hatzimanikatis, "lumpgem: Systematic generation of subnetworks and elementally balanced lumped reactions for the biosynthesis of target metabolites," *PLoS computational biology*, vol. 13, no. 7, p. e1005513, 2017.

[149] M. Ataman, D. F. H. Gardiol, G. Fengos, and V. Hatzimanikatis, "redgem: Systematic reduction and analysis of genome-scale metabolic reconstructions for development of consistent core metabolic models," *PLoS computational biology*, vol. 13, no. 7, p. e1005444, 2017.

[150] T. Kieser, "Factors affecting the isolation of ccc dna from streptomyces lividans and escherichia coli," *Plasmid*, vol. 12, no. 1, pp. 19–36, 1984.

[151] T. Hu, K. Oksanen, W. Zhang, E. Randell, A. Furey, G. Sun, and G. Zhai, "An evolutionary learning and network approach to identifying key metabolites for osteoarthritis," *PLoS computational biology*, vol. 14, no. 3, p. e1005986, 2018.

[152] O. K. Kamneva, "Genome composition and phylogeny of microbes predict their co-occurrence in the environment," *PLoS computational biology*, vol. 13, no. 2, p. e1005366, 2017.

[153] N. M. Shchegolkova, G. S. Krasnov, A. A. Belova, A. A. Dmitriev, S. L. Kharitonov, K. M. Klimina, N. V. Melnikova, and A. V. Kudryavtseva, "Microbial community structure of activated sludge in treatment plants with different wastewater compositions," *Frontiers in microbiology*, vol. 7, p. 90, 2016.

[154] G. Yu and Q.-Y. He, "Reactomepa: an r/bioconductor package for reactome pathway analysis and visualization," *Molecular BioSystems*, vol. 12, no. 2, pp. 477–479, 2016.

[155] J. Fayn and P. Rubel, "Caviar: a serial ecg processing system for the comparative analysis of vcgs and their interpretation with autoreference to the patient," *Journal of electrocardiology*, vol. 21, pp. S173–S176, 1988.

[156] H. Mendes-Soares, M. Mundy, L. M. Soares, and N. Chia, "Mminte: an application for predicting metabolic interactions among the microbial species in a community," *BMC bioinformatics*, vol. 17, no. 1, p. 343, 2016.

[157] D. Waltemath, R. Adams, F. T. Bergmann, M. Hucka, F. Kolpakov, A. K. Miller, I. I. Moraru, D. Nickerson, S. Sahle, J. L. Snoep *et al.*, "Reproducible computational biology experiments with sed-ml-the simulation experiment description markup language," *BMC systems biology*, vol. 5, no. 1, p. 198, 2011.

[158] R. Levy, R. Carr, A. Kreimer, S. Freilich, and E. Borenstein, "Netcooperate: a network-based tool for inferring host-microbe and microbe-microbe cooperation," *BMC bioinformatics*, vol. 16, no. 1, p. 164, 2015.

[159] Y. Zhu, L. Sun, A. Garbarino, C. Schmidt, J. Fang, and J. Chen, "Pathrings: a web-based tool for exploration of ortholog and expression data in biological pathways," *BMC bioinformatics*, vol. 16, no. 1, p. 165, 2015.

[160] H. Zhou, J. Jin, H. Zhang, B. Yi, M. Wozniak, and L. Wong, "Intpath–an integrated pathway gene relationship database for model organisms and important pathogens," in *BMC systems biology*, vol. 6, no. 2. BioMed Central, 2012, p. S2.

[161] M. F. Gerard, G. Stegmayer, and D. H. Milone, "Evoms: An evolutionary tool to find de novo metabolic pathways," *Biosystems*, vol. 134, pp. 43–47, 2015.

[162] K. Choi, L. P. Smith, J. K. Medley, and H. M. Sauro, "phrased-ml: A paraphrased, human-readable adaptation of sed-ml," *Journal of bioinformatics and computational biology*, vol. 14, no. 06, p. 1650035, 2016.

[163] J. Schneider, F.-J. Vorhölter, E. Trost, J. Blom, Y. Musa, H. Neuweger, K. Niehaus, S. Schatschneider, A. Tauch, A. Goesmann *et al.*, "Carmen-comparative analysis and in silico reconstruction of organism-specific metabolic networks," *Genet Mol Res*, vol. 9, no. 3, pp. 1660–1672, 2010.

