USER MANUAL

ChemCom: a Software Program for searching and Comparing Chemical Libraries

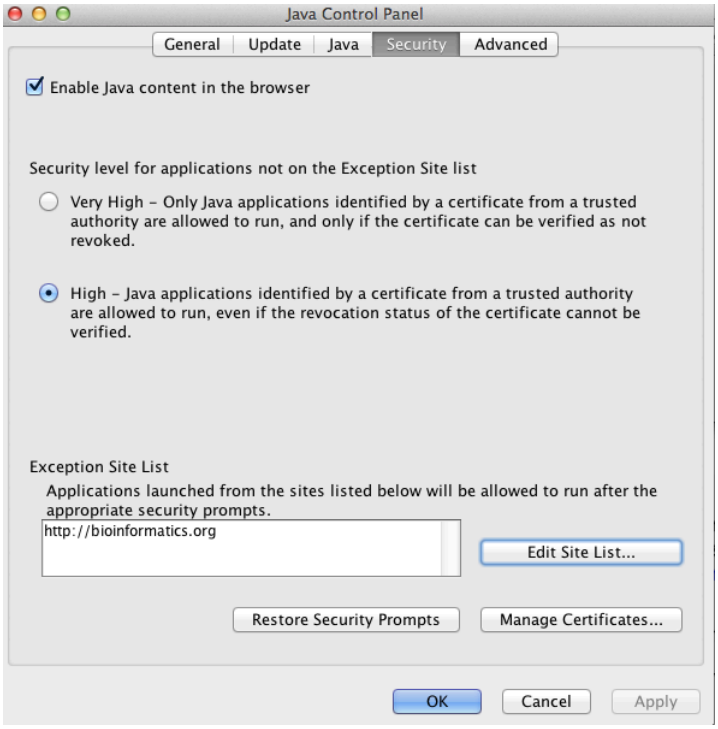
Sirus Saeedipour, David Tai, Jianwen Fang

Contact: jianwen.fang@nih.gov

Availability

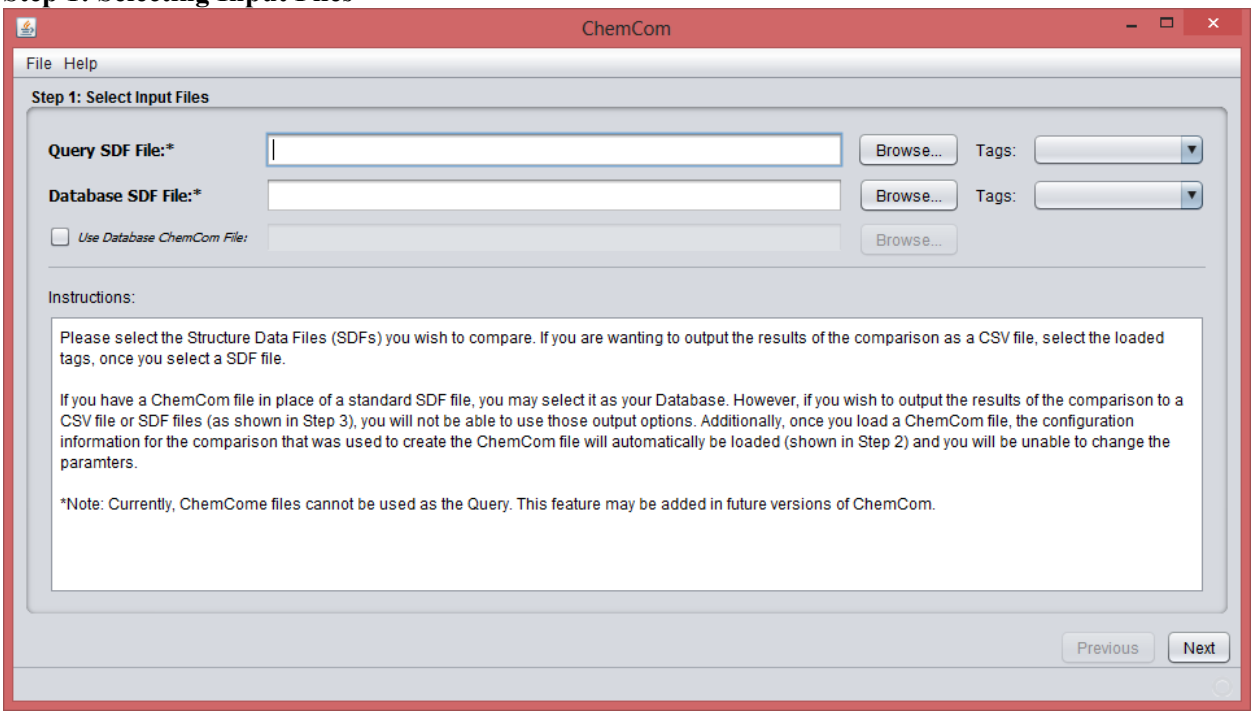
ChemCom can be run as either a JWS application via this link: **http://www.bioinformatics.org/chemcom/launchChemCom.jnlp**, or a desktop Java application by downloading the whole package here. If you choose the second option, we recommend NetBeans V7 or up to compile and execute ChemCom.

