# NEBNext Immune Sequencing data analysis user guide

## **pRESTO** tool introduction

The NEBNext Immune Sequencing data analysis workflow is based on the pRESTO tool suite, the Repertoire Sequencing TOolkit. pRESTO performs all stages of raw sequence processing prior to alignment against reference germline sequences. pRESTO is flexible and customizable and is composed of multiple modules. This tutorial is meant to be concise and allow you to understand and run an example workflow that is configured for the standard NEBNext Immune Sequencing workflow. For a more detailed tutorial please consider reading the pRESTO tutorial.

In an example of an Illumina MiSeq paired-end 2x300 cycle run, Figure 1 (adapted from pResto documentation figures) shows the read schematic. Each read was sequenced from one end of the target cDNA so that the two reads together cover the entire variable region of BCR and TCR. The V(D)J reading frame proceeds from the start of read 2 to the start of read 1. Read 1 is in the opposite orientation (reverse complement), contains a partial C-region, and is 300 nucleotides in length. Read 2 contains the 5'RACE template switch site with a 17 nucleotide UMI barcode preceding it and is 300 nucleotides in length.



Figure 1. NEBNext Immune Sequencing library read schematic example of a run with Illumina MiSeq paired-end 2x300 reads.

#### pRESTO workflow overview



- 1. **Quality filter**: remove reads with mean Phred quality scores (Q) less than a user defined, sensitive value (set to 20%)
- 2. **Remove primers**: remove PCR primers and annotate reads with UMI (with maximum allowable error rate set to 20%)
- 3. **UMI consensus\***: generate consensus sequences for each UMI barcode by multiple alignment
- 4. **Sequence assembly**: assemble paired-reads on the UMI consensus into a long sequence (by default, when the reads do not intercept each other, the full fragment can be inferred using IgBlast)
- 5. **Filters**: quality-filtersequences
- 6. **Final repertoire**: obtain the final repertoire output that can be used for VDJ alignment tools like IgBlast. Only unique sequences with more than N (set to 2) representative reads are used in downstream analysis.

#### \*Troubleshooting consensus generation

Workflow parameters have been selected based on our experience with Illumina MiSeq instruments and V3 reagent kits. Since each experiment accumulates experimental errors differently, the parameters may require tuning based on the following factors.

- Filter reads if there is a high rate of mismatches between UMIs in a UMI consensus (maxerror parameter).
- Filter reads that do not share common primers within a UMI consensus group (**prcons parameter**).
- Align C-regions to a C-InternalRegions.fasta file, filtering for no more than 30% errors and limited to 100 nucleotides.
- The annotation specifying the number of raw reads used to build each sequence, is updated to the minimum of Fwd and Rev reads.
- Duplicate sequences, sharing the same constant region, are removed.
- Sequences with more than n (20) "N" are also removed.

For more UMI consensus troubleshooting details, please go to the pRESTO tool webpage for <u>UMI consensus</u> or the following reference.

# Dysregulation of B Cell Repertoire Formation in Myasthenia Gravis Patients Revealed through Deep Sequencing.

Vander Heiden JA, et al. J Immunol. 2017 198(4):1460-1473. doi:10.4049/jimmunol.1601415.

## How to use the pRESTO workflow via Galaxy

NEB has implemented a pRESTO workflow on <u>Galaxy</u> for an easy start to users of immune sequencing data analysis. Galaxy is an open source, web-based platform for data intensive biomedical research. New users to Galaxy can follow the <u>Galaxy Tours</u> to start learning the platform. The pRESTO workflow published by NEB on Galaxy can be found <u>here</u>.

1. Import files before running pRESTO workflow

Input files for running pRESTO and example dataset can be found at an <u>example history</u> published by NEB. Follow the steps below to import your own files.

1.1 Create a new history and upload fastq files into the history.



1.2 Select all of the fastq files using the check button, click the "For all selected..." button and choose <u>Build List of Dataset</u> Pairs.



1.3 Enter the filters for forward and reverse fastq files, click <u>Auto-pair</u>, type the name of the <u>collection</u>, check that each R1 and R2 are correctly paired, and click <u>Create list</u>.

Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Subcide of critical datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse read		Success a selle stimulation of mains of shares of states and state					3 0 78
Could not automatically create any pairs from the given dataset names. You may want to choose or enter different filters and try auto-pairing   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an v   Langained forward - 1 filtered out   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an v   Mouse-IO-TCR.11strg   Oppirs   Ungain all   Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an v   Oungained forward - 0 filtered out   Clear Filters   Oungained forward - 0 filtered out   Alto-pair   Zfastig   Index of the current filters.  House-IO-TCR.11stg → Mouse-IO-TCR. ← Mouse-IO-TCR.21stg Konce file extension? Konce file extension? Konce file extension? Cancel Konce file extension? Center intered intered intered intered interestion interestic interestication i		Create a collection of paired datasets				. + 1	□ \$
Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an *   Impaired forward - 1 filtered out   Clear Filters   Impaired forward - 1 filtered out   Clear Filters   Mouse-IO-TCR.1 fastq   O pairs: Unpair all    Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an * Ourpaired forward - 0 filtered out Organs: Unpair all    Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an * Ourpaired forward - 0 filtered out Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an * Ourpaired forward - 0 filtered out Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an * Ourpaired forward - 0 filtered out Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an * O unpaired forward - 0 filtered out Clear Filters O unpaired reverse - 0 filtered out Auto-pair Zestq Mouse-IG-TCR.1 fastq > Mouse-IG-TCR.4 feet Mouse-IG-TCR.2 fastq Kence file original elements? Remove file extensions? Kence file original elements? Remove file extensions? Cancel Concel Cence is the original elements?	n to						8
Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an 1       umpaired forward - 1 filtered out       Clear Filters       1 umpaired reverse - 1 filtered out         1       1       Pair these datasets       Mouse-IO-TCR.2.fastq         0       pairs       Uppair all         Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an          0       opairs       Uppair all         Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an          0       unpaired forward - 0 filtered out       Clear Filters         0       unpaired forward - 0 filtered out       Clear Filters         0       unpaired forward - 0 filtered out       Clear Filters         0       unpaired forward - 0 filtered out       Auto-pair         1       Internet filters.       0 unpaired reverse - 0 filtered out         1       Mouse-IG-TCR.1.fastq > Mouse-IG-TCR. (et Mouse-IG-TCR.2.fastq       &         Hide original elements?       Remove file extensions?       Name         Name       Mouse-IG-TCR       Create list		Could not automatically create any pairs fro again. cancel and reselect new elements.	om the given dataset names. You may w	vant to choose or ente	r different filters and try auto	o-pairing ×	
Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an v tunpaired forward - 1 filtered out Copairs Unpair all Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an v Opairs Unpair all Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an v Ounpaired forward - 0 filtered out Clear Filters Ounpaired forward - 0 filtered out Clear Filters Ounpaired forward - 0 filtered out Clear Filters Ounpaired reverse - 0 filtered out Auto-pair No datasets were found matching the current filters. Notate-IG-TCR.1fastq > Mouse-IG-TCR.2 fastq Notate-IG-TCR.1fastq > Mouse-IG-TCR.2 fastq Notate-IG-TCR.1fastq > Mouse-IG-TCR.2 fastq Name Name Name Name Notate-IG-TCR Name Name Name Name Notate-IG-TCR	•						
1 unpaired forward - 1 filtered out       Clar Filters Auto-pair       1 unpaired reverse - 1 filtered out         Wouse-IG-TCR.1fastq       Pair these datasets       Mouse-IG-TCR.2fastq         0 pairs       Unpair       Unpair         0 unpaired forward - 0 filtered out       Clear Filters Auto-pair       0 unpaired reverse - 0 filtered out         0 unpaired forward - 0 filtered out       Clear Filters Auto-pair       0 unpaired reverse - 0 filtered out         1 fasta	ta	Collections of paired datasets are ordered li	lists of dataset pairs (often forward and	d reverse reads). Thes	e collections can be passed	to tools an 🗸	