[164] M. Poolman, "Scrumpy: metabolic modelling with python," *IEE Proceedings-Systems Biology*, vol. 153, no. 5, pp. 375–378, 2006.

[165] S. Waldherr, "A guideline to model reduction by stoichiometric decomposition for biochemical network analysis," in *Proc. of the 21st International Symposium on Mathematical Theory of Networks and Systems pages*, 2014, pp. 490–495.

[166] N. Fyson, M. K. Kim, D. Lun, and C. Colijn, "Gene-centric constraint of metabolic models," *bioRxiv*, p. 116558, 2017.

[167] L. Corrêa, R. Alves, F. Goés, C. Chaparro, and L. Thom, "Funn-mg: A metagenomic systems biology computational framework," in *Brazilian Symposium on Bioinformatics*. Springer, 2014, pp. 25–32.

[168] I. L. Hofacker and P. F. Stadler, "Vienna-rnl," 2005.

[169] L. Angeles-Martinez and C. Theodoropoulos, "Estimation of flux distribution in metabolic networks accounting for thermodynamic constraints: The effect of equilibrium vs. blocked reactions," *Biochemical engineering journal*, vol. 105, pp. 347–357, 2016.

[170] S. Ekins, A. Bugrim, L. Brovold, E. Kirillov, Y. Nikolsky, E. Rakhmatulin, S. Sorokina, A. Ryabov, T. Serebryiskaya, A. Melnikov *et al.*, "Algorithms for network analysis in systems-adme/tox using the metacore and metadrug platforms," *Xenobiotica*, vol. 36, no. 10-11, pp. 877–901, 2006.

[171] Q. Zhang, S. Bhattacharya, D. E. Kline, R. B. Crawford, R. B. Conolly, R. S. Thomas, N. E. Kaminski, and M. E. Andersen, "Stochastic modeling of b lymphocyte terminal differentiation and its suppression by dioxin," *BMC systems biology*, vol. 4, no. 1, p. 40, 2010.

[172] A. Karnovsky, T. Weymouth, T. Hull, V. G. Tarcea, G. Scardoni, C. Laudanna, M. A. Sartor, K. A. Stringer, H. Jagadish, C. Burant *et al.*, "Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data," *Bioinformatics*, vol. 28, no. 3, pp. 373–380, 2011.

[173] M. Kankainen, P. Gopalacharyulu, L. Holm, and M. Orešič, "Mpeametabolite pathway enrichment analysis," *Bioinformatics*, vol. 27, no. 13, pp. 1878–1879, 2011.

[174] A. Kamburov, R. Cavill, T. M. Ebbels, R. Herwig, and H. C. Keun, "Integrated pathway-level analysis of transcriptomics and metabolomics data with impala," *Bioinformatics*, vol. 27, no. 20, pp. 2917–2918, 2011.

[175] M. Chagoyen and F. Pazos, "Mbrole: enrichment analysis of metabolomic data," *Bioinformatics*, vol. 27, no. 5, pp. 730–731, 2011.

[176] H. U. Zacharias, T. Rehberg, S. Mehrl, D. Richtmann, T. Wettig, P. J. Oefner, R. Spang, W. Gronwald, and M. Altenbuchinger, "Scale-invariant biomarker discovery in urine and plasma metabolite fingerprints," *Journal of proteome research*, vol. 16, no. 10, pp. 3596–3605, 2017.

[177] D. K. Barupal and O. Fiehn, "Chemical similarity enrichment analysis (chemrich) as alternative to biochemical pathway mapping for metabolomic datasets," *Scientific reports*, vol. 7, no. 1, p. 14567, 2017.

[178] S. Picart-Armada, F. Fernández-Albert, M. Vinaixa, O. Yanes, and A. Perera-Lluna, "Fella: an r package to enrich metabolomics data," *BMC bioinformatics*, vol. 19, no. 1, p. 538, 2018.

[179] M. Kaliszewski, F. Athias-Binche, and E. E. Lindquist, "Parasitism and parasitoidism in tarsonemina (acari: Heterostigmata) and evolutionary

considerations," in *Advances in Parasitology*. Elsevier, 1995, vol. 35, pp. 335–367.

