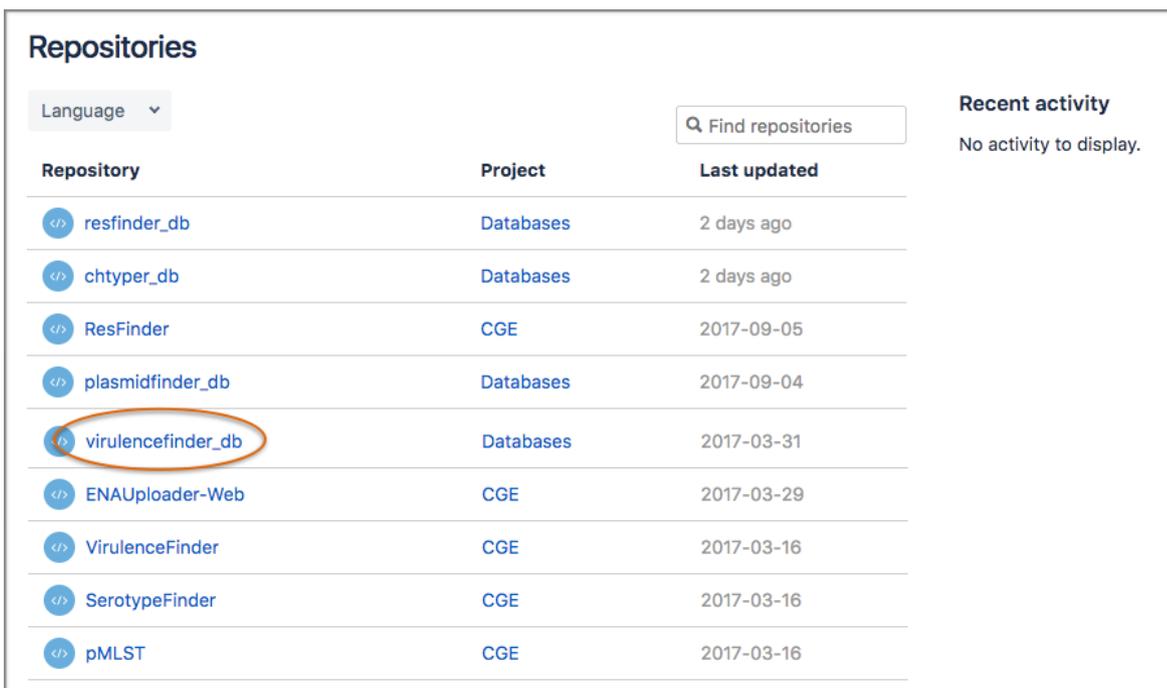


Databases on BitBucket

BitBucket is a web-based hosting service meant for storing primarily computer software (scripts). It enables having one common location (it is called a repository), where the software is stored. From this location all members in a team can download (it is called “pull”) the latest version, work on it on their own computer, and afterwards upload (it is called “push”) the changes to the common repository. Bitbucket further allows you to track changes and go back to older versions. Center for Genomic Epidemiology uses BitBucket both for storing the software behind the different web-services as well as for storing the databases.

For an overview of the repositories from CGE see: <https://bitbucket.org/genomicepidemiology/>. In the column with the header “Project”, databases are marked with “Databases”. Note that the column “Last updated” tells you when the database was last updated.

You can download the data from the databases manually, which is a bit tedious, but possible. For this option, click on the name of the database you want to download, e.g., virulencefinder_db :



The screenshot shows the BitBucket 'Repositories' page. At the top left, there is a 'Language' dropdown menu. To the right is a search box labeled 'Find repositories'. On the far right, there is a 'Recent activity' section with the text 'No activity to display.' Below these elements is a table with three columns: 'Repository', 'Project', and 'Last updated'. The table lists several repositories, with 'virulencefinder_db' highlighted by an orange circle.

Repository	Project	Last updated
resfinder_db	Databases	2 days ago
chtyper_db	Databases	2 days ago
ResFinder	CGE	2017-09-05
plasmidfinder_db	Databases	2017-09-04
virulencefinder_db	Databases	2017-03-31
ENAUuploader-Web	CGE	2017-03-29
VirulenceFinder	CGE	2017-03-16
SerotypeFinder	CGE	2017-03-16
pMLST	CGE	2017-03-16

From there click on “Source”

Genomic Epidemiology / Databases / virulencefinder_db

Overview

git clone https://bitbucket.org/genor

Last updated	2017-03-31	0	3
Access level	Read	Open PRs	Watchers
		1	0
		Branch	Forks

THERE ISN'T A README YET
A README file tells people where to start and how to contribute.

Recent activity

- 1 commit Pushed to genomicepidemiology/virulencefinder_db
bbcc0fe Update of config
Rosa Lundbye Allesee · 2017-03-31
- 1 commit Pushed to genomicepidemiology/virulencefinder_db
1588329 Update of available databases
Rosa Lundbye Allesee · 2017-03-31
- 1 commit Pushed to genomicepidemiology/virulencefinder_db
91527be Update of config
Rosa Lundbye Allesee · 2017-03-31

From here you can see a list of the individual databases that together make up the VirulenceFinder database:

Genomic Epidemiology / Databases / virulencefinder_db

Source

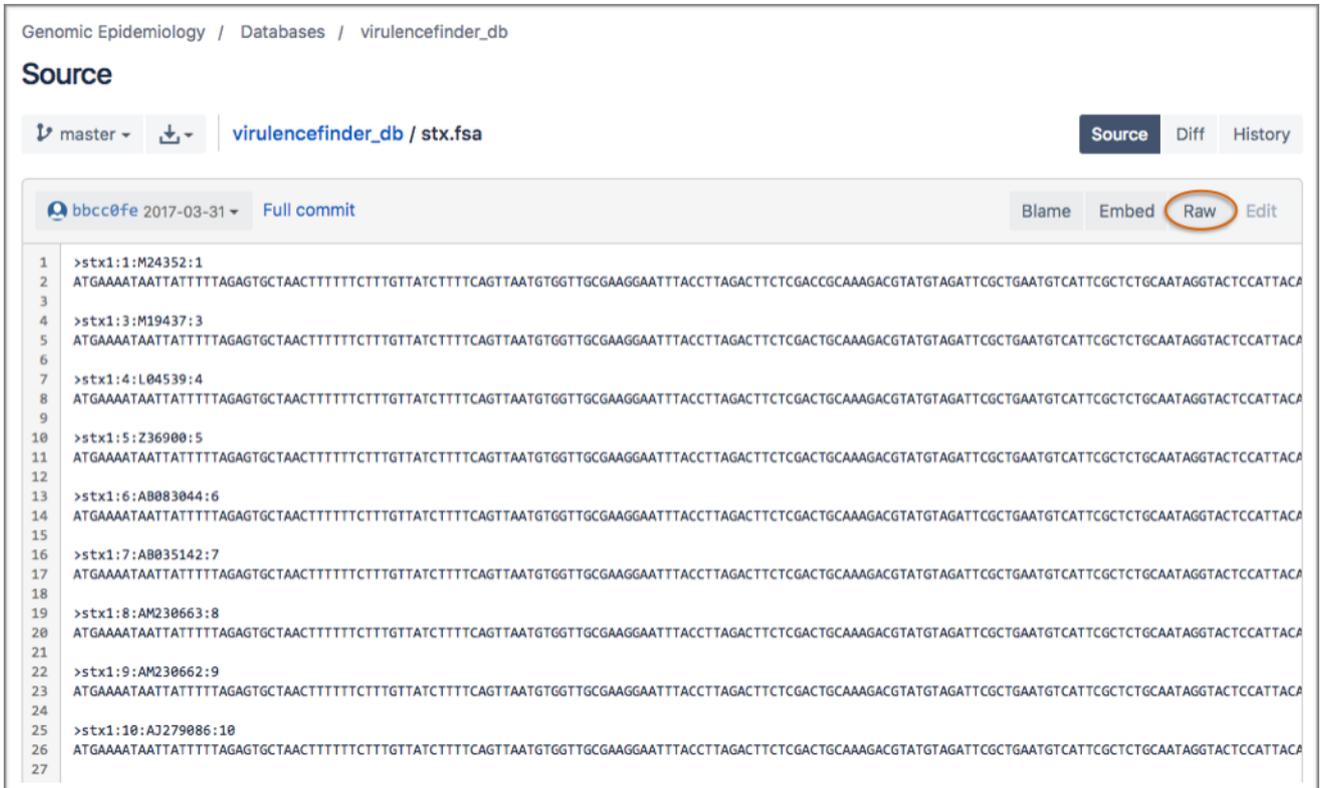
master | virulencefinder_db /

config	587 B
listeria.fsa	125.2 KB
notes.txt	17.6 KB
s.aureus_exoenzyme.fsa	39.9 KB
s.aureus_hostimm.fsa	51.2 KB
s.aureus_toxin.fsa	217.9 KB
stx.fsa	182.0 KB
virulence_ecoli.fsa	1.4 MB
virulence_ent.fsa	194.4 KB

According to the config files, the files that are related to *Escherichia coli* are virulence_ecoli.fsa and stx.fsa.

If you click on either of these files, you can see all the genes they contain in FASTA format.

To download the FASTA file, you should click “Raw”:



The screenshot shows a BitBucket repository page for 'virulencefinder_db / stx.fsa'. The page title is 'Source'. Below the title, there are navigation buttons: 'master', 'virulencefinder_db / stx.fsa', 'Source', 'Diff', and 'History'. Below these, there are buttons for 'Blame', 'Embed', 'Raw' (circled in red), and 'Edit'. The main content area displays the FASTA file content, which is a list of 10 stx genes (stx1:1 to stx1:10) with their corresponding DNA sequences. The 'Raw' button is highlighted with a red circle, indicating it is the button to click to download the file.

This will open up the file in another tab. You can now download the file by right-clicking and selecting “Save As...”

If you want to look at the content and download more databases, each of the steps have to be repeated. If you prefer a more automatic method, we recommend using the software “SourceTree”. It can be a bit tricky to set up, but once it runs, it makes keeping track of changes and downloading databases *very* convenient.

If you do not plan to use SourceTree in the future, you can skip the rest.

If you want to use SourceTree, you initially must create an account on BitBucket. It is free: <https://bitbucket.org/account/signup/>

SourceTree can be downloaded from here: <https://www.sourcetreeapp.com/> (it is also free).

Go through this online tutorial to set up SourceTree: <https://confluence.atlassian.com/get-started-with-sourcetree/install-sourcetree-847359094.html>

You should go through the steps “Install SourceTree”, “Connect your BitBucket or GitHub account”, and “Clone a remote repository” (in this case the remote repository is the CGE database you want a local copy of on your own computer, e.g., virulencefinder_db).