1unpaired forward - 1 filtered out       Clear Filters       1unpaired reverse - 1 filtered out         Mouse-IO-TCR.1fastq       Pair these datasets       Mouse-IO-TCR.2fastq         0 pairs       Unpair all         0 unpaired forward - 0 filtered out       Clear Filters       0 unpaired reverse - 0 filtered out         1 filterad       0 unpaired forward - 0 filtered out       Clear Filters       0 unpaired reverse - 0 filtered out         1 filterad       0       Unpair all       0 unpaired forward - 0 filtered out       Clear Filters         0 unpaired forward - 0 filtered out       Clear Filters       0 unpaired reverse - 0 filtered out       1 filtered         1 filterad	ior						
Lastq       • <td>. 1</td> <td>1 unpaired forward - 1 filtered out</td> <td>Clear Filters</td> <td></td> <td>1 unpaired reverse - 1 filte</td> <td>red out</td> <td></td>	. 1	1 unpaired forward - 1 filtered out	Clear Filters		1 unpaired reverse - 1 filte	red out	
Mouse-IG-TCR.1/astq       Pair these datasets       Mouse-IG-TCR.2/astq         0 pairs       Unpair all         Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an ∨         0 unpaired forward - 0 filtered out       Clear Filters       0 unpaired reverse - 0 filtered out         1/astq       =       0 unpaired forward - 0 filtered out       Clear Filters         No datasets were found matching the current filters.       0 unpaired reverse - 0 filtered out       2/astq         1/astq       =       1 pairs       Unpair all         Mouse-IG-TCR.1/astq       Mouse-IG-TCR.        Mouse-IG-TCR.2/astq         Name       Mouse-IG-TCR       Mouse-IG-TCR       Create list	ini	1.fastq -	hato par	2.faste	¥	lected	i •
Opairs       Unpair all         Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an          Ounpaired forward - 0 filtered out       Clear Filters       Ounpaired reverse - 0 filtered out         Itastq       =       Ounpaired forward - 0 filtered out       Clear Filters       Ounpaired reverse - 0 filtered out         No datasets were found matching the current filters.       No datasets were found matching the current filters.       Image: Consel: CR:1/astq → Mouse-IG-TCR:        Mouse-IG-TCR: / Mous	d	Mouse-IG-TCR.1.fastq	Pair these dataset	ts Mous	e-IG-TCR.2.fastq	q	
O pairs       Unpair all         Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an          O unpaired forward - 0 filtered out       Clear Filters       O unpaired reverse - 0 filtered out         I.1astq       O       Auto-pair       O unpaired reverse - 0 filtered out         No datasets were found matching the current filters.       Notable current filters.       O unpair all         Mouse-IG-TCR.1fastq → Mouse-IG-TCR.       Mouse-IG-TCR.2fastq       &         Hide original elements?       Remove file extensions?          Name       Mouse-IG-TCR       Create list	b						
O pairs       Unpair all         Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an          O unpaired forward - 0 filtered out       Clear Filters       O unpaired reverse - 0 filtered out         I.1astq       O unpaired reverse - 0 filtered out       Clear Filters       O unpaired reverse - 0 filtered out         No datasets were found matching the current filters.       Image: Clear CR1.1astq → Mouse-IG-TCR.1        Mouse-IG-TCR.1        K         Hide original elements?       Remove file extensions?       Mouse-IG-TCR       Mouse-IG-TCR       K	sł						-
Opairs       Unpair all         Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an >         Ounpaired forward - 0 filtered out       Clear Filters       Ounpaired reverse - 0 filtered out         Instr	E F						
O pairs       Unpair all         Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an          O unpaired forward - 0 filtered out         Lifastq       O unpaired reverse - 0 filtered out         Auto-pair       Lastq         No datasets were found matching the current filters.         Image:	· / ·						
Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an ✓ O unpaired forward - 0 filtered out Clear Filters O unpaired reverse - 0 filtered out Lifastq No datasets were found matching the current filters. No datasets were found matching the current filters. No datasets were found matching the current filters. Hide original elements?  Remove file extensions?  Remove file extensions?  Remove file extensions?  Cancel Cance							
Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an O unpaired forward - 0 filtered out Lifast O unpaired reverse - 0 filtered out Lifast No datasets were found matching the current filters. No datasets were found matching the current filters. Hide original elements? Remove file extensions? Remove file	.IV		0 pairs Unpair a	II			
Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an O unpaired forward – O filtered out Lfastq No datasets were found matching the current filters. No datasets were found matching the current filters. Hide original elements? Remove file extensions?							
Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an O unpaired forward - 0 filtered out I.fastq - 0 filtered out Auto-pair No datasets were found matching the current filters. No datasets were found matching the current filters. Mouse-IG-TCR.1.fastq → Mouse-IG-TCR.2.fastq & Hide original elements? Remove file extensions? Hide original elements? Remove file extensions? Name Mouse-IG-TCR Cancel Cancel							
Collections of paired datasets are ordered lists of dataset pairs (often forward and reverse reads). These collections can be passed to tools an O unpaired forward - 0 filtered out Lifastq I fastq I fastq							
No datasets were found matching the current filters.         1		Collections of paired datasets are ord	lered lists of dataset pairs (often for	rward and reverse re	eads). These collections ca	an be passed to tools an.	~ ~
a 1 pairs Unpair all Mouse-IG-TCR.1.fastq → Mouse-IG-TCR.2.fastq Hide original elements? Remove file extensions? Name Mouse-IG-TCR Cancel Create list	tc	O unpaired forward - 0 filtered	ered lists of dataset pairs (often for out Clear Auto	rward and reverse re Filters p-pair	o unpaired re 2.fastq	werse - 0 filtered out	~~ ~
1 pairs       Unpair all         Mouse-IG-TCR.1.fastq →       Mouse-IG-TCR.2.fastq         K       Hide original elements?         K       Name         Mouse-IG-TCR       Cancel	n tc	O unpaired forward - 0 filtered	lered lists of dataset pairs (often for out Clear Auto e current filters.	rward and reverse re Filters o-pair	ads). These collections ca 0 unpaired re 2.fastq	n be passed to tools an. verse - 0 filtered out	~~~~
1 pairs       Unpair all         Mouse-IG-TCR.1.fastq →       Mouse-IG-TCR.2.fastq         K       Mouse-IG-TCR.2.fastq         K       Hide original elements?         K       Name         Mouse-IG-TCR         K       Cancel	n to	O unpaired forward - 0 filtered 1.fastq	lered lists of dataset pairs (often for out Clear Auto e current filters.	rward and reverse re Filters o-pair	ouds). These collections ca 0 unpaired re 2.fastq	n be passed to tools an.	¥
1 pairs       Unpair all         Mouse-IG-TCR.1.fastq →       Mouse-IG-TCR.2.fastq         K       Hide original elements?         K       Name         Mouse-IG-TCR         K         Cancel	ta	O unpaired forward - 0 filtered 1.fastq	lered lists of dataset pairs (often for out Clear Auto e current filters.	rward and reverse re Filters o-pair	ounpaired re	werse - 0 filtered out	···· ¥
1 pairs       Unpair all         Mouse-IG-TCR.1.fastq →       Mouse-IG-TCR.2.fastq         Mouse-IG-TCR.1.fastq →       Mouse-IG-TCR.2.fastq         Hide original elements?          Remove file extensions?          Name       Mouse-IG-TCR         Cancel       Create list	ta L T	O unpaired forward - 0 filtered 1.fastq	lered lists of dataset pairs (often for out Clear Auto	ward and reverse re Filters 5-pair	ounpaired re	werse - 0 filtered out	~ ~
b 1 pairs Unpair all Mouse-IG-TCR.1.fastq → Mouse-IG-TCR.2.fastq Hide original elements? Name Mouse-IG-TCR Cancel Create list	ta ta L T	O unpaired forward - 0 filtered 1.fastq	ered lists of dataset pairs (often for out Clear Auto	ward and reverse re Filters 5-pair	ads). These collections ca 0 unpaired re 2.fastq	werse - 0 filtered out	~ ~
SI Mouse-IG-TCR.1.fastq → Mouse-IG-TCR. ← Mouse-IG-TCR.2.fastq A Discrete discret	ta ta L T ani	O unpaired forward - 0 filtered 1.fastq	ered lists of dataset pairs (often for out Clear Auto	ward and reverse re Filters 5-pair	ads). These collections ca 0 unpaired re 2.fastq	In be passed to tools an.	~ ~
E Hide original elements? ✓ Remove file extensions? ✓ Name Mouse-IG-TCR Cancel Create list	i tc a ior L T ani nd	O unpaired forward - 0 filtered 1.fastq	lered lists of dataset pairs (often for out Clear Auto e current filters.	Ward and reverse re Filters 5-pair Unpair all	o unpaired re 0 unpaired re 2.fastq	In be passed to tools an.	~~~