[180] D. Verdegem, D. Lambrechts, P. Carmeliet, and B. Ghesquière, "Improved metabolite identification with midas and magma through ms/ms spectral dataset-driven parameter optimization," *Metabolomics*, vol. 12, no. 6, p. 98, 2016.

[181] J. Chong, O. Soufan, C. Li, I. Caraus, S. Li, G. Bourque, D. S. Wishart, and J. Xia, "Metaboanalyst 4.0: towards more transparent and integrative metabolomics analysis," *Nucleic acids research*, vol. 46, no. W1, pp. W486–W494, 2018.

[182] B. Usadel, F. Poree, A. Nagel, M. Lohse, A. CZEDIK-EYSENBERG, and M. Stitt, "A guide to using mapman to visualize and compare omics data in plants: a case study in the crop species, maize," *Plant, cell & environment*, vol. 32, no. 9, pp. 1211–1229, 2009.

[183] T.-C. Kuo, T.-F. Tian, and Y. J. Tseng, "3omics: a web-based systems biology tool for analysis, integration and visualization of human transcriptomic, proteomic and metabolomic data," *BMC systems biology*, vol. 7, no. 1, p. 64, 2013.

[184] K.-A. Lê Cao, I. González, and S. Déjean, "integromics: an r package to unravel relationships between two omics datasets," *Bioinformatics*, vol. 25, no. 21, pp. 2855–2856, 2009.

[185] W. Duren, T. Weymouth, T. Hull, G. S. Omenn, B. Athey, C. Burant, and A. Karnovsky, "Metdiseaseconnecting metabolites to diseases via literature," *Bioinformatics*, vol. 30, no. 15, pp. 2239–2241, 2014.

[186] S. J. Hinshaw, A. HY Lee, E. E. Gill, and R. EW Hancock, "Metabridge: enabling network-based integrative analysis via direct protein interactors of metabolites," *Bioinformatics*, vol. 34, no. 18, pp. 3225–3227, 2018.

[187] H. Redestig, M. Kusano, A. Fukushima, F. Matsuda, K. Saito, and M. Arita, "Consolidating metabolite identifiers to enable contextual and multi-platform metabolomics data analysis," *BMC bioinformatics*, vol. 11, no. 1, p. 214, 2010.

[188] H. Neuweger, M. Persicke, S. P. Albaum, T. Bekel, M. Dondrup, A. T. Hüser, J. Winnebald, J. Schneider, J. Kalinowski, and A. Goesmann, "Visualizing post genomics data-sets on customized pathway maps by prometra–aeration-dependent gene expression and metabolism of corynebacterium glutamicum as an example," *BMC systems biology*, vol. 3, no. 1, p. 82, 2009.

[189] E. K. Speliotes, C. J. Willer, S. I. Berndt, K. L. Monda, G. Thorleifsson, A. U. Jackson, H. L. Allen, C. M. Lindgren, J. Luan, R. Mägi *et al.*, "Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index," *Nature genetics*, vol. 42, no. 11, p. 937, 2010.

[190] H. Ji, F. Zeng, Y. Xu, H. Lu, and Z. Zhang, "Kpic2: an effective framework for mass spectrometry-based metabolomics using pure ion chromatograms," *Analytical chemistry*, vol. 89, no. 14, pp. 7631–7640, 2017.

[191] A. Kaever, M. Landesfeind, K. Feussner, A. Mosblech, I. Heilmann, B. Morgenstern, I. Feussner, and P. Meinicke, "Marvis-pathway: integrative and exploratory pathway analysis of non-targeted metabolomics data," *Metabolomics*, vol. 11, no. 3, pp. 764–777, 2015.

[192] Y. Tikunov, S. Laptenok, R. Hall, A. Bovy, and R. De Vos, "Msclust: a tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data," *Metabolomics*, vol. 8, no. 4, pp. 714–718, 2012.

[193] W. M. Edmands, D. K. Barupal, and A. Scalbert, "Metmsline: an automated and fully integrated pipeline for rapid processing of high-resolution lc–ms metabolomic datasets," *Bioinformatics*, vol. 31, no. 5, pp. 788–790, 2014.

[194] W. M. Edmands, J. Hayes, and S. M. Rappaport, "Simextargid: a comprehensive package for real-time lc-ms data acquisition and analysis," *Bioinformatics*, vol. 34, no. 20, pp. 3589–3590, 2018.

