

Jun 22, 2023

Version 1

Library Aligner V.1

DOI

dx.doi.org/10.17504/protocols.io.5qpvo3n89v4o/v1

Colin Kremitzki¹

¹FIVE@MGI



Colin Kremitzki

Washington University School of Medicine

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS



DOI: <https://dx.doi.org/10.17504/protocols.io.5qpvo3n89v4o/v1>

Protocol Citation: Colin Kremitzki 2023. Library Aligner. **protocols.io**
<https://dx.doi.org/10.17504/protocols.io.5qpvo3n89v4o/v1>

License: This is an open access protocol distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: June 22, 2023

Last Modified: June 22, 2023

Protocol Integer ID: 83840

Keywords: library aligner in fivtool, library aligner protocol, library aligner, fivtool, sequencing core, sequencing data, grna library, sequencing

Disclaimer

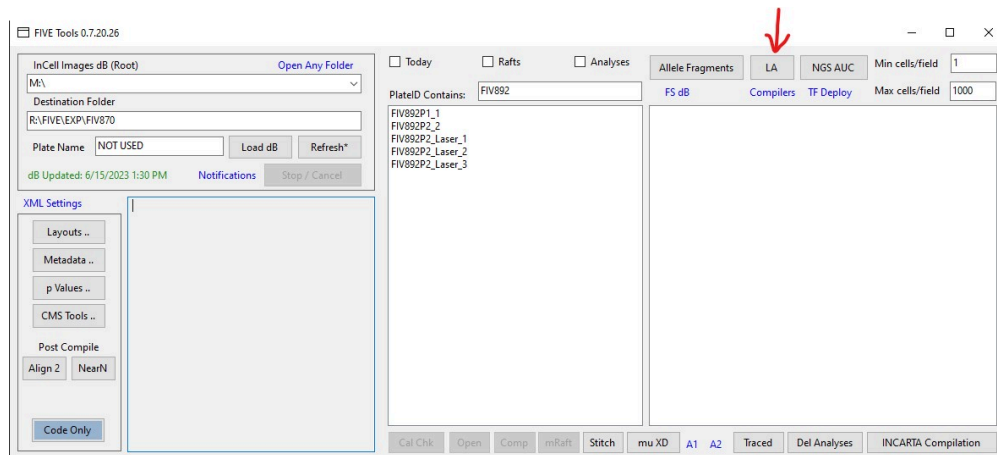
This protocol will be adjusted to the best of our ability when FIVTools is updated, but it may not match exactly.

Abstract

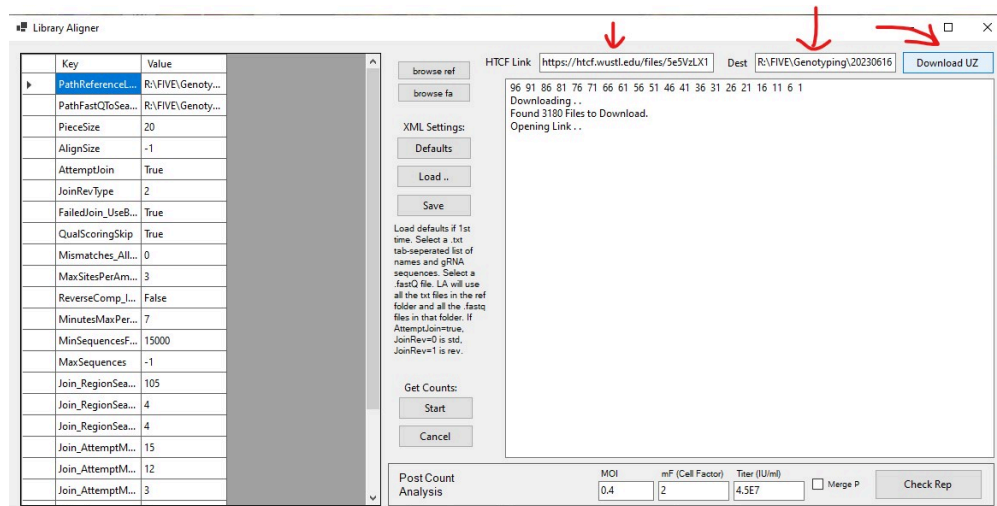
Protocol on how to run Library Aligner in FIVTools once sequencing data has been returned to us from the Sequencing Core. It also has an option for checking the Representation of a gRNA library.

Troubleshooting

- 1 Open FIVTools and click on the LA tab in the upper left section of the program.



- 2 Enter in your returned NGS link into the HTCF Link tab and the destination to download the files into Dest.



- 3 Press Download UZ to download the files and automatically unzip them. This will deposit reads into your destination folder inside folders called FastQ and FastQ_gz.
- 4 Once downloaded, you can press the button Defaults under XML Settings: to return the left criteria back to the defaults.



