5	FOOD SAFETY LAB / MILK QUALITY IMPROVEMENT PROGRAM Standard Operating Procedure							
Title: Creati	ng a consensus DNA s	sequence from A	BI sequen	ce data using Sequencher				
SOP #: 8.4.5	Revision: 00	Revision Date	: n/a	Effective Date: 4/27/2020				
Authors: Ahmed Andrus	Gaballa, Jordan Skee	ns and A.	Approved	by:				

<u>Creating a consensus DNA sequence from ABI</u> <u>sequence data using Sequencher</u>

FILE NAME: 8.4.5 Creating Consensus DNA-Sequencher.docx



TABLE OF CONTENTS

1.	INTRODUCTION	3
	1.1. Purpose	
	1.2. Scope	
	1.3. Definitions	
2.	MATERIALS	4
3.	PROCEDURE	5
4.	METHOD REVIEWS & CHANGES	11



SECTION 1 - INTRODUCTION

Purpose

To create and edit consensus sequence from ABI files generated from Sanger DNA sequencing. The obtained sequence can be used for BLAST analysis, e.g., *rpoB* and *sigB* allelic typing.

Scope

This SOP may be used by all lab members. For automated editing of large number of sequences using SeqTrace, please refer to <u>8.4.1-Automated Sequence Editing</u> SOP.

Definitions

ABI (or AB1): A file produced by Sanger DNA sequencing instrument that contains sequence chromatogram (also known as electropherogram, or electrophoregram). This file is used to determine DNA bases identity (base calling).

Consensus: a single sequence derived from a set of overlapping DNA segments originating from one genetic source.



SECTION 2 - MATERIALS

- 1. Computer.
- Sequence files: Sequence files obtained from BRC facility (Biotechnology Research Center). Raw data files are in .ab1 format and edited consensus sequences are saved as .fas files.
- 3. **Sequencher:** DNA editing software. The lab has a license to use Sequencher through Cornell BRC.



SECTION 3- PROCEDURE

3.1 Installation of Sequencher:

 Download and install the latest keyserver client for your OS: Windows: <u>http://www.sassafras.com/links/K2Client.exe</u> MacOS: <u>http://www.sassafras.com/links/K2Client.pkg</u>

During the install, be sure to enter Keyserver Address : brcks.biotech.cornell.edu Please ask a lab member for the username and password (Hint: same username and password as for Geneious).

 Download Sequencher software: <u>https://www.genecodes.com/free-download</u> You may be prompted for a license file. Download and use the file from: <u>https://cornell.box.com/s/celf2wufrjt7hpnone8kf3ww9khmnnly</u>

3.2 <u>Sequence editing in Sequencher</u>:

- 1. Import Raw sequence files:
 - a. Log-in to Sequencher.
 - b. Create new project: in the toolbar, go to "File \rightarrow New Project".
 - c. To import raw files right click, or use the toolbar menu "File \rightarrow Import \rightarrow sequences", browse to the folder containing the raw sequence files.

- Asse	mble Automatically Assemble Interactiv	ely Assemble to Re	ference 💼 🗚 +A		A Sequencher				
ap placement): Min C a Qu	Nerlap = 20, Min Match = 85% Bity Kind Label	Modified	Comments		File Edit Select Assemble C New Project New Project From Template Open Project Open Recent Clore Project	ontig Sequer	view Window Help Assemble Automatically Assemble Automatically Assemble Xutomatically Kind	Assemble Intera h = 85% Label	actively Assemble to Ref
					Import Export	>	Sequences Folder Of Sequences		
	Import Get Info	> Ctrial	Sequences Folder Of Sequencer		Open Window Close Window Get Info	Ctrl+O Ctrl+W Ctrl+I	ACE Project CAF Project	,	
	Print Select All Select All Items Tha	o CHI-1 Polder Of Sequences CHI-P From Template >> All ChrI+A ACE Project All Items That >> CAP Project Bave Project As II. Save Project As II. Bave To Scame P	Save Project Save Project As Save Project As Template Revert To Saved Project	Ctrl+S	CEP Project FASTA - aligned GCG Contig Sequencher Project				
	Sort/Cleanup Show Columns Expand All Collapse All	>	CEP Project FASTA - aligned GCG Contig Sequencher Project Sequence From Vechase		Set Header & Footer Print Setup Print Print Trace in One Page	Ctrl+P	Sequence From Vecbase		

d. Select Forward and Reverse files with the .ab1 file extension (not the .seq files) and hit open.



- e. Save Project: in the toolbar menu, go to "File \rightarrow Save project, or Ctrl+S", browse to your sequence-designated folder, choose a file name and click save.
- 2. Trim ambiguous sequences:

Important note: this step may trim bases that might have been possible to call manually. If the final edited consensus sequence is shorter than the minimum required sequence for a specific application (e.g., 632bp for *rpoB*-based allelic typing of bacillales or 660bp for *sigB*-based allelic typing of *Listeria*), re-analyze your sequence with omitting this step. Attempt to base-call ambiguous sequence manually. If base calling is not possible, re-send the DNA to BRC for sequencing after troubleshooting possible issues related to DNA quality and quantity.