[195] K. Hiller, J. Hangebrauk, C. Jager, J. Spura, K. Schreiber, and D. Schomburg, "Metabolitedetector: comprehensive analysis tool for targeted and nontargeted gc/ms based metabolome analysis," *Analytical chemistry*, vol. 81, no. 9, pp. 3429–3439, 2009.

[196] C. Costa, M. Maraschin, and M. Rocha, "An r package for the integrated analysis of metabolomics and spectral data," *Computer methods and programs in biomedicine*, vol. 129, pp. 117–124, 2016.

[197] F. Giacomoni, G. Le Corguillé, M. Monsoor, M. Landi, P. Pericard, M. Pétéra, C. Duperier, M. Tremblay-Franco, J.-F. Martin, D. Jacob *et al.*, "Workflow4metabolomics: a collaborative research infrastructure for computational metabolomics," *Bioinformatics*, vol. 31, no. 9, pp. 1493–1495, 2014.

[198] N. Kessler, H. Neuweger, A. Bonte, G. Langenkämper, K. Niehaus, T. W. Nattkemper, and A. Goesmann, "Meltdb 2.0–advances of the metabolomics software system," *Bioinformatics*, vol. 29, no. 19, pp. 2452–2459, 2013.

[199] K. Uppal, Q. A. Soltow, F. H. Strobel, W. S. Pittard, K. M. Gernert, T. Yu, and D. P. Jones, "xmsanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data," *BMC bioinformatics*, vol. 14, no. 1, p. 15, 2013.

[200] A. M. Miguez, M. P. McNerney, and M. P. Styczynski, "Metabolomics analysis of the toxic effects of the production of lycopene and its precursors," *Frontiers in microbiology*, vol. 9, p. 760, 2018.

[201] A. J. Carroll, M. R. Badger, and A. H. Millar, "The metabolomeexpress project: enabling web-based processing, analysis and transparent dissemination of gc/ms metabolomics datasets," *BMC bioinformatics*, vol. 11, no. 1, p. 376, 2010.

[202] K. Wanichthanarak, S. Fan, D. Grapov, D. K. Barupal, and O. Fiehn, "Metabox: A toolbox for metabolomic data analysis, interpretation and integrative exploration," *PLoS One*, vol. 12, no. 1, p. e0171046, 2017.

[203] C. Williamson and T. Cluse-Tolar, "Pimp-controlled prostitution: Still an integral part of street life," *Violence against women*, vol. 8, no. 9, pp. 1074–1092, 2002.

[204] L. P. de Souza, T. Naake, T. Tohge, and A. R. Fernie, "From chromatogram to analyte to metabolite. how to pick horses for courses from the massive web-resources for mass spectral plant metabolomics," *GigaScience*, 2017.

[205] M. Rasmussen, Y. Li, S. Lindgreen, J. S. Pedersen, A. Albrechtsen, I. Moltke, M. Metspalu, E. Metspalu, T. Kivisild, R. Gupta *et al.*, "Ancient human genome sequence of an extinct palaeo-eskimo," *Nature*, vol. 463, no. 7282, p. 757, 2010.

[206] Z. Lai, H. Tsugawa, G. Wohlgemuth, S. Mehta, M. Mueller, Y. Zheng, A. Ogiwara, J. Meissen, M. Showalter, K. Takeuchi *et al.*, "Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics," *Nature methods*, vol. 15, no. 1, p. 53, 2018.

[207] A. G. de la Fuente, J. Godzien, M. F. López, F. J. Rupérez, C. Barbas, and A. Otero, "Knowledge-based metabolite annotation tool: Ceu mass mediator," *Journal of pharmaceutical and biomedical analysis*, vol. 154, pp. 138–149, 2018.

[208] J. A. Winchester and P. A. Floyd, "Geochemical discrimination of different magma series and their differentiation products using immobile elements," *Chemical geology*, vol. 20, pp. 325–343, 1977.

[209] K. Dührkop, H. Shen, M. Meusel, J. Rousu, and S. Böcker, "Searching molecular structure databases with tandem mass spectra using csi: Fingerid," *Proceedings of the National Academy of Sciences*, vol. 112, no. 41, pp. 12 580–12 585, 2015.

[210] J. Wandy, Y. Zhu, J. J. van der Hooft, R. Daly, M. P. Barrett, and S. Rogers, "Ms2lda. org: web-based topic modelling for substructure discovery in mass spectrometry," *Bioinformatics*, vol. 34, no. 2, pp. 317–318, 2017.