Hide original elements? Hide original elements? Remove file extensions? Remove file extensions? Remove file extensions? Create list	i tc ani ani ub	Ounpaired forward - 0 filtered 1.fastq	ered lists of dataset pairs (often for out Clear Auto e current filters. Mouse-IG-TCR.1.fastq → Mouse-	Unpair all -IG-TCR. ← Mous-	o unpaired re 2.fastq 9-IG-TCR.2.fastq	In be passed to tools an.	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
P       Hide original elements? ✓       Remove file extensions? ✓         Name       Mouse-IG-TCR       Create list         P       Cancel       Create list	a tc ior L T ani ub ash	O unpaired forward - 0 filtered 1.fastq	ered lists of dataset pairs (often for out Clear Auto e current filters. Mouse-IG-TCR:1.fastq → Mouse-	Filters o-pair Unpair all -IG-TCR. 🗲 Mous	eads). These collections ca 0 unpaired re 2.fastq 9-IG-TCR.2.fastq	n be passed to tools an.	
Hide original elements?  Remove file extensions? Hide original elements? Remove file extensions? Create list	a tc ior L T nd ub c F	O unpaired forward - 0 filtered           1.fastq         •	lered lists of dataset pairs (often for out Clear Auto e current filters. ■ pairs Mouse-IG-TCR.1.fastq → Mouse-	Unpair all	eads). These collections ca 0 unpaired re 2.fastq 9-IG-TCR.2.fastq	n be passed to tools an.	
Hide original elements?  Remove file extensions? Hide original elements? Remove file extensions? Create list Create list	ta ior L 1 ani ub ash c F FA Qi	O unpaired forward - 0 filtered 1.fastq	lered lists of dataset pairs (often for out Clear Auto e current filters. Mouse-IG-TCR:1.fastq → Mouse-	Filters o-pair Unpair all -IG-TCR. ← Mous	o unpaired re 2.fastq 9-IG-TCR.2.fastq	n be passed to tools an.	¥
Hide original elements? Remove file extensions? Name Mouse-IG-TCR Create list	a to a into a into into into into into into into into	O unpaired forward - 0 filtered 1.fastq	lered lists of dataset pairs (often for out Clear Auto e current filters. Mouse-IG-TCR:1.fastq → Mouse-	Filters o-pair Unpair all -IG-TCR.	o unpaired re 2.fastq 9-IG-TCR.2.fastq	werse - 0 filtered out         -	
Name Mouse-IG-TCR Create list	ta ta ta ta ta ta ta ta ta ta ta ta ta t	O unpaired forward - 0 filtered 1.fastq	lered lists of dataset pairs (often for out Clear Auto e current filters. Mouse-IG-TCR:1.fastq → Mouse-	Filters o-pair Unpair all -IG-TCR.   ← Mous	o unpaired re 2.fastq a-IG-TCR.2.fastq	In be passed to tools an.	] &
F Cancel Create list	ta ior L 1 ani asi FA Qi AN	O unpaired forward - 0 filtered 1.fastq No datasets were found matching the	lered lists of dataset pairs (often for out Clear Auto e current filters. Mouse-IG-TCR:1.fastq → Mouse-	Filters ⊳-pair -IG-TCR.   ← Mouse H	eads). These collections ca 0 unpaired re 2.fastq a-IG-TCR.2.fastq de original elements? ✓	werse - 0 filtered out         -          -         -         -         -         -         -         -         -         -         -         -         -         -         - <td> ~</td>	~
r Create list	ta ior L T Jb ash FA Q AN	O unpaired forward - 0 filtered 1.fastq No datasets were found matching the	lered lists of dataset pairs (often for out Clear Auto e current filters. I pairs Mouse-IG-TCR:1.fastq → Mouse- Mouse-IG-TCR:1.fastq → Mouse-	rward and reverse re Filters ⊳-pair -IG-TCR.   ← Mouse H Mouse-IG-TCR	eads). These collections ca 0 unpaired re 2.fastq e-IG-TCR.2.fastq de original elements?	n be passed to tools an.	s; ✓
ar a state and a	ta ior L T ani ub asl C F Qu Qu C F	Ounpaired forward - 0 filtered 1.fastq No datasets were found matching the	lered lists of dataset pairs (often for out Clear Auto e current filters. Mouse-IG-TCR.1.fastq → Mouse- Mouse-IG-TCR.1.fastq → Mouse-	vard and reverse re Filters ⊳-pair -IG-TCR. ← Mous H Mouse-IG-TCR	e-IG-TCR.2.fastq	werse - 0 filtered out         -          -         -         -         -         -         -         -         -         -         -         -         -         -         - <td> ✓  ✓</td>	✓ ✓
	ta ior L T ani ub asl C F Q Q AN CF CF CF CF	O unpaired forward - 0 filtered          1.fastq       •         No datasets were found matching the	lered lists of dataset pairs (often for out Clear Auto e current filters. Mouse-IG-TCR:1.fastq → Mouse- Mouse-IG-TCR:1.fastq → Mouse-	Vunpair all -IG-TCR.   ← Mouse House-IG-TCR	eads). These collections ca 0 unpaired re 2.fastq e-IG-TCR.2.fastq de original elements?	werse - 0 filtered out         -          -         -         -         -         -         -         -         -         -         -         -         -         -         - <td> ×</td>	×



