Supporting Information for

CCSfind: A tool for chemically informed LC-IM-MS database building

Sangeeta Kumari[1], Tim Causon[2*]

---

[1] University of Natural Resources and Life Sciences Vienna, Core Facility Bioinformatics, Muthgasse 11, 1190 Vienna, Austria

[2] University of Natural Resources and Life Sciences Vienna, Department of Chemistry, Institute of Analytical Chemistry, Muthgasse 18, 1190 Vienna, Austria

*Corresponding author

Tel: +43 1 47654-77187
E-Mail: Tim.Causon@boku.ac.at
University of Natural Resources and Life Sciences Vienna
Department of Chemistry
Institute of Analytical Chemistry
Muthgasse 18, Vienna, 1190
Austria

S1: CCSfind step-by-step

1. Before using CCSfind, convert your raw .d datafiles to mzML using ProteoWizard with default settings (note: do not select “peak picking” for centroiding).

2. Download CCSfind, Choose path of the folder for installation inside the user directory only (e.g., Desktop, Downloads, Documents). Do not choose a default Program Files location as the directory for installation as it can make access of files difficult due to administrator rights.

3. Open **CCSfind.exe**

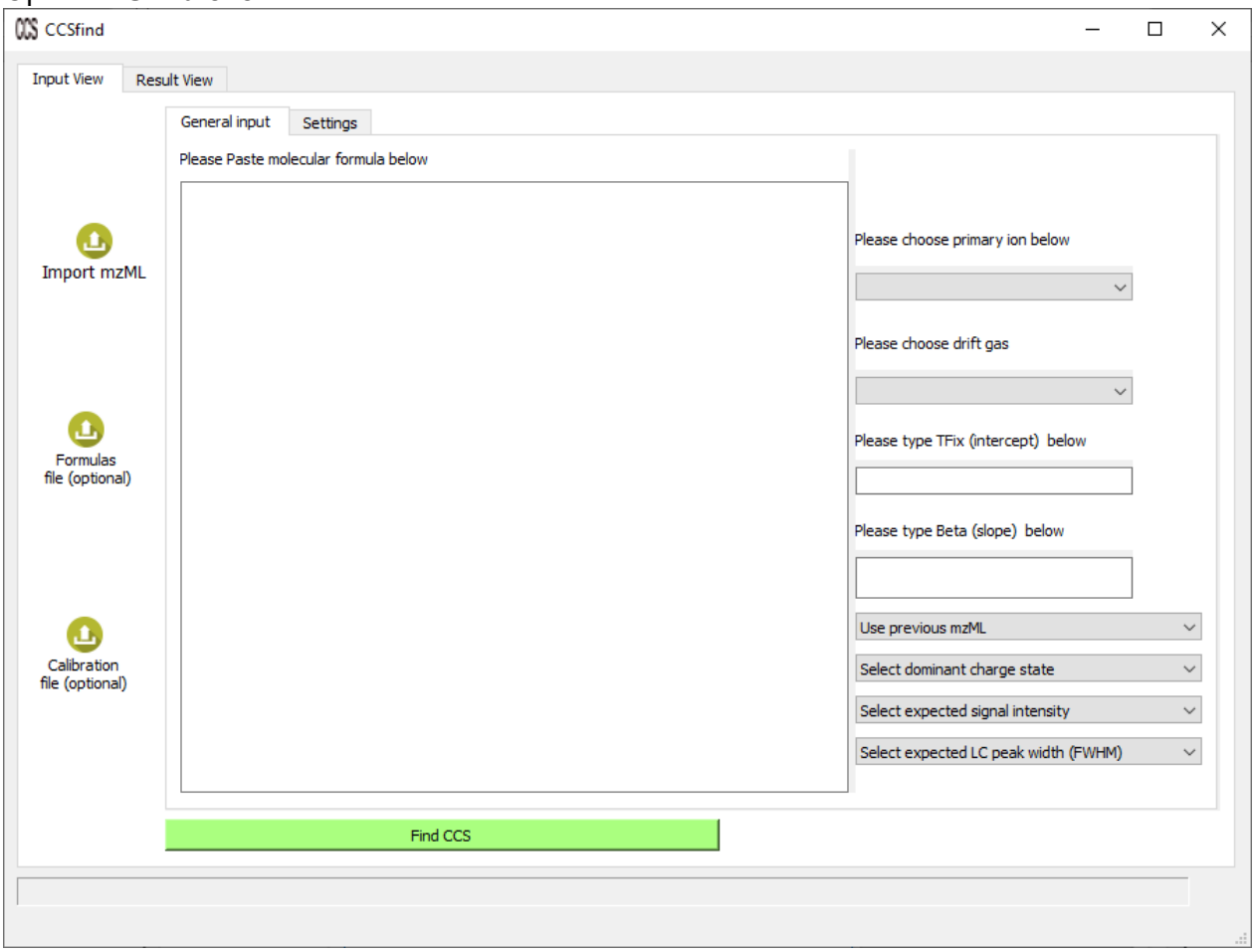


Figure S1. CCSfind start screen after opening.

4. Choose “*Import mzML*” and select your datafile.

5. Enter the appropriate settings for the file including the target molecular formulae (or use the import function), primary ion species, drift gas, single-field calibration parameters (note that these can also be imported from the raw .d file using “*OverrideImsCal.xml*” found in the *AcqData* subfolder).