[211] C. Bueschl, B. Kluger, N. K. Neumann, M. Doppler, V. Maschietto, G. G. Thallinger, J. Meng-Reiterer, R. Krska, and R. Schuhmacher, "Metextract ii: A software suite for stable isotope-assisted untargeted metabolomics," *Analytical chemistry*, vol. 89, no. 17, pp. 9518–9526, 2017.

[212] E. Kligun, B. Ostretsov, A. Titievsky, M. Farkov, S. Alamouti, and L. Brodsky, "Adaptation of the pine fungal pathogen grosmannia clavigera to monoterpenes: Biochemical mechanisms revealed by rna-seq analysis," *Forest Pathology*, vol. 47, no. 6, p. e12372, 2017.

[213] A. Lommen and H. J. Kools, "Metalign 3.0: performance enhancement by efficient use of advances in computer hardware," *Metabolomics*, vol. 8, no. 4, pp. 719–726, 2012.

[214] F. Allen, A. Pon, M. Wilson, R. Greiner, and D. Wishart, "Cfm-id: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra," *Nucleic acids research*, vol. 42, no. W1, pp. W94–W99, 2014.

[215] D. J. Creek, A. Jankevics, K. E. Burgess, R. Breitling, and M. P. Barrett, "Ideom: an excel interface for analysis of lc–ms-based metabolomics data," *Bioinformatics*, vol. 28, no. 7, pp. 1048–1049, 2012.

[216] A. Alonso, A. Julià, A. Beltran, M. Vinaixa, M. Díaz, L. Ibañez, X. Correig, and S. Marsal, "Astream: an r package for annotating lc/ms metabolomic data," *Bioinformatics*, vol. 27, no. 9, pp. 1339–1340, 2011.

[217] R. Spicer, R. M. Salek, P. Moreno, D. Cañueto, and C. Steinbeck, "Navigating freely-available software tools for metabolomics analysis," *Metabolomics*, vol. 13, no. 9, p. 106, 2017.

[218] F. Wilhelms, "Explaining the dielectric properties of firn as a density-and-conductivity mixed permittivity (decomp)," *Geophysical Research Letters*, vol. 32, no. 16, 2005.

[219] M. Rojas-Chertó, M. van Vliet, J. E. Peironcely, R. van Doorn, M. Kooyman, T. te Beek, M. A. van Driel, T. Hankemeier, and

T. Reijmers, "Metitree: a web application to organize and process high-resolution multi-stage mass spectrometry metabolomics data," *Bioinformatics*, vol. 28, no. 20, pp. 2707–2709, 2012.

[220] J. B. Jackson, M. X. Kirby, W. H. Berger, K. A. Bjorndal, L. W. Botsford, B. J. Bourque, R. H. Bradbury, R. Cooke, J. Erlandson, J. A. Estes *et al.*, "Historical overfishing and the recent collapse of coastal ecosystems," *science*, vol. 293, no. 5530, pp. 629–637, 2001.

[221] N. Sakurai, T. Narise, J.-S. Sim, C.-M. Lee, C. Ikeda, N. Akimoto, and S. Kanaya, "Uc2 search: using unique connectivity of uncharged compounds for metabolite annotation by database searching in mass spectrometry-based metabolomics," *Bioinformatics*, vol. 34, no. 4, pp. 698–700, 2017.

[222] I. Laponogov, N. Sadawi, D. Galea, R. Mirnezami, and K. A. Veselkov, "Chemdistiller: an engine for metabolite annotation in mass spectrometry," *Bioinformatics*, vol. 34, no. 12, pp. 2096–2102, 2018.

[223] F. Nicolè, Y. Guitton, E. A. Courtois, S. Moja, L. Legendre, and M. Hossaert-McKey, "Mseasy: unsupervised and untargeted gc-ms data processing," *Bioinformatics*, vol. 28, no. 17, pp. 2278–2280, 2012.

[224] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Physical review letters*, vol. 77, no. 18, p. 3865, 1996.

[225] L. F. Feick and L. L. Price, "The market maven: A diffuser of marketplace information," *Journal of marketing*, vol. 51, no. 1, pp. 83–97, 1987.