Depending on the security settings of your computer, your computer may block ChemCom from running because it is a self-signed application (see above). You may need to change Java security settings by adding bioinformatics.org to the exception site list (see below). You should save the lauchChemCom.jnlp file to your computer and then open it using right mouse.

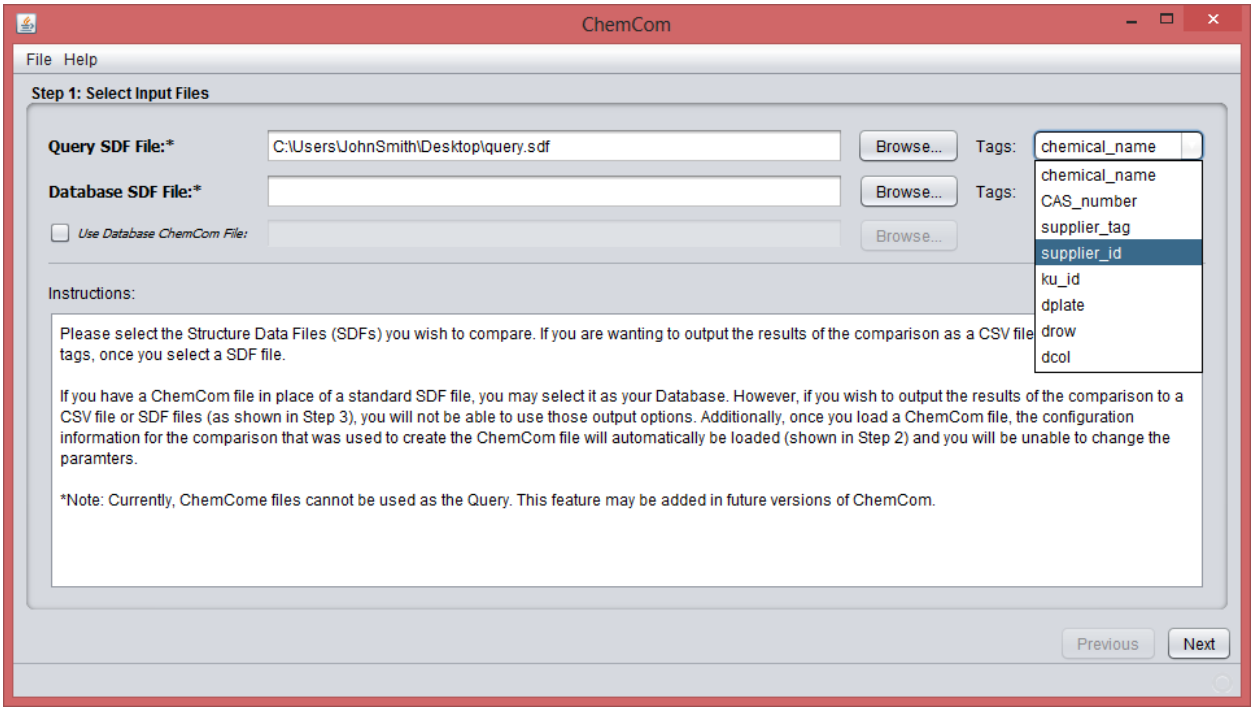


Step-by-step Execution

Step 1: Selecting Input Files



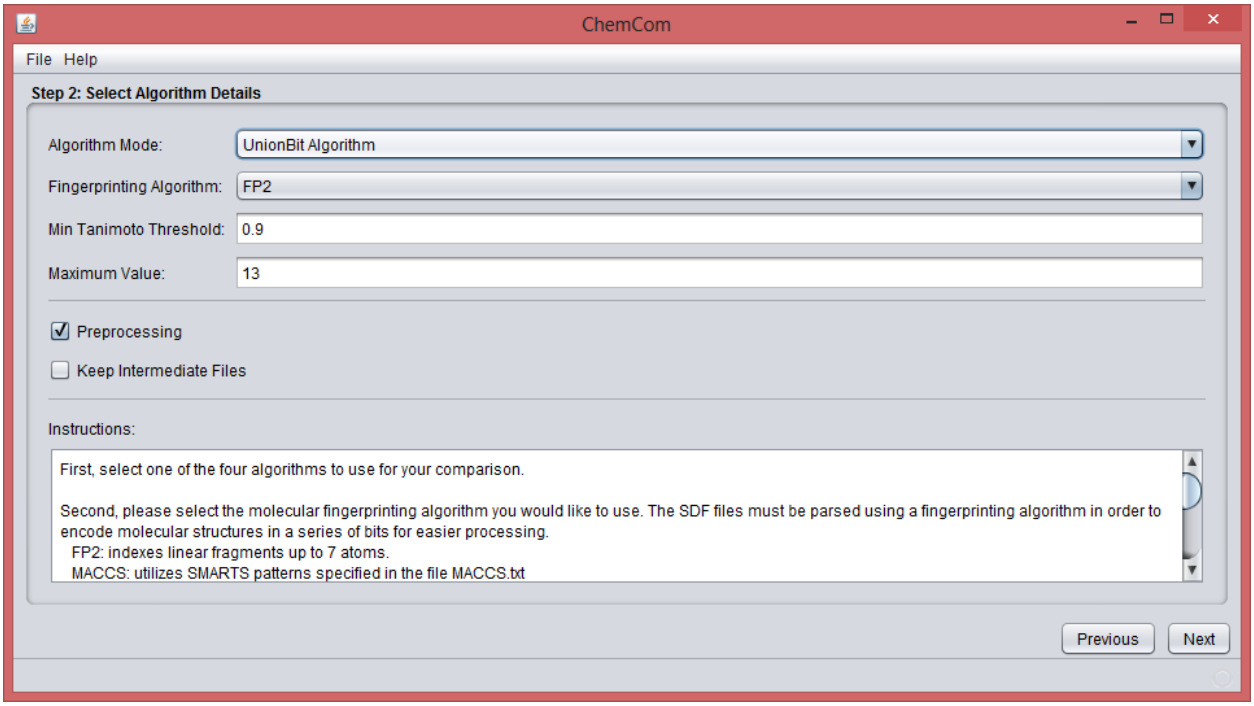
Upon opening ChemCom, you are prompted this window which presents you with *Step 1*. For this step, you must browse and select your Query and Database SDF Files.



Upon loading a SDF file, all SDF tags of the respective SDF file are displayed. Here, you need to select which tags you want your output files to use (as seen in Step 3). These tags should be unique for each chemical.

Optionally, you may choose to use a Database ChemCom file to supplement your Database SDF file. Doing so will considerably speed up your comparison at the cost of limiting your comparison options (as seen in Step 2).