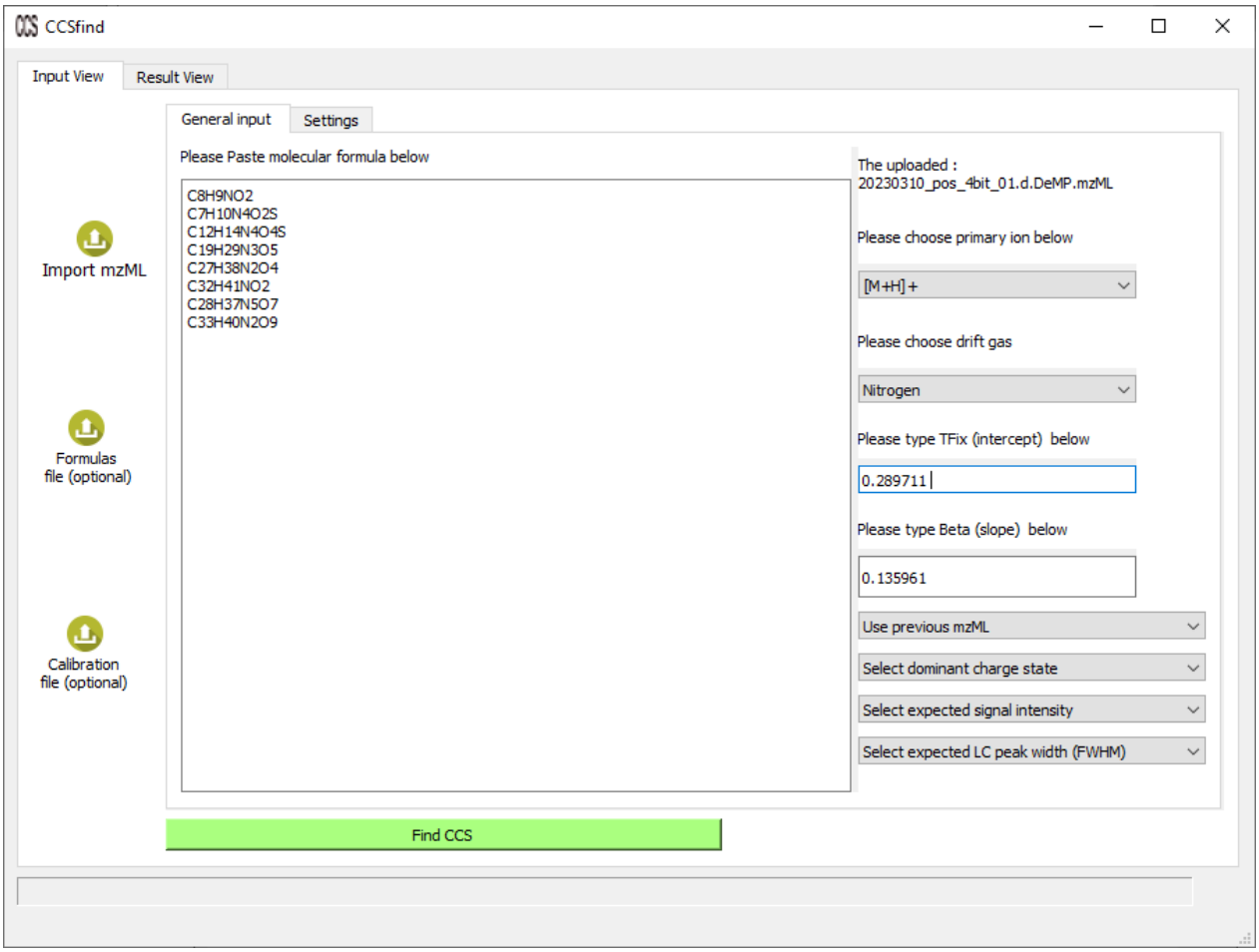


Figure S2. Exemplary general input for evaluation of positive mode LC-IM-MS measurement of a QC standard mix.

6. Additional settings (i.e., charge state, expected signal intensity, expected LC peak width) can be adjusted if the first evaluation results are not satisfactory.

7. Under the "*Settings*" tab, add any additional ion species of interest, or use the "*Formula to be added*" or "*Formula to be subtracted fields*" to search for related species (e.g., in-source fragments, solvent adducts). Note: it is not necessary to add the primary ion species again here as an additional adduct.