1.4 Import input files required to run pRESTO. The input files have been uploaded onto the <u>example history</u> as file number 4-7. The files can be copied to a new history or downloaded to local computer. To copy the files to a user history, click <u>view all histories</u> and click <u>Switch to</u> to make the destination history current, then select the files to be copied and drag them to the current history.

			🔁 Galaxy	Anal	yze Data Workflow Visualiz	e 🔹 Shared Data 🕇
			search histories	8	search all datasets	<b>8 8</b>
			Current History	•		Switch to 👻
			User new history		NEBNext Immune Sec example data_Mouse	quencing
11-1			(empty)	۲	7 shown	
History	<del>5</del> +щФ		search datasets	00	1.46 GB	Image: Second
search datasets		View all histories		••	search datasets	00
NEBNext Immune Sequene example data_Mouse	cing		Drag datasets here to copy them to the c	current history	All None	
7 shown			1 This history is empty		For	r all selected
1.46 GB	Image: Second				✓ 7: Immune_mouse_Re	ef.fasta
7: Immune_mouse_Ref.fasta	● # ×				6: IS_Mouse_C-Regio	n.txt
6: IS_Mouse_C-Region.txt	⊛ # ×				5: IS_Mouse_R2_Prin	ners.txt
5: IS_Mouse_R2_Primers.txt	● # ×				4: IS_Mouse_R1_Prim	ners.txt
4: IS_Mouse_R1_Primers.txt	● # ×				<b>3: Mouse-IG-TCR</b> a list of pairs with 1 iter	n
3: Mouse-IG-TCR a list of pairs with 1 item	×				2: Mouse-IG-TCR.2.fa	istq
2: Mouse-IG-TCR.2.fastq	● # ×				1: Mouse-IG-TCR.1.fa	stq
1: Mouse-IG-TCR.1.fastq	⊛ # ×					