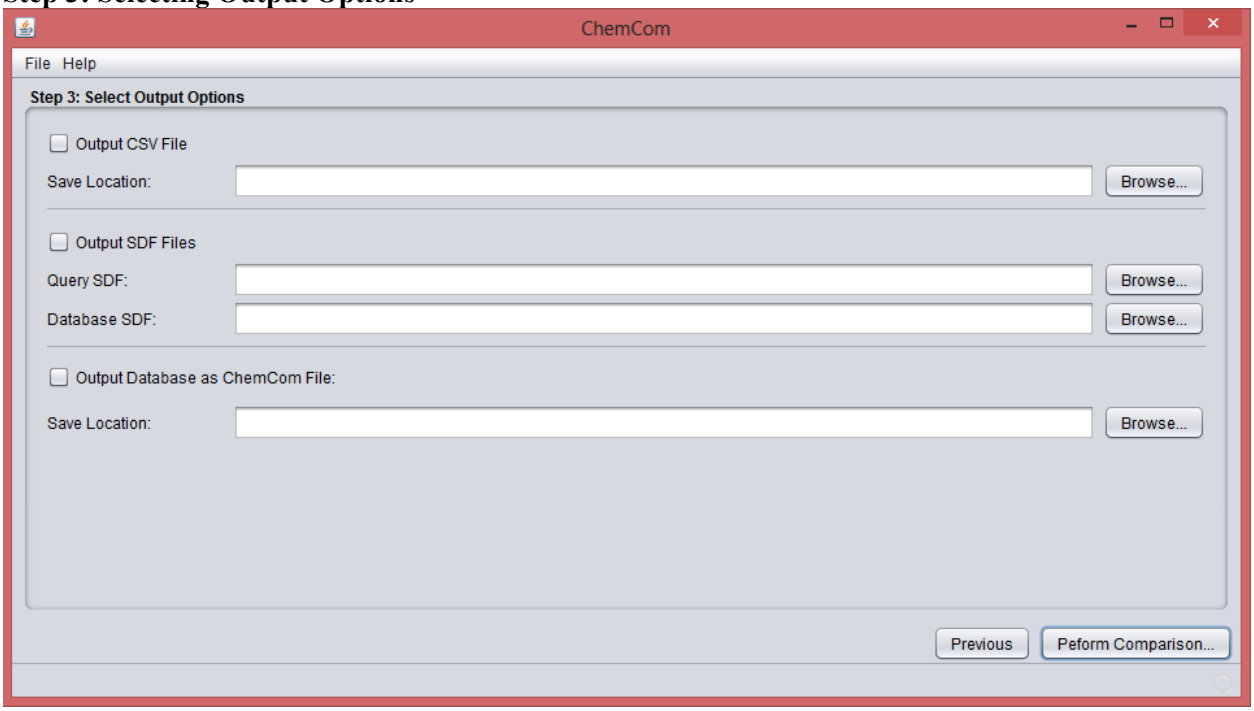
Step 2: Selecting Algorithm Details



The next step involves choosing the desired comparison algorithm along with its parameters. Additional information regarding the different fingerprinting algorithms are shown as tooltips *and* shown at the bottom of the screen. Default values are loaded for "Minimum Tanimoto Threshold" and "Maximum/k value" independent from the loaded SDF files.

- Enabling "Preprocessing" enables SymDex'd versions of the algorithms to be used for the current comparison. The description of SymDex algorithm can be found at http://pubs.acs.org/doi/pdf/10.1021/ci200606t.
- Enabling "Keep Intermediate Files" will allow ChemCom to store intermediate files that are calculated by ChemCom on the computer you are using to be reused (thus avoiding unnecessary recalculations and significantly reducing comparison times) when the same SDF files are loaded in future comparisons.