- a. Select all sequences.
- b. In the toolbar menu choose "Sequence \rightarrow Trim Ends Without Preview".

Sequencher	Mary Maday Hele
tet:SPF Astembly Parameters Standard Parameters (Dity Data, With ReAligner, ? of Sizettalia 668 Sizettalia 66	Create New Sequence Generate Shuffled Sequence Compare Bases To Compare Translation To Analyzes NCBI Blast Search Call Secondary Peaks Collect Gaps Insert Gaps & Move Bases Revet To Experimental Data Spit Attre Selection
	Trim Ends Trim Ends Without Preview
	Batch Revert Trim Ends Trim Vector
	Edit Features Feature Listing Mark Selection As Feature Ctrl+'
	Reference Sequence Set Circular Genome Size
	Rename Sequence Set Base Number
	Speech >

- 3. Assemble sequences:
 - a. Select both forward and reverse files for one sequence (select one file, hold Ctrl key and select the second) and click "Assemble Automatically" radio button (or from the toolbar



go to "Assemble \rightarrow Automatically").

🕍 Sequencher				
File Edit Select	Assemble Contig Sequence View Wine	dow Help		
test.SPF				
Costion 1				
Assembly Parame	ers Standard Assemble A	utomatically Asser	Assemble to Reference	A +A
Parameters: (Dirty 0	Data, With ReAligner, 3' gap placement): Min Overlap	- 20, Min Match = 85	5%	
Name	Size Quality	Kind L	abel Modified	Comments
(B) Sample 1 F	658 BPs	AutoSeq Frag	Wed, Apr 15, 2020 2:59:52 PM	
()S Sample 1 R	680 BPs	AutoSeq Frag	Wed, Apr 15, 2020 2:59:52 PM	
Sample 2 F	080 BPs	AutoSeq Frag	Wed, Apr 15, 2020 3:00:42 PM	
(E) Sample 2 R	1220 BPs	AutoSeq Frag	Thu Aug 22, 2019 2:29:00 PM	
ES Sample 3 R	1251 BPs	AutoSeq Frag	Thu, Aug 22, 2019 3:39:02 PM	
18				
1.4 a				
Sequencher				
File Edit Select	Assemble Contig Sequence View Windo	w Help		
E test.SPF	Assembly Parameters			
Assembly Paramete	Automatically	mble	Interactively Assemble to Reference	
Parameters: (Dirty D	Assemble to Reference	594		
Name	Add Selected Items To Others	abe	Modified	Comments
AS Sample 1 F			Wed, Apr 15, 2020 2:59:52 PM	•
Sample 1 R	Interactively		Wed, Apr 15, 2020 2:59:52 PM	
Sample 2 F	Mindlessly Join		Wed, Apr 15, 2020 3:00:42 PM	
Sample 2 R	Build Reference Database or Index	>	Wed, Apr 15, 2020 3:00:48 PM Thu, Aug 22, 2019 3:39:00 PM	
Sample_3_R	Alian Union		Thu, Aug 22, 2019 3:39:02 PM	
	Align Using	,		
	Align Data Files to Ref Using	>		

- b. Change the name of the Contig, right click on the Contig name and select "Rename contig" from the popup menu or from the toolbar go to "Contig → Rename contig". Choose a unique name to reflect the FSL and the gene you have sequenced.
- c. To view the forward and reverse sequences in a Contig click the + to the far-left of the contig name.
- 4. Editing sequences:
 - a. Double-click on the symbol to the left of the Contig name.

A Sequencher - [test.SPF]							
/ Eile Edit Select Asse	mble <u>C</u> ontig	Se <u>q</u> uence <u>V</u> iew	<u>W</u> indow	<u>H</u> elp			
Assembly Parameters Standard		 Assemble A 	utomatically	Assemble Interactiv	vely Assemble to Reference	1 -A +A	
Parameters: (Dirty Data, With ReA	Parameters: (Dirty Data, With ReAligner, 3' gap placement): Min Overlap = 20, Min Match = 85%						
Name	Size	Quality	Kind	Label	Modified	Comments	
E = FSL# gene name	735 BPs		Contig of 2	-	Wed, Apr 15, 2020 4:35:00 PM		

b. When the new window pops up, click the radio button "Bases" then "Show

Chromatograms".

L Sequencher - [FSL# gene name]	
Select Assemble Contig Seguence View Window Help	
Bases Jummary Sort Options Find Get Info ReAligner	
	Sample 1 R Sample 1 F
1 67	
	Hole in contig Hole in contig Diagram Keu Politipie frammet zame direction Both strands Both st
La Sequencher - [FSL# gene name]	
E File Edit Select Assemble Contig Sequence View Window Help	
Overview Summary Cut Map Find Show Chromatograms ReAligner	
42 Sample 1 R GACCIGAAGAAAATINACGCCGI GAINIIC NCCGAAIGITI 42 Sample 1 F	GGAGAAGATGCACTTCGTAACCTTTG



c. The forward and reverse sequences are at the top of the page with the consensus sequence

below them. Make edits only to the consensus sequence.



d. You can change how the bases are displayed from the "View" menu. This includes color vs monochrome "View → Display Color Bases"; shading by base confidences values "View → Display base confidences".



