



Oct 22, 2021

Bioinformatics in Sudan-Publications

Sabah Ibrahim¹, Sofia Ali¹, Sumaya Kambal¹

¹(Bioinformatics and Biostatistics department:National University Biomedical Research Institute:National University:Khartoum:Sudan)

1



dx.doi.org/10.17504/protocols.io.bf4njqve



Sabah Ibrahim

This protocol shows in details how to extract Bioinformatics-related Publications from Sudan. It consists of six sections: PubMed Database Advance Search Bulider, Science Direct Database Advance Search Builder, Citation Manager, Exporting to Excel, Inclusion and exclusion Criteria and Data analysis.

DOI

dx.doi.org/10.17504/protocols.io.bf4njqve

<https://doi.org/10.1371/journal.pcbi.1009462>

Sabah Ibrahim, Sofia Ali, Sumaya Kambal 2021. Bioinformatics in Sudan-Publications. **protocols.io**

<https://dx.doi.org/10.17504/protocols.io.bf4njqve>



protocol

Mohamed SB, Kambal S, Ibrahim SAE, Abdalwhab E, Munir A, Ibrahim A, Ali QM (2021) Bioinformatics in Sudan: Status and challenges case study: The National University-Sudan. PLoS Comput Biol 17(10): e1009462. doi: [10.1371/journal.pcbi.1009462](https://doi.org/10.1371/journal.pcbi.1009462)

Publication, PubMed, Science direct, Literature review, Bioinformatics, Sudan

protocol ,

May 07, 2020

Oct 22, 2021

36718

The time for completing this protocol could differ because it uses Online database and thus may be affected by the internet speed.

PubMed Database Advance Search Builder

- 1 Open PubMed Advance search builder (<https://pubmed.ncbi.nlm.nih.gov/advanced/>)
- 2 Start building your search by adding Sudan (the country under subject) to the query box under the **"Affiliation field"**, then use the Boolean Operators (Add with AND).
- 3 To include papers from 2003 to 2020, under the Date publication field write "2003/01/01" to "present" (at the time writing this paper 2020 was not ended yet). Then use the Boolean operator (Add with AND) and click "add to history".
The outcome will be the following search query:
 - ((Sudan[Affiliation]) AND (("2003/01/01"[Date - Publication] : "3000"[Date - Publication])))
- 4 Add the query search you created in step 3 to the query box by clicking the **"Action button"** then choose **"add query"**.
- 5 Add the query terms ("Next generation sequencing", "Genomics", "Bioinformatics", "Sequencing", "Computational Biology" and "in silico") under **"All fields"** each one at a time (separated from the other) to the query box, then click **"add to history"** (or search if you want to move directly to the search results). This will result in the following search queries:
 - ((Sudan[Affiliation]) AND (("2003/01/01"[Date - Publication] : "3000"[Date - Publication]))) AND (Next generation Sequencing)
 - ((Sudan[Affiliation]) AND (("2003/01/01"[Date - Publication] : "3000"[Date - Publication]))) AND (Genomics)
 - ((Sudan[Affiliation]) AND (("2003/01/01"[Date - Publication] : "3000"[Date - Publication]))) AND (Bioinformatics)
 - ((Sudan[Affiliation]) AND (("2003/01/01"[Date - Publication] : "3000"[Date - Publication]))) AND (Sequencing)
 - ((Sudan[Affiliation]) AND (("2003/01/01"[Date - Publication] : "3000"[Date - Publication]))) AND (Computational Biology)
 - ((Sudan[Affiliation]) AND (("2003/01/01"[Date - Publication] : "3000"[Date - Publication]))) AND (In Silico)
- 6 Go to the search page of each **"query search"**, choose sorted by **"Best match"** and then click **"send to citation Manager"**, select **"All results"** and then **"create file"**. Save all the files

to a proper repository.

Science Direct Database Advance Search Builder

- 7 Go to the Sience direct advance search (<https://www.sciencedirect.com/search>).
- 8 Start Buliding your seach according to the below sub-steps
 - 8.1 Add the query terms (mentioned in step 4) each one at a time to the "**Find articles with these terms**" Field.
 - 8.2 Then under the **Year** field write "2003-2020"
 - 8.3 Write Sudan under the "**Author Affiliation**" Field. Finallyt, click on "**search**".
- 9 In the search results page, select **ALL RESULTS** and then click on export "**export citation to RIS**" (or any other choice you prefer). Save the results to a proper repository.

Citation Manager

- 10
 - Choose your favourite Citation Manager (we used **EndNote x9.3.3"Bld 13966"**) and upload all the resulted files from steps 5 and 8 to EndNote library "Bioinfromatics_in_Sudan_All_Search_terms.enl"
- 11 Remove all the duplicated articles (there is an option in EndNote that can "**find duplication**", then you can compare all duplicated and remove the similar ones).

Notice that: some articles have different versions and it can include Journals as full names or as abbreviation and EndNote may fail to discard these duplicated articles, thus you need to complete the rest manually.

Exporting to Excel

- 12 To export the results to an excel sheet Change the Output style of the atilcles according to the following Sub-steps:

- 12.1 Move to the following path: **Edit/Output styles/New style/Bibliography/Templates**
- 12.2 Open **"Reference types"** and choose **"Journal Article"**
- 12.3 Click on **"Insert field"** and choose the following in order:
 - **Year+Tab+Title+Tab+Author+Tab+Journal+ Tab+Volume+ Tab+Issue+ Tab+ DOI.**
- 12.4 Exit the window and Save the new style as (Bioinformatics_in_Sudan_All_Search_terms_excel).
- 12.5 Open **"Edit/ Find and replace"** , a window will appear click on **"any field"** then click on **"search"** and **"enter special carriage return"**. Add a **"semicolon plus a space"** to the **"change text"** tool and just **click ok.**
- 12.6 Make a copy from the Original file "Bioinformatics_in_Sudan_All_Search_terms" and then click on **"export references"** and save the file as a **text file (*.txt)**, then under **"Output style"** choose the style you created in the above step.
- 12.7 Open a Blank workbook Excel spreadsheet and click on **File/Open** and then choose the text file you generated in the above step, a **"Text import wizard"** window will appear, choose the file type as delimited and then leave every option as default, click next and finish. All results will appear in excel.

Inclusion and exclusion Criteria

2w

2w

13



Open each article alone and apply the following inclusion and exclusion criteria, you need to focus in each article to extract the required information. These are listed below:

- 13.1 Choose articles that are applied to all the below:
 1. Written in English
 2. Authored by at least one scientist affiliated with a Sudanese institution (Be aware that "Sudan" is mentioned as affiliation because some authors are named "Sudan")

3. Incorporating bioinformatics techniques according to the definition by authors [Luscombe et al. 2001].

- *Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? A proposed definition and overview of the field. Methods of information in medicine. 2001;40(04):346-58.*

13.2 Discard any articles that consist at least one of the followings:
Affiliated with South Sudan (now a separate country from Sudan since 2011),
Available only as abstracts and thus we were not able to confirm whether a
Bioinformatics tool was used or not
Review papers (40 papers) due to the lack of a Bioinformatics tool application.

14 Save the final output of the articles and start the data analysis

Data Analysis

3d 0h 10m

15 Order the articles by years and count the **frequency of articles per year**, list them in an excel sheet to create the figure.

16 Categorize each article to a research area using the **article title, keywords and the Journal scope**, then count the **frequency of each research area**, list them in an excel sheet to create the figure.

3d