[226] M. P. Styczynski, J. F. Moxley, L. V. Tong, J. L. Walther, K. L. Jensen, and G. N. Stephanopoulos, "Systematic identification of conserved metabolites in gc/ms data for metabolomics and biomarker discovery," *Analytical chemistry*, vol. 79, no. 3, pp. 966–973, 2007.

[227] C. D. Broeckling, F. Afsar, S. Neumann, A. Ben-Hur, and J. Prenni, "Ramclust: a novel feature clustering method enables spectral-matching-based annotation for metabolomics data," *Analytical chemistry*, vol. 86, no. 14, pp. 6812–6817, 2014.

[228] L. C. Menikarachchi, S. Cawley, D. W. Hill, L. M. Hall, L. Hall, S. Lai, J. Wilder, and D. F. Grant, "Molfind: a software package enabling hplc/ms-based identification of unknown chemical structures," *Analytical chemistry*, vol. 84, no. 21, pp. 9388–9394, 2012.

[229] Y. Ma, T. Kind, D. Yang, C. Leon, and O. Fiehn, "Ms2analyzer: a software for small molecule substructure annotations from accurate tandem mass spectra," *Analytical chemistry*, vol. 86, no. 21, pp. 10 724–10 731, 2014.

[230] H. Tsugawa, T. Kind, R. Nakabayashi, D. Yukihira, W. Tanaka, T. Cajka, K. Saito, O. Fiehn, and M. Arita, "Hydrogen rearrangement rules: computational ms/ms fragmentation and structure elucidation using ms-finder software," *Analytical chemistry*, vol. 88, no. 16, pp. 7946–7958, 2016.

[231] J. Capellades, M. Navarro, S. Samino, M. Garcia-Ramirez, C. Hernandez, R. Simo, M. Vinaixa, and O. Yanes, "george: a computational tool to detect the presence of stable isotope labeling in lc/ms-based untargeted metabolomics," *Analytical chemistry*, vol. 88, no. 1, pp. 621–628, 2015.

[232] H. Treutler, H. Tsugawa, A. Porzel, K. Gorzolka, A. Tissier, S. Neumann, and G. U. Balcke, "Discovering regulated metabolite families in untargeted metabolomics studies," *Analytical chemistry*, vol. 88, no. 16, pp. 8082–8090, 2016.

[233] X. Domingo-Almenara, J. Brezmes, M. Vinaixa, S. Samino, N. Ramirez, M. Ramon-Krauel, C. Lerin, M. Diaz, L. Ibanez, X. Correig *et al.*, "erah: a computational tool integrating spectral deconvolution and alignment with quantification and identification of metabolites in gc/ms-based metabolomics," *Analytical chemistry*, vol. 88, no. 19, pp. 9821–9829, 2016.

[234] R. Zhou, C.-L. Tseng, T. Huan, and L. Li, "Isoms: automated processing of lc-ms data generated by a chemical isotope labeling metabolomics platform," *Analytical chemistry*, vol. 86, no. 10, pp. 4675–4679, 2014.

[235] H. Li, Y. Cai, Y. Guo, F. Chen, and Z.-J. Zhu, "Metdia: targeted metabolite extraction of multiplexed ms/ms spectra generated by data-independent acquisition," *Analytical chemistry*, vol. 88, no. 17, pp. 8757–8764, 2016.

[236] H. Salvesen, S. Carter, M. Mannelqvist, A. Dutt, G. Getz, I. a. Stefansson, M. Raeder, M. L. Sos, I. Engelsen, J. Trovik *et al.*, "Integrated genomic profiling of endometrial carcinoma associates aggressive tumors with indicators of pi3 kinase activation," *Proceedings of the National Academy of Sciences*, vol. 106, no. 12, pp. 4834–4839, 2009.

[237] S. J. McIlroy, R. H. Kirkegaard, B. McIlroy, M. Nierychlo, J. M. Kristensen, S. M. Karst, M. Albertsen, and P. H. Nielsen, "Midas 2.0: an ecosystem-specific taxonomy and online database for the organisms of wastewater treatment systems expanded for anaerobic digester groups," *Database*, vol. 2017, 2017.

[238] C. Jaeger, F. Hoffmann, C. A. Schmitt, and J. Lisec, "Automated annotation and evaluation of in-source mass spectra in gc/atmospheric pressure chemical ionization-ms-based metabolomics," *Analytical chemistry*, vol. 88, no. 19, pp. 9386–9390, 2016.

