Trimming bases and filtering reads

Based on pre-alignment QA/QC, we need to trim low quality bases from the 3' end of reads.

- Click the **Unaligned reads** data node
- Click **Pre-alignment tools** in the task menu
- Click **Trim bases** (Figure 1)

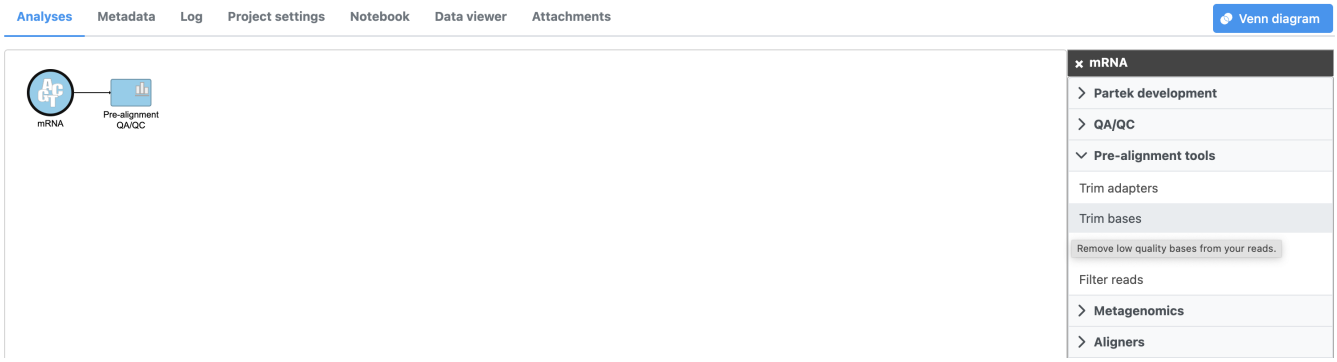


Figure 5. Invoking the Trim bases task

By default, *Trim bases* removes bases starting at the 3' end and continuing until it finds a base pair call with a Phred score of equal to or greater than 35 (Figure 2).

- Click **Finish** to run *Trim bases* with default settings

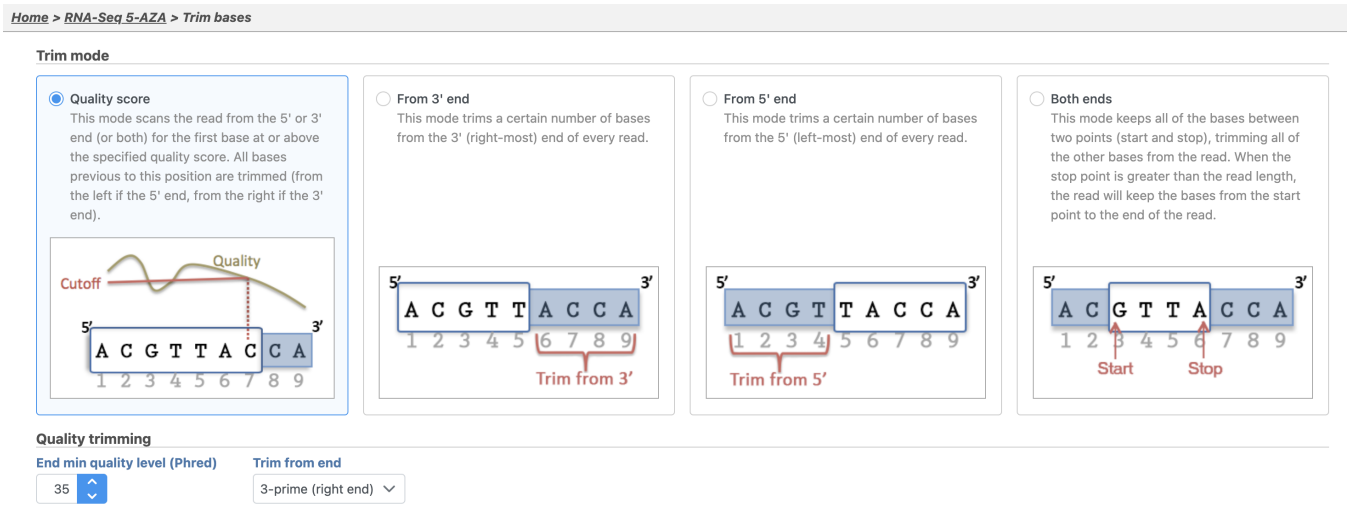


Figure 6. Configuring Trim bases

The *Trim bases* task will generate a new data node, *Trimmed reads* (Figure 3). We can view the task report for *Trim bases* by double-clicking either the *Trim bases* task node or the *Trimmed reads* data node or choosing *Task report* from the task menu.

