

Criteria	Ratings				
Section 1: <i>E. coli</i> assembly	Excellent	Good	Acceptable	Needs Improvement	Not Done
	<p>Correctly identifies the large section in the middle of the graph as the genome.</p> <p>Identifies the sections around node 4 and 18 as plasmid and phage DNA, respectively.</p> <p>Explains how they identified this (probably with BLAST).</p> <p>Identifies the plasmid as <i>E. coli</i> plasmid F.</p> <p>Describes the most common path for the plasmid as one of: 4, 199, 200, 201, 206, 207, 200, 201, 61, 62, 201, 206, 65, 156, 40, 158, 41, 259, 63, 35, 17, 198, 20, 35, (4) OR 4, 199, 200, 201, 61, 62, 201, 206, 207,</p>	<p>Correctly identifies the large section in the middle of the graph as the genome.</p> <p>Identifies the sections around node 4 and 18 as plasmid and phage DNA, respectively.</p> <p>Explains how they identified this (probably with BLAST).</p> <p>Does NOT identify the plasmid as <i>E. coli</i> plasmid F.</p> <p>Describes the most common path for the plasmid as one of: 4, 199, 200, 201, 206, 207, 200, 201, 61, 62, 201, 206, 65, 156, 40, 158, 41, 259, 63, 35, 17, 198, 20, 35, (4) OR 4, 199, 200, 201, 61, 62, 201, 206, 207,</p>	<p>Correctly identifies the large section in the middle of the graph as the genome.</p> <p>Identifies the sections around node 4 and 18 as plasmid and phage DNA, respectively.</p> <p>Explains how they identified this (probably with BLAST).</p> <p>Does NOT identify the plasmid as <i>E. coli</i> plasmid F.</p> <p>Does not indicate the loops, and describes the most common path for the plasmid as: 4, 199, 200, 201, 206, 65, 156, 40, 158, 41, 259, 63, 35, (4) Identifies that the less common paths are sections with mutations in these</p>	<p>Does not correctly identify the large section in the middle of the graph as the genome.</p> <p>Identifies the sections around node 4 and 18 as plasmid and phage DNA, respectively.</p> <p>Provides no explanation for how they identified this (probably with BLAST).</p> <p>Does NOT identify the plasmid as <i>E. coli</i> plasmid F.</p> <p>Does not indicate the loops, and describes the most common path for the plasmid as: 4, 199, 200, 201, 206, 65, 156, 40, 158, 41, 259, 63, 35, (4) Does not identify that the less common paths are</p>	<p>Does not answer at least 4 of these questions.</p>

	<p>200, 201, 206, 65, 156, 40, 158, 41, 259, 63, 35, 17, 198, 20, 35, (4)</p> <p>Identifies that the less common paths are sections with mutations in these plasmids.</p> <p>Should identify that they can not determine which loop (206, 207, 200, 201 or 201, 61, 62) comes first from this data.</p> <p>They should indicate that that might be able to resolve this with BLAST. Or another suggestion that would work.</p> <p>If previously unsequenced, they would need to get long reads from Pac Bio or ONT or mate pair library to resolve.</p> <p>Indicate that the small, unconnected pieces are likely</p>	<p>200, 201, 206, 65, 156, 40, 158, 41, 259, 63, 35, 17, 198, 20, 35, (4)</p> <p>Identifies that the less common paths are sections with mutations in these plasmids.</p> <p>Should identify that they can not determine which loop (206, 207, 200, 201 or 201, 61, 62) comes first from this data.</p> <p>Does not provide a way to resolve these that would work.</p> <p>If previously unsequenced, they would need to get long reads from Pac Bio or ONT or mate pair library to resolve.</p> <p>Indicates that the small, unconnected pieces are likely contamination or regions of poor sequencing quality.</p>	<p>plasmids.</p> <p>Does not indicate that they can not resolve the loops.</p> <p>Does not provide a way to resolve these that would work in sequenced or unsequenced data.</p> <p>Indicates that the small, unconnected pieces are likely contamination or regions of poor sequencing quality.</p>	<p>sections with mutations in these plasmids.</p> <p>Does not indicate that they can not resolve the loops.</p> <p>Does not provide a way to resolve these that would work in sequenced or unsequenced data.</p> <p>Indicates that the small, unconnected pieces are likely contamination or regions of poor sequencing quality.</p>	
--	---	--	---	--	--