[239] J. Wills, J. Edwards-Hicks, and A. J. Finch, "Assayr: A simple mass spectrometry software tool for targeted metabolic and stable isotope tracer analyses," *Analytical chemistry*, vol. 89, no. 18, pp. 9616–9619, 2017.

[240] P. W. Jones, "St. george's respiratory questionnaire: Mcid," *COPD: Journal of Chronic Obstructive Pulmonary Disease*, vol. 2, no. 1, pp. 75–79, 2005.

[241] W. M. Edmands, L. Petrick, D. K. Barupal, A. Scalbert, M. J. Wilson, J. K. Wickliffe, and S. M. Rappaport, "compms2miner: An automatable metabolite identification, visualization, and data-sharing r package for high-resolution lc–ms data sets," *Analytical chemistry*, vol. 89, no. 7, pp. 3919–3928, 2017.

[242] S. Neumann, A. Thum, and C. Böttcher, "Nearline acquisition and processing of liquid chromatography-tandem mass spectrometry data," *Metabolomics*, vol. 9, no. 1, pp. 84–91, 2013.

[243] Y. Fukuda, T. Hayakawa, E. Ichihara, K. Inoue, K. Ishihara, H. Ishino, Y. Itow, T. Kajita, J. Kameda, S. Kasuga *et al.*, "Evidence for oscillation of atmospheric neutrinos," *Physical Review Letters*, vol. 81, no. 8, p. 1562, 1998.

[244] Y.-C. Harn, B.-H. Su, Y.-L. Ku, O. A. Lin, C.-F. Chou, and Y. J. Tseng, "Np-structurepredictor: Prediction of unknown natural products in plant mixtures," *Journal of chemical information and modeling*, vol. 57, no. 12, pp. 3138–3148, 2017.

[245] B. Zhou, J. Wang, and H. W. Ressom, "Metabosearch: tool for mass-based metabolite identification using multiple databases," *PLoS One*, vol. 7, no. 6, p. e40096, 2012.

[246] N. Kessler, F. Walter, M. Persicke, S. P. Albaum, J. Kalinowski, A. Goesmann, K. Niehaus, and T. W. Nattkemper, "Allocator: an interactive web platform for the analysis of metabolomic lc-esi-ms datasets, enabling semi-automated, user-revised compound annotation and mass isotopomer ratio analysis," *PLoS One*, vol. 9, no. 11, p. e113909, 2014.

[247] D. Shang, C. Li, Q. Yao, H. Yang, Y. Xu, J. Han, J. Li, F. Su, Y. Zhang, C. Zhang *et al.*, "Prioritizing candidate disease metabolites based on global functional relationships between metabolites in the context of metabolic pathways," *PloS one*, vol. 9, no. 8, p. e104934, 2014.

[248] R. Winkler, "Spidermass: semantic database creation and tripartite metabolite identification strategy," *Journal of Mass Spectrometry*, vol. 50, no. 3, pp. 538–541, 2015.

[249] J. Draper, D. P. Enot, D. Parker, M. Beckmann, S. Snowdon, W. Lin, and H. Zubair, "Metabolite signal identification in accurate mass metabolomics data with mzeddb, an interactive m/z annotation tool utilising predicted ionisation behaviour'rules'," *BMC bioinformatics*, vol. 10, no. 1, p. 227, 2009.

[250] K. Skogerson, G. Wohlgemuth, D. K. Barupal, and O. Fiehn, "The volatile compound binbase mass spectral database," *BMC bioinformatics*, vol. 12, no. 1, p. 321, 2011.

[251] N. Sakurai, T. Ara, M. Enomoto, T. Motegi, Y. Morishita, A. Kurabayashi, Y. Iijima, Y. Ogata, D. Nakajima, H. Suzuki *et al.*, "Tools and databases of the komics web portal for preprocessing, mining, and dissemination of metabolomics data," *BioMed research international*, vol. 2014, 2014.

[252] H. Takahashi, T. Morimoto, N. Ogasawara, and S. Kanaya, "Amdorap: Non-targeted metabolic profiling based on high-resolution lc-ms," *BMC bioinformatics*, vol. 12, no. 1, p. 259, 2011.

[253] J. Jeong, X. Shi, X. Zhang, S. Kim, and C. Shen, "An empirical bayes model using a competition score for metabolite identification in gas chromatography mass spectrometry," *BMC bioinformatics*, vol. 12, no. 1, p. 392, 2011.