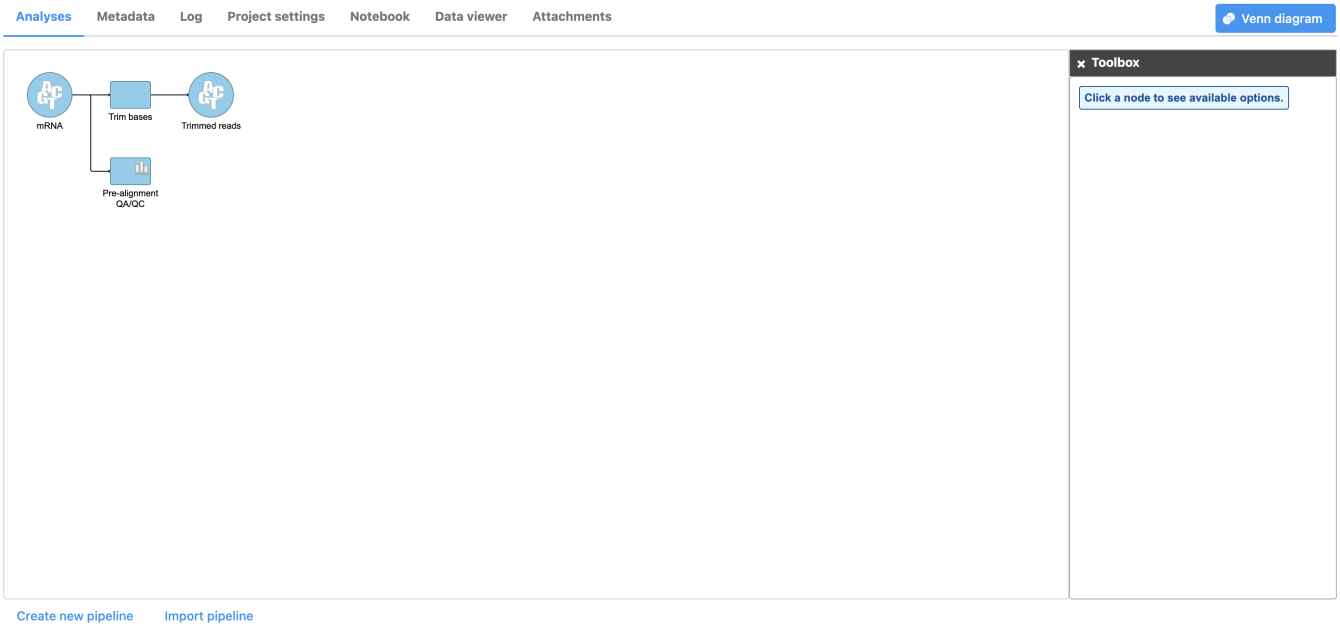


Figure 7. A Task and a Data node are created from the Trim bases task. Task and Data nodes that have been queued or are in progress are shown in a lighter color than completed tasks.

- Double-click the **Trimmed reads** data node to open the task report

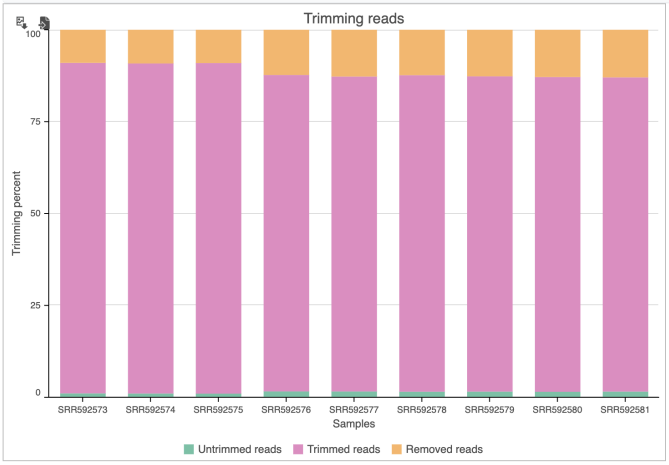
The report shows the percentage of trimmed reads and reads removed in a spreadsheet and a two graphs (Figure 4).

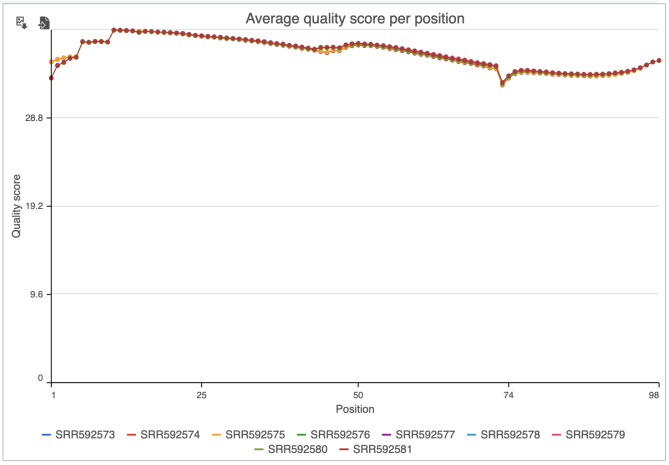
Optional columns

Sample name	Total reads	Reads trimmed	% Reads trimmed	Reads removed	% Reads removed	Average bases trimmed	Pre-trim quality	Post-trim quality
SRR592573	116,350	104,628	89.93%	10,482	9.01%	27.19	32.19	35.77
SRR592574	173,849	156,106	89.79%	15,949	9.17%	27.20	32.09	35.74
SRR592575	242,360	217,965	89.93%	21,979	9.07%	27.26	32.20	35.80
SRR592576	281,368	242,250	86.10%	34,624	12.31%	26.72	31.73	35.84
SRR592577	251,571	215,577	85.69%	32,015	12.73%	26.26	31.67	35.85
SRR592578	293,754	252,971	86.12%	36,314	12.36%	26.76	31.73	35.84
SRR592579	141,924	121,737	85.78%	17,988	12.67%	26.25	31.71	35.86
SRR592580	239,377	205,018	85.65%	30,788	12.86%	27.48	31.49	35.80
SRR592581	206,711	176,709	85.49%	26,792	12.96%	27.05	31.50	35.87

Rows per page 25 (1 of 1)

Download





Task details

Figure 8. Results of the Trim bases task

The results are fairly consistent across samples with ~2% of reads untrimmed, ~86% trimmed, and ~12% removed for each. The average quality score for each sample is increased with higher average quality scores at the 3' ends.

- Click **RNA-Seq 5-AZA** to return to the *Analyses* tab

« Running pre-alignment QA/QC Aligning to a reference genome »

Additional Assistance

If you need additional assistance, please visit our support page to submit a help ticket or find phone numbers for regional support.

Your Rating:  Results:  30 rates