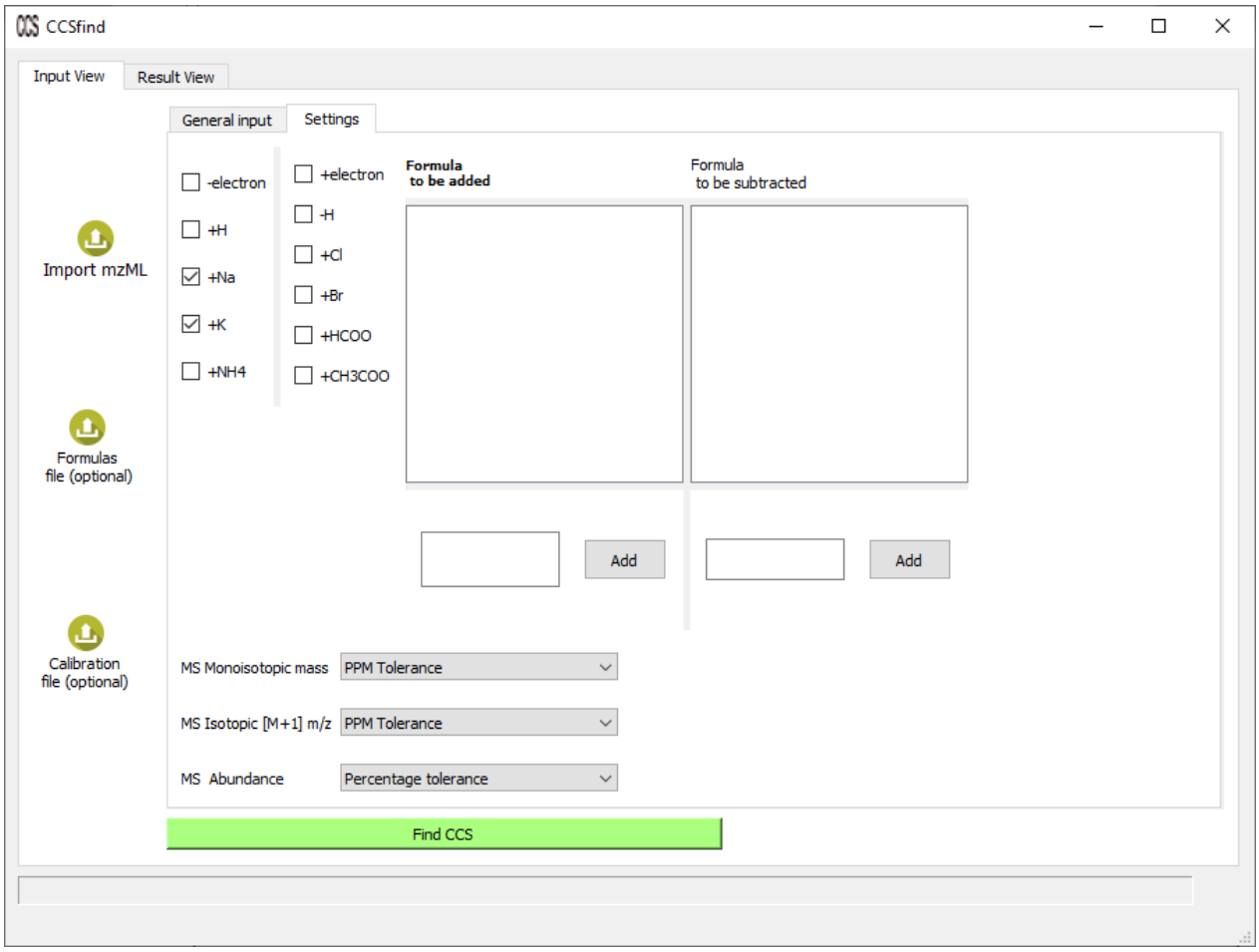


Figure S3. Exemplary settings for evaluation of positive mode LC-IM-MS measurement of a QC standard mix.

8. The "*MS Monoisotopic mass*", "*MS Isotopic [M+1] m/z*" and "*MS Abundance*" settings can be adjusted if the first evaluation results are not satisfactory.

