# **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.

00	Java Application Blocked	
Applicatio	on Blocked by Java Security	
For securit Very High allowed to	ty, applications must now meet the requirements for the Hi security settings, or be part of the Exception Site List, to b o run. More Information.	gh oi e
Name:	chemcom.ChemComApp	
Location:	http://bioinformatics.org	
Reason:	Your security settings have blocked a self-signed applica from running	tion

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.

00	Java Control Panel			
	General Update Java Security Advanced			
Enable Java content in the browser				
Security level for	applications not on the Exception Site list			
<ul> <li>Very High - Only Java applications identified by a certificate from a trusted authority are allowed to run, and only if the certificate can be verified as not revoked.</li> </ul>				
• High - Java applications identified by a certificate from a trusted authority are allowed to run, even if the revocation status of the certificate cannot be verified.				
Exception Site Li Applications I appropriate se http://bioinform	st aunched from the sites listed below will be allowed to run after the scurity prompts. natics.org			
	Edit Site List			
	Restore Security Prompts Manage Certificates			
	OK Cancel Apply			

#### **Step-by-step Execution Step 1: Selecting Input Files**

File Help         Step 1: Select Input Files         Query SDF File:*         Browse         Database SDF File:*         Browse         Tags:         Instructions:         Please select the Structure Data Files (SDFs) you wish to compare. If you are wanting to output the results of the comparison as a CSV file, select the loc tags, once you select a SDF File.         If you have a ChemCom file in place of a standard SDF file, you may select it as your Database. However, if you wish to output the results of the comparison file, the configuration information for the comparison that was used to create the ChemCom file will automatically be loaded (shown in Step 2) and you will be unable to chang paramters.         *Note: Currently, ChemCome files cannot be used as the Query. This feature may be added in future versions of ChemCom.	□ ×
Step 1: Select Input Files         Query SDF File:*         Database SDF File:*         Database ChemCom File:         Drowse Tags:         Use Database ChemCom File:         Instructions:         Please select the Structure Data Files (SDFs) you wish to compare. If you are wanting to output the results of the comparison as a CSV file, select the lost tags, once you select a SDF file.         If you have a ChemCom file in place of a standard SDF file, you may select it as your Database. However, if you wish to output the results of the comparison file, the configuration information for the comparison that was used to create the ChemCom file will automatically be loaded (shown in Step 2) and you will be unable to change paramters.         *Note: Currently, ChemCome files cannot be used as the Query. This feature may be added in future versions of ChemCom.	
Query SDF File:*       Browse       Tags:         Database SDF File:*       Browse       Tags:         Use Database ChemCom File:       Browse       Tags:         Instructions:         Please select the Structure Data Files (SDFs) you wish to compare. If you are wanting to output the results of the comparison as a CSV file, select the loat tags, once you select a SDF file.         If you have a ChemCom file in place of a standard SDF file, you may select it as your Database. However, if you wish to output the results of the comparison file, the configuration information for the comparison that was used to create the ChemCom file will automatically be loaded (shown in Step 2) and you will be unable to change paramters.         *Note: Currently, ChemCome files cannot be used as the Query. This feature may be added in future versions of ChemCom.	
Database SDF File:*       Browse       Tags:         Use Database ChemCom File:       Browse       Browse         Instructions:       Please select the Structure Data Files (SDFs) you wish to compare. If you are wanting to output the results of the comparison as a CSV file, select the load tags, once you select a SDF file.         If you have a ChemCom file in place of a standard SDF file, you may select it as your Database. However, if you wish to output the results of the comparison information for the comparison that was used to create the ChemCom file will automatically be loaded (shown in Step 2) and you will be unable to chang paramters.         *Note: Currently, ChemCome files cannot be used as the Query. This feature may be added in future versions of ChemCom.	•
Use Database ChemCom File:  Instructions:  Please select the Structure Data Files (SDFs) you wish to compare. If you are wanting to output the results of the comparison as a CSV file, select the loat tags, once you select a SDF file.  If you have a ChemCom file in place of a standard SDF file, you may select it as your Database. However, if you wish to output the results of the comparison SCSV file or SDF files (as shown in Step 3), you will not be able to use those output options. Additionally, once you load a ChemCom file, the configuration information for the comparison that was used to create the ChemCom file will automatically be loaded (shown in Step 2) and you will be unable to chang paramters.  *Note: Currently, ChemCome files cannot be used as the Query. This feature may be added in future versions of ChemCom.	•
Instructions: Please select the Structure Data Files (SDFs) you wish to compare. If you are wanting to output the results of the comparison as a CSV file, select the load tags, once you select a SDF file. If you have a ChemCom file in place of a standard SDF file, you may select it as your Database. However, if you wish to output the results of the comparison SCSV file or SDF files (as shown in Step 3), you will not be able to use those output options. Additionally, once you load a ChemCom file, the configuration information for the comparison that was used to create the ChemCom file will automatically be loaded (shown in Step 2) and you will be unable to chang paramters. *Note: Currently, ChemCome files cannot be used as the Query. This feature may be added in future versions of ChemCom.	
Previous	n to a
Li revou	Next