You can view and change confidence-shading ranges from the toolbar menu "Window \rightarrow User preferences (or Ctrl+U)" and select "confidences" in the popup window from the menu on the left.



Creating a consensus DNA sequence from ABI sequence data using Sequencher FSL/MQIP @ CORNELL UNIVERSITY

Revision

Effective 04/27/2020

Revised 04/27/2020

	Add to Connections Session			
- Assem	Open Existing Connections Session Delete Existing Connections Session	1		
3' gap placement): Min Ov Size Qual	Chromatogram	Ctrl+T	Comments	(1 ite
12 (19)	Ambiguity/Key Codes Genetic Code Motif Definitions Specify Restriction Enzymes Specify Vector Insertion Sites		Monorality Confidence Ranges X Anotania 0 10 1 1 1 0 1	
	Project Window Variance Table Translated Variance Table Chromatograms Contig Editors Sequence Editors	>	Line I fore New New New Particle Denviry Committy	
	Open Trim Log Folder Open External Data Folder Open External Data Browser		Billen stop Viriance Table ⊡ Ingent®Oxfput File Import Report ↓	
	Remember Window Layout	>		<u></u>
	Ambiguity Helper			
	User Preferences	Ctrl+U		

- e. Trim messy sequence from 5' and 3' ends by highlighting these bases in the consensus sequence and hit delete.
- f. Make edits in the consensus sequence by comparing peaks in the chromatograms. Change base editing mode (i.e., replace base vs insert base) from the toolbar menu "Edit →
 Replace When Editing or Insert When Editing".

🔬 Seq	quencher						
File	Edit Se	ect Assemble	Contig	Sequence	View	Window	Help
Ξ.	Und	0		Ctrl+Z			
+	Cut	Selection		Ctrl+X	Show Ch	romatograms	ReAligner
	Cop	y Selection		Ctrl+C			
12 F	Cop	y As		>	CACTIO	GTAACC	TTTGACGAGCGCGGTATTNATTCGCGTCG
Bi	Past	e		Ctrl+V			GNCGNGCGCGGTATT: ATTCGCGTCG
	Clea	r					
	Edit	Comments		Ctrl+,			
	Aut	o Match					
	Lab	el					
	Refr	igerate					
Ь	Dup	licate Seq Fragn	nent		20 CACTTO	GTAACC	TTTGACGAGCGCGGTATTNATTCGCGTCG
CI DI	Ente	r Selection As N	lotif		TS	. P	L T S A V ? I R V G
	Shif	t			H F	V T	F D E R G I ? S R R
	Rep	ace When Editir	g				
	Ove	rstrike When Edi	ting				
	Inse	rt When Editing					
Ľ,	Ren	iove From Proje	ct		э т	ттте	
븉	Ren	iove From Table			Ľ.	Á.	
	$\overline{\Lambda}$	<u> </u>	Λ	~^	$\Lambda /$		
م ا	$^{\vee}$	$\land \land \land$		V ~ V	IV-	$- \sum f$	
GT		- \	XX	4 X	. Λ -	X	
	m	rd /	YV.	X	\mathcal{V}	n.H	XX-AXX/VAXXAV
	× 100 -	Ne von	. <i>u</i> V	<u></u>			

<u>Hint</u>: <u>8.4.1-Automated Sequence Editing</u> SOP describes valuable background details about quality (confidence) score and its use in base calling (introduction and 3.7 sections).

g. Undo a change from toolbar menu "Edit \rightarrow Undo or Ctrl+Z".



- h. Sequencher automatically saves edits, so you only need to save your project.
- At this step, it is good idea to estimate if you have the minimum sequence length for a specific application (e.g., 632bp for *rpoB*-based allelic typing of bacillales or 660bp for *sigB*-based allelic typing of *Listeria*). However, BLAST analysis will be required to determine if you have the complete sequence of the allelic typing region.
- j. Dissolving Contigs and reverting to experimental data:
 - Undo function can be used only to revert one edit. To revert to the original sequence, select the contig from the main project window then from the toolbar menu choose "Contig → dissolve Contig". This option can also be done by right click on the contig name and select dissolve Contig from the popup menu.
 - ii. Select the individual sequence files and from the toolbar menu choose "Sequence
 - \rightarrow Revert to Experimental Data"



- k. Exporting edited Contig as a FASTA (.fas) file:
 - To export consensus, select the edited contigs. From the toolbar menu choose
 "File → Export → consensus" or right click on the contig and choose Export, then Consensus from the popup menu.



ii. In the popup window, choose Fasta from the drop menu of "Format".

Browse to your designated sequence folder then click OK.



SECTION 4- METHOD REVIEWS & CHANGES

VERSION	DATE	EDITOR	COMMENTS
Version: draft	05/30/13	A. Andrus	Draft SOP
Version 1	4/22/20	Ahmed	
(SOP		Gaballa	
8.1.1.1.21)		and Jordan	
		Skeens	