[254] T. Suvitaival, S. Rogers, and S. Kaski, "Stronger findings from mass spectral data through multi-peak modeling," *BMC bioinformatics*, vol. 15, no. 1, p. 208, 2014.

[255] R. J. Weber and M. R. Viant, "Mi-pack: Increased confidence of metabolite identification in mass spectra by integrating accurate masses and metabolic pathways," *Chemometrics and Intelligent Laboratory Systems*, vol. 104, no. 1, pp. 75–82, 2010.

[256] M. Scholz and O. Fiehn, "Setup x–a public study design database for metabolomic projects," in *Biocomputing 2007*. World Scientific, 2007, pp. 169–180.

[257] E. Kenar, H. Franken, S. Forcisi, K. Wörmann, H.-U. Häring, R. Lehmann, P. Schmitt-Kopplin, A. Zell, and O. Kohlbacher,

"Automated label-free quantification of metabolites from liquid chromatography–mass spectrometry data," *Molecular & cellular proteomics*, vol. 13, no. 1, pp. 348–359, 2014.

[258] B. Bonn, C. Leandersson, F. Fontaine, and I. Zamora, "Enhanced metabolite identification with mse and a semi-automated software for structural elucidation," *Rapid Communications in Mass Spectrometry*, vol. 24, no. 21, pp. 3127–3138, 2010.

[259] E. Compe and J.-M. Egly, "Tfiih: when transcription met dna repair," *Nature reviews Molecular cell biology*, vol. 13, no. 6, p. 343, 2012.

[260] R. Bannuru, N. Natov, U. Dasi, C. Schmid, and T. McAlindon, "Therapeutic trajectory following intra-articular hyaluronic acid injection in knee osteoarthritis–meta-analysis," *Osteoarthritis and cartilage*, vol. 19, no. 6, pp. 611–619, 2011.

[261] E. Sentandreu, M. D. Peris-Diaz, S. R. Sweeney, J. Chiou, N. Munoz, and S. Tiziani, "A survey of orbitrap all ion fragmentation analysis assessed by an r metabolist package to study small-molecule metabolites," *Chromatographia*, vol. 81, no. 7, pp. 981–994, 2018.

[262] R. Wehrens, G. Weingart, and F. Mattivi, "metams: An open-source pipeline for gc–ms-based untargeted metabolomics," *Journal of Chromatography B*, vol. 966, pp. 109–116, 2014.

[263] A. Corma, "From microporous to mesoporous molecular sieve materials and their use in catalysis," *Chemical reviews*, vol. 97, no. 6, pp. 2373–2420, 1997.

[264] S. Claus, I. Puscalau-Girtu, P. Walther, T. Syrovets, T. Simmet, C. Haupt, and M. Fändrich, "Cell-to-cell transfer of saa1 protein in a cell culture model of systemic aa amyloidosis," *Scientific reports*, vol. 7, p. 45683, 2017.

[265] J. W. Schopf, "Microfossils of the early archean apex chert: new evidence of the antiquity of life," *Science*, vol. 260, no. 5108, pp. 640–646, 1993.

[266] H. Mizuno, K. Ueda, Y. Kobayashi, N. Tsuyama, K. Todoroki, J. Z. Min, and T. Toyo'oka, "The great importance of normalization of lc–ms data for highly-accurate non-targeted metabolomics," *Biomedical Chromatography*, vol. 31, no. 1, p. e3864, 2017.

[267] Y. Simon-Manso, M. S. Lowenthal, L. E. Kilpatrick, M. L. Sampson, K. H. Telu, P. A. Rudnick, W. G. Mallard, D. W. Bearden, T. B. Schock, D. V. Tchekhovskoi *et al.*, "Metabolite profiling of a nist standard reference material for human plasma (srm 1950): Gc-ms, lc-ms, nmr, and clinical laboratory analyses, libraries, and web-based resources," *Analytical chemistry*, vol. 85, no. 24, pp. 11 725–11 731, 2013.

[268] M. F. Clasquin, E. Melamud, and J. D. Rabinowitz, "Lc-ms data processing with maven: a metabolomic analysis and visualization engine," *Current protocols in bioinformatics*, vol. 37, no. 1, pp. 14–11, 2012.