- 5 In the left tab for PathReferenceList, enter in the pathway which contains the reference .txt file. This file should contain the gRNA sequences that you wish to align to your reads. If you wish to align to multiple different .txt files, just list the directory and not a specific file.
- 6 In the left tab for PathFastQToSearch, enter in the pathway to your FastQ files.

Key	Value
PathReferenceList	R:\DB\Sequences\gRNA Pools\
PathFastQToSearch	R:\FIVE\Genotyping\20230616\
PieceSize	20
AlignSize	-1
AttemptJoin	True
JoinRevType	2
FailedJoin_UseBestRead	True
QualScoringSkip	True
Mismatches_Allowed	0
MaxSitesPerAmplicon	3
ReverseComp_IncomingRefSeq...	False
MinutesMaxPerIndex	7
MinSequencesForParallelAlign...	15000
MaxSequences	-1
Join_RegionSearch	105
Join_RegionSearchStep	4
Join_RegionSearchStart	4
Join_AttemptMatch_Longest	15
Join_AttemptMatch_Shortest	12
Join_AttemptMatch_Step	3

HTCF Link: Dest: Download UZ

Finished Downloading and currently UnZipping (to two separate folders). Delete the FastQ version to save space once LA is finished.

R:\FIVE\Genotyping\20230616\FastQ

XML Settings: Defaults Load .. Save

Get Counts: Start Cancel

Post Count Analysis: MOI: mF (Cell Factor): Titer (U/ml): Merge P: ☐ Check Rep

- 7 The rest of the settings can be adjusted to fit your schema.
- 8 Press Start once you are ready to align your reads to your reference sequence list.
- 9 When LA is running, you will see several things happen on your screen. You will want to be looking for the % of Unmapped Reads. If this is 100%, then your reads are not matching anything in your reference list. If you have a low % of Unmapped Reads, then that means they are hitting to reference reads, and that's a good thing.
- 10 Once finished, LA will deposit a lot of files in your destination folder. FastA files, ref_f.txt (what most people will want to look at - includes a counts table), ref_long.txt (fancier version used in Spotfire analysis), ref_TopSeq.txt (good for looking at your top hits if you get a low number of matches - answers the question: what are you matching to?).

LA_Settings.xml	6/5/2023 10:39 AM	XML Document	2 KB
ref_f.txt	6/5/2023 10:40 AM	TXT File	13 KB
ref_long.txt	6/5/2023 10:40 AM	TXT File	183 KB
ref_TopSeq.txt	6/5/2023 10:40 AM	TXT File	179 KB
Report1.txt	6/5/2023 10:40 AM	TXT File	11 KB
res_1f.txt	6/5/2023 10:40 AM	TXT File	12,405 KB

- 11 If you wish to check the Representation of the gRNA libraries you've put into cells, you can do this at the bottom of Library Aligner.

Library Aligner

Key	Value
PathReferenceList	R:\db\Sequences\gRNA Pools\
PathFastQtoSearch	R:\FIVE\Genotyping\20230616\
PieceSize	20
AlignSize	-1
AttemptJoin	True
JoinRevType	2
FailedJoin_UseBestRead	True
QualScoringSkip	True
Mismatches_Allowed	0
MaxSitesPerAmplicon	3
ReverseComp_IncomingRefSeq...	False
MinutesMaxPerIndex	7
MinSequencesForParallelAlign...	15000
MaxSequences	-1
Join_RegionSearch	105
Join_RegionSearchStep	4
Join_RegionSearchStart	4
Join_AttemptMatch_Longest	15
Join_AttemptMatch_Shortest	12
Join_AttemptMatch_Step	3

HTCF Link: <https://htcf.wustl.edu/files/Se5VzLX1> Dest: R:\FIVE\Genotyping\20230616 Download UZ

XML Settings: Defaults Load ... Save

Load defaults if 1st time. Select a txt tab-separated list of names and gRNA sequences. Select a fastQ file. LA will use all the txt files in the ref folder and all the fastq files in that folder. If AttemptJoin=true, JoinRev=0 is std, JoinRev=1 is rev.

Get Counts: Start Cancel

Post Count Analysis

MOI: 0.4 mF (Cell Factor): 2 Titer (U/ml): 4.5E7 Merge P ☐ Check Rep

- 12 Enter in the MOI you used to infect your cell line, the titer, and the Cell Factor. Cell Factor is 1 if you were using hearty cells like U2OS or HeLa cells. Cell Factor will be 2 if you were using delicate cells like iPSC cells. Select Merge P if you wish to merge SubPools together in the Representation Check, ie. MSPH1 through MSPH7 will be joined together. Press Check Rep when ready.

- 13 You can also check your representation via the following website:
<https://libalignerminiature20221013181154.azurewebsites.net/>