9. Select "*Find CCS*".

10. After completion, close the dialog box and choose "*Result View*".

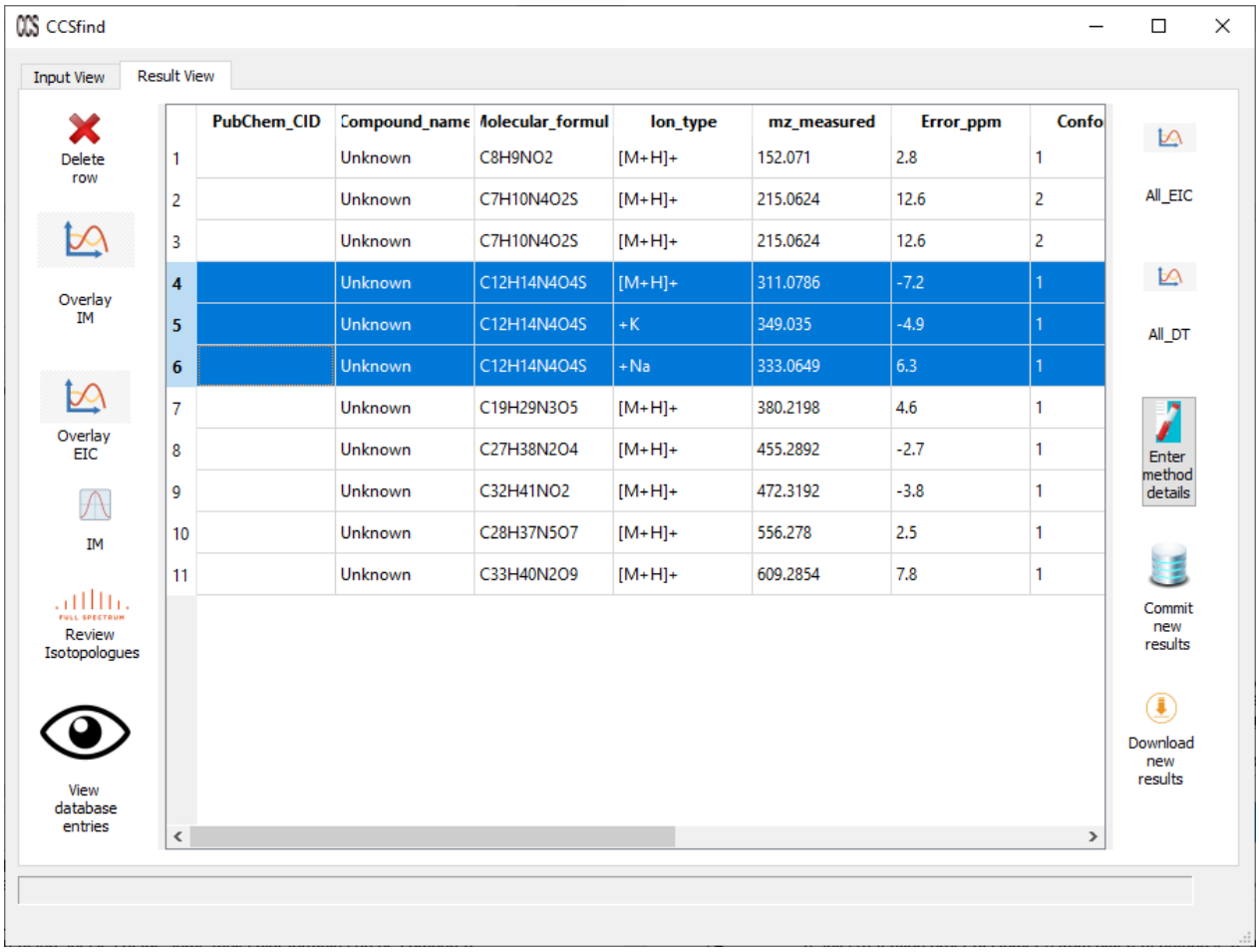


Figure S4. Exemplary results following evaluation of positive mode LC-IM-MS measurement of a QC standard mix.

11. Results of individual compounds can be reviewed with the "*Overlay*" options. EICs or IM spectra from different ion species of the same molecular formula can be compared.

12. Multi-selection of overlay IM yields a direct comparison (overlay) of results corresponding to the same target molecular formula.