	contamination or regions of poor sequencing quality. Include a fasta file that has the regions indicated, and contains the loops.	Include a fasta file that has the regions indicated, and contains the loops..			
Section 2: <i>S. aureus</i> assembly	Excellent	Good	Acceptable	Needs Improvement	Not Done
	Explains what the N50 and N75 are: the contig in which the assembly reaches 50% of the length of the assembly and 75%, respectively. Explains that these are good for comparing between assemblies of the same genome as the genomes are comparable, and the difference is how large the contigs assembled are. Explains that they are bad comparing between genomes as different genomes are different	Explains what the N50 and N75 are: the contig in which the assembly reaches 50% of the length of the assembly and 75%, respectively. Explains that these are good for comparing between assemblies of the same genome as the genomes are comparable, and the difference is how large the contigs assembled are. Explains that they are bad comparing between genomes as different genomes are	Explains what the N50 and N75 are: the contig in which the assembly reaches 50% of the length of the assembly and 75%, respectively. Explains that these are good for comparing between assemblies of the same genome, but not why. Does not, or incorrectly, explains that they are bad comparing between genomes as different genomes are different lengths and have differences in	Incorrectly explains what the N50 and N75 are: the contig in which the assembly reaches 50% of the length of the assembly and 75%, respectively. Does not provide a correct explanation of why these are good for comparing between assemblies of the same genome. Does not, or incorrectly, explains that they are bad comparing between genomes as different genomes are different lengths and have	At least 3 questions not answered.

	<p>lengths and have different fractions of repetitive DNA, and so will present different challenges and not be comparable.</p> <p>N50=16, N75=31</p> <p>Identified the size of a <i>S. aureus</i> as 2.8 Mbp. So this assembly is about the right size.</p> <p>They should indicate that it is similar in that it has a large central genome with small pieces of non genomic data.</p> <p>There are no complete assemblies of extra-genomic DNA like in the <i>E. coli</i> genome assembly.</p> <p>They should describe that there is one large region of DNA outside the genome (213, 100, 212, 100, 230), but that a BLAST of it</p>	<p>different lengths, but does NOT indicate that the differences in amount of repetitive DNA will vary and make genomes not comparable.</p> <p>N50=16, N75=31</p> <p>Identified the size of a <i>S. aureus</i> as 2.8 Mbp. So this assembly is about the right size.</p> <p>They should indicate that it is similar in that it has a large central genome with small pieces of non genomic data.</p> <p>There are no complete assemblies of extra-genomic DNA like in the <i>E. coli</i> genome assembly.</p> <p>They either do NOT describe that there is one large region of DNA outside the genome (213, 100,</p>	<p>amount of repetitive DNA making genomes not comparable.</p> <p>N50=16, N75=31</p> <p>Identified the size of a <i>S. aureus</i> as 2.8 Mbp. So this assembly is about the right size.</p> <p>Do not indicate similarities or differences.</p> <p>They both do NOT describe that there is one large region of DNA outside the genome (213, 100, 212, 100, 230), and they did not BLAST it and identify that it is found in many <i>S. aureus</i>, and so is probably part of the genome.</p> <p>They trace from 207 to a subset of the nodes:</p> <p>41, 72, &amp; 65</p> <p>OR</p> <p>423, 451, 42, 115, 81, 279, &amp;, 404.</p>	<p>differences in amount of repetitive DNA making genomes not comparable.</p> <p>N50=16, N75=31</p> <p>Does not identify the size of a <i>S. aureus</i> as 2.8 Mbp. So this assembly is about the right size.</p> <p>Do not indicate similarities or differences.</p> <p>They both do NOT describe that there is one large region of DNA outside the genome (213, 100, 212, 100, 230), and they did not BLAST it and identify that it is found in many <i>S. aureus</i>, and so is probably part of the genome.</p> <p>They do not trace from 207 the correct nodes:</p> <p>41, 72, &amp; 65</p> <p>OR</p> <p>423, 451, 42, 115, 81, 279, &amp;, 404.</p>	
--	---	--	---	---	--

	<p>indicates that is it found in many <i>S. aureus</i>, and so is probably part of the genome.  They should trace from 207 to 41, 72, &amp; 65  OR  423, 451, 42, 115, 81, 279, &amp;, 404.</p>	<p>212, 100, 230), or they did not BLAST it and identify that it is found in many <i>S. aureus</i>, and so is probably part of the genome.  They trace from 207 to a subset of the nodes: 41, 72, &amp; 65  OR  423, 451, 42, 115, 81, 279, &amp;, 404.</p>			
Citations	Excellent	Acceptable		Needs Improvement	
	Cites BLAST, QUAST, and Bandage and all other sources	Cites two of the 3 tools, and all other sources.		Does not cite tools.	
Well written and clear	Excellent	Acceptable		Needs Improvement	
	Answers to each section written in paragraph form. Clearly written and easy to understand.	Answers to each section written mainly in paragraph form. Mostly clearly written and easy to understand.		Answers to each section written in bullet form. Difficult to understand.	