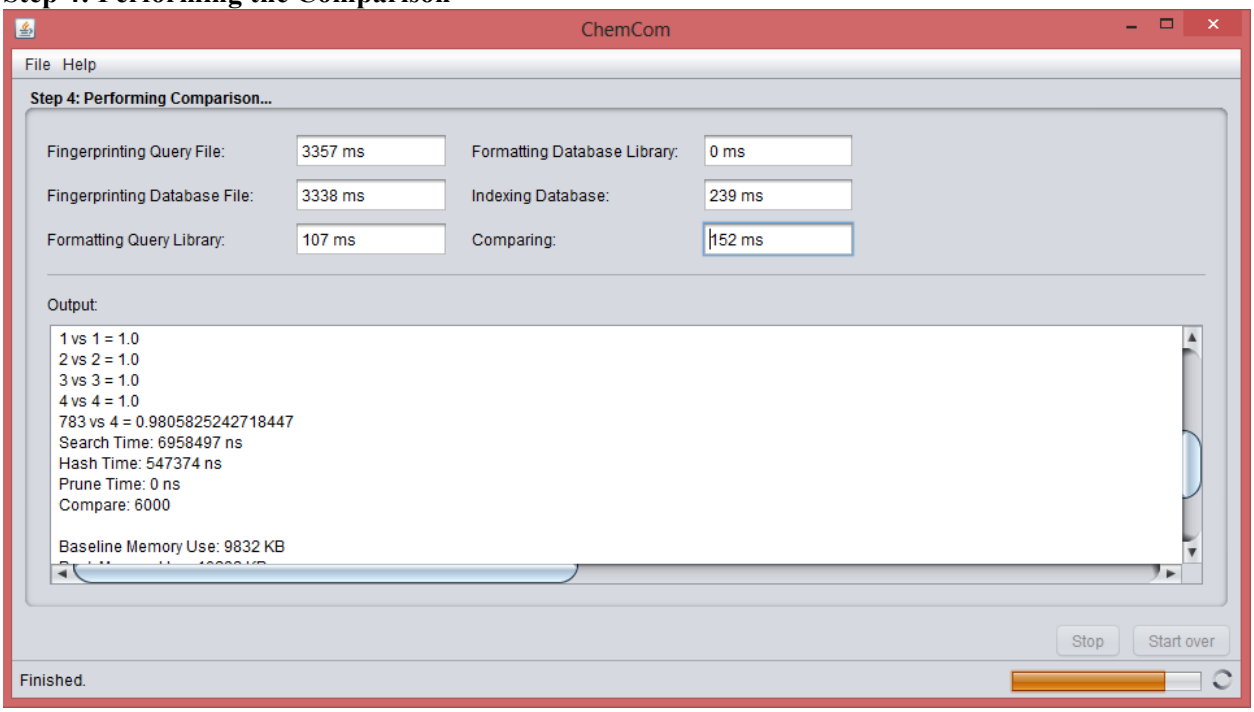
Step 3: Selecting Output Options



The last step prior to performing the actual comparison is to select locations and names for the output files. Four output files can be saved. The output CSV file has three columns: two for the query

and database chemical tags and the third is for Tanimoto scores. Each row represents a pair of chemicals with similarity above the threshold chosen in the previous step. The structures of the distinct chemicals in the CSV file can be saved in the output SDF files. In addition, the database can be saved as a binary ChemCom file. This file can be reused for searching the same database using the same algorithm and parameters in future. Thus reduce the time to run searches.

Step 4: Performing the Comparison



The last pane will show you ChemCom's progress as it performs the comparison. After each step, it will show you how long that respective process took.