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.

<u>\$</u>	ChemCom			- 🗆 ×		
File Help						
Step 1: Select Input Files						
Query SDF File:* Database SDF File:* Use Database ChemCom File:	C:\Users\JohnSmith\Desktop\query.sdf	Browse Browse	Tags: Tags:	chemical_name chemical_name CAS_number supplier_tag		
Instructions:	ata Eilee (SDEe) you wish to compare. If you are wanting to output the results of the	comparison as	a CSV fil	supplier_id ku_id dplate drow		
Please select the Structure Data Files (SDFs) you wish to compare. If you are wanting to output the results of the comparison as a CSV file drow dcol If you have a ChemCom file in place of a standard SDF file, you may select it as your Database. However, if you wish to output the results of the comparison to a CSV file or SDF files (as shown in Step 3), you will not be able to use those output options. Additionally, once you load a ChemCom file, the configuration information for the comparison that was used to create the ChemCom file will automatically be loaded (shown in Step 2) and you will be unable to change the paramters. *Note: Currently, ChemCome files cannot be used as the Query. This feature may be added in future versions of ChemCom.						
				Previous		
				Ó		

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**

<b>\$</b>	ChemCom – 🗖	×		
File Help				
Step 2: Select Algorithm De	tails			
Algorithm Mode:	UnionBit Algorithm	•		
Fingerprinting Algorithm:	(FP2			
Min Tanimoto Threshold:	0.9			
Maximum Value:	13			
Preprocessing  Keep Intermediate Files  Instructions:				
First, select one of the four algorithms to use for your comparison. Second, please select the molecular fingerprinting algorithm you would like to use. The SDF files must be parsed using a fingerprinting algorithm in order to encode molecular structures in a series of bits for easier processing. FP2: indexes linear fragments up to 7 atoms. MACCS: utilizes SMARTS patterns specified in the file MACCS.txt				
	Previous	Next		

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 <a href="http://pubs.acs.org/doi/pdf/10.1021/ci200606t">http://pubs.acs.org/doi/pdf/10.1021/ci200606t</a>.
- 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.

ChemCom	×
	Browse
	Browse
	Browse
	Browse
	Previous Peform Comparison
	ChemCom

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.

<u>چ</u>		ChemCom		- <b>-</b> ×
File Help				
Step 4: Performing Comparison				
Fingerprinting Query File:	3357 ms	Formatting Database Library:	0 ms	
Fingerprinting Database File:	3338 ms	Indexing Database:	239 ms	
Formatting Query Library:	107 ms	Comparing:	152 ms	
Output:				
1 vs 1 = 1.0 2 vs 2 = 1.0 3 vs 3 = 1.0 4 vs 4 = 1.0 783 vs 4 = 0.980582524271844 Search Time: 6958497 ns Hash Time: 547374 ns Prune Time: 0 ns Compare: 6000	7			
Baseline Memory Use: 9832 KB				
				Stop Start over
Finished.				0

## **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.