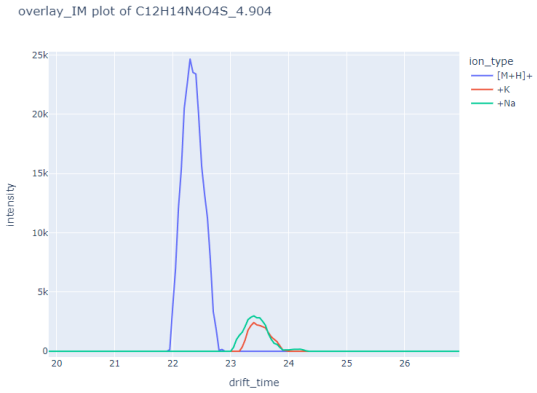


Figure S5. Overlays of IM spectra results corresponding to the input molecular formula $C_{12}H_{14}N_4O_4S$ obtained using the "*Overlay IM*" option.

13. "*Review isotopologues*" can be used to check that confirming signals are detected as co-eluting LC signals for the target molecular formula and ion species.

14. Selection of "*All EIC*" or "*All_DT*" yields a comparison of all results.

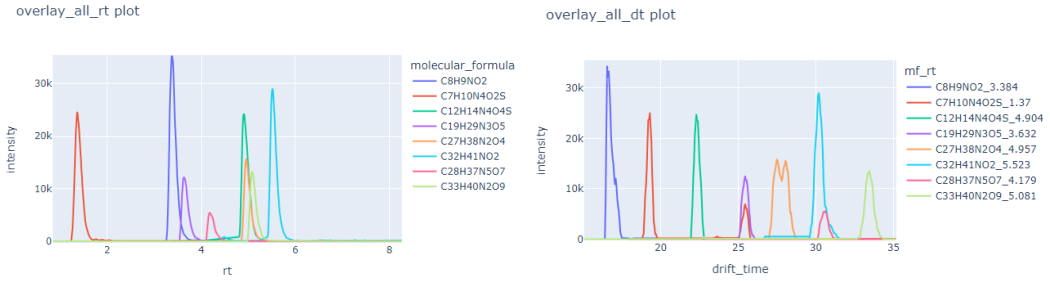


Figure S6. Results from "*All_EIC*" and "*All_DT*" options.