- 2. Run the pRESTO workflow
  - 2.1. Import the NEB <u>pRESTO workflow</u> to your Galaxy account.

2.2. Activate the history containing your data and input files. Click <u>Workflow</u> on the top menu and chose the <u>Run Workflow</u> button next to the pRESTO workflow.

Search Workflows						+ Create	1 Import	History	<b>€+</b> □ 1
						1 create	in port	search datasets	00
lame imported: pRESTO NEBNext Immune S (it Workflow v3.2.0 IEADME: Example workflow for processing	¢ equencing	Tags NEBNext ×	÷	Updated	Sharing 🌲	Bookmarl	ked 🌲 Run Workflow	NEBNext Immune Seque example history_Mouse pRESTO	ncing _run
nmune Sequencing data with pRESTO. CH 3.1.0: Copied from pRESTO Abseq Workfl Collections). Auto re-layout for clarity. v3:	HANGES: ow v3 1.1: Try to							1.46 GB	
x workflow issue where it stops after pRE ilterSeq without errors in UI. Change 2 pR	STO ESTO							9: Immune_mouse_Ref.fasta	• • •
terSeq tools right after seqtk to: generat g = yes. v3.1.2: Try to fix issue with all pR	e detailed ESTO							8: IS_Mouse_C-Region.txt	• /
rseLog tools failing. Add missing values tion to all pRESTO ParseLog tools using	for -f							7: IS_Mouse_R2_Primers.txt	• /
estov5.3_AbSeqV3_html.sh as a templat y to fix workflow issue where it stops afte	er it after							6: IS_Mouse_R1_Primers.txt	• /
askPrimers, BuildConsensus, AssembleP imer sequences tools to: generate detail	airs, mask ed log =							5: Mouse-IG-TCR a list of pairs with 1 item	
ter it after pRESTO FilterSeq without erro	irs in UI.							2: Mouse-IG-TCR.2.fastq	۲
lange Assemblerans, mask low quality b	ases tools								

#### 2.3. Run the pRESTO workflow

• Fill out Workflow Parameters and History Options

<u>Num Pairs</u>: number of paired end reads to include in the analysis. Consistent subsampling ensures that each library has the same power to detect transcripts but selecting a very large number will cause all available reads to be used if that is preferred. Example number: 500000.

Send results to a new history: select Yes or No

1: Input dataset collection: select the dataset collection in the history

2: R1 Primer FASTA: select input file IS\_Mouse\_R1\_Primers in the history

<u>3: R2 Primer FASTA</u>: select input file IS\_Mouse\_R2\_Primers in the history

4: C-Region FASTA: select input file IS\_Mouse\_C-Region in the history

5: Immune Ref FASTA: select input file Immune\_mouse\_Ref in the history

• Click the <u>Run Workflow</u> button.

Workflow: imported: pRESTO NEBNext Immune Sequencing Kit Workflow v3.2.0	History 😂 🕂 🖽 🌣	
► Run Workflow	search datasets	00
Input Datasets	NEBNext Immune Sequencing example history_Mouse_run	e pRESTO
Paired Fastq Dataset Collection	1.46 GB	v 📎 🗩
R1 Primer FASTA		
Image: Contract of the second seco	9: Immune_mouse_Ref. sta	.fa 🕑 🥒 🗙
Read 1 Primer Fasta	8: IS_Mouse_C-Region. t	tx • / ×
C C 5: IS_Mouse_R2_Primers.txt	7: IS_Mouse_R2_Prime txt	rs. @ 🖋 🗙
Read 2 Primer Fasta	6: IS_Mouse_R1_Prime txt	rs. @ / X
C-Region FASTA	5: Mouse-IG-TCR	×
C C 6: IS_Mouse_C-Region.txt	a list of pairs with 1 item	n Oly
C-Region Easta	2. 110030 10 101.2.103	4 0 9 H
Immune Ref FASTA	1: Mouse-IG-TCR.1.tasto	1 @ # X
C C 7: Immune. mouse. Ref.fasta		
Fasta containing known immune sequnces (used to assemble read pairs that do not overlap)		
Numkeaus		
500000		
decimal 0-0.9999 = fraction of total reads intever 1-N = number of reads		
Expand to full workflow form.		

3. QC report and output.

The sequencing reads QC metrics are visualized in the <u>pRESTOr AbSeq3 Report</u>. For further reads alignment, the output files included in <u>Unique sequences</u> and <u>Unique</u> <u>sequences (>=2 reads</u>) can be used as input for V(D)J alignment tools, for example IgBlast. An example history after running pRESTO is published <u>here</u>.

History	S + 🗆 🕈
search datasets	88
NEBNext Immune Sequencing exa history_Mouse_run pRESTO	mple
14 shown, 83 hidden	
7.08 GB	
96: Report on Unique sequences (>=2 r a list with 1 item	eads) X
93: Unique sequences (>=2 reads) a list with Titem	×
90: Report on Unique Sequences a list with 1 item	×
88: Unique sequences a list with 1 item	×
86: Report on final sequences a list with 1 item	×
84: Final sequences a list with 1 item	×
82: pRESTOr AbSeq3 Report on collecti ollection 54, and others a list with 1 item	on 70, c 🗙