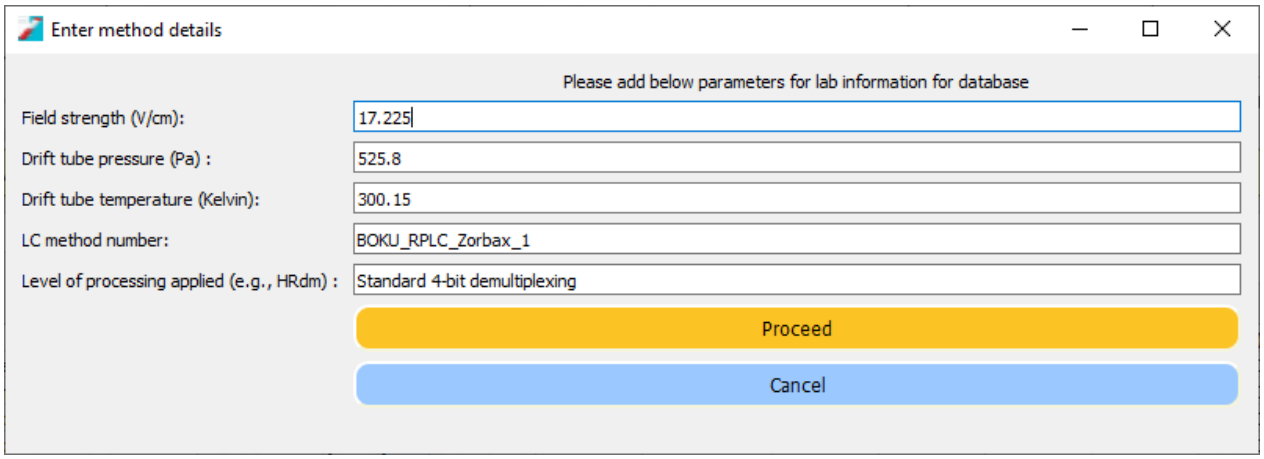
15. Incorrectly picked results can be removed with "*Delete Row*".

16. If results are not satisfactory or additional formulas need to be found, the input settings can be adjusted and the option "*Use previous mzML*" (see **Figure S1**) set to "yes" to avoid repeating the parsing of the datafile.

17. "*PubChem_CID*" and "*Compound_name*" can be updated for the confirmed results.

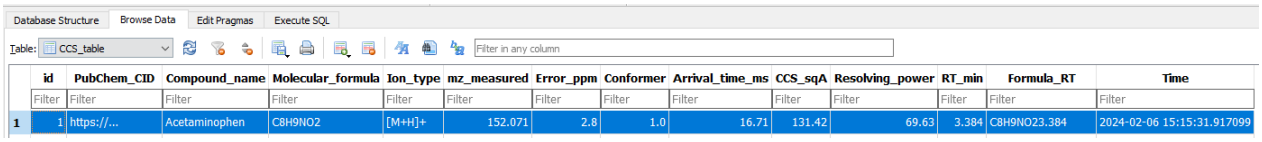
	PubChem_CID	Compound_name	Molecular_formula	Ion_type	mz_measured	Error_ppm	Conformer	Arrival_time_ms	CCS_sqA	Resolving_power	RT_min
1	1983	Acetaminophen	C8H9NO2	[M+H]+	152.071	2.8	1	16.71	131.42	69.63	3.384

18. Update the "*Enter method details*" and choose "*proceed*". Note that standardization of formatting for these fields is recommended. In addition to reporting of key IM-MS reporting parameters, curation of LC methods in a separate list can also be used to compare LC results (retention order of isomers) from different database results.



19. Commit the results to the database "*Commit new results*" and then select "*Download new results*" to the location selected by the user.

20. Review the results using a suitable tool (e.g. SQLite Browser).



	id	PubChem_CID	Compound_name	Molecular_formula	Ion_type	mz_measured	Error_ppm	Conformer	Arrival_time_ms	CCS_sqA	Resolving_power	RT_min	Formula_RT	Time
	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	1	https://...	Acetaminophen	C8H9NO2	[M+H]+	152.071	2.8	1.0	16.71	131.42	69.63	3.384	C8H9NO23.384	2024-02-06 15:15:31.917099

21. All results (EICs, IM spectra, isotopologue comparison) in HTML format can be opened directly from the results folder in a browser for viewing after downloading.