

Abstract

Recent studies have shown that SMRT sequencing by Pacific Biosciences is a rapid, effective, and highly accurate platform for generation of complete microbial genome sequences. As early-adopters of the RS II sequencer upgrade, we conducted an extensive and broad comparison to evaluate the new platform and chemistries for simultaneous generation of complete or nearly complete microbial genome sequences and analysis of epigenetic base modifications. Comparing more than 120 bacterial genomes from more than 16 species ranging in genome size from 900 Kbp to 7 Mbp and in GC-content from 30.2% - 64.3%, we generated complete genome sequences at twice the rate for isolates sequenced on RS II compared to isolates sequenced on RS. Overall, when combined with longer insert libraries and rigid size-selection using the Blue Pippin by Sage Science, the RS II upgrade yielded an increase in mean read length and tripling of total per-SMRTcell yield. This significant increase in read length and throughput has enabled more rapid and efficient generation of finished microbial genomes and has rendered this approach the *de facto* standard for small genome sequencing in our center. Further, using comparative Illumina sequencing, we found a median of one putative consensus basecall error per finished genome. Here, we present our experiences with RS II sequencing, a comparison of SMRT sequencing based generation of complete genomes of diverse microbial species using RS and RS II, and a comparison of available genome assemblers for these data.

Discussion

The improvement in both sequencing and assembly results using RS II is significant. Comparing PacBio bacterial genome sequencing data metrics from our last six months (11/2012 – 4/2013) using RS with our first nine months using RS II, we achieved more than a doubling in passed filter reads per SMRT Cell, a nearly 200% increase in total base pair yield, and a 50% increase in subread lengths. Increased polymerase read length and longer insert libraries enabled by BluePippin size selection both contributed to the observed increased subread lengths.

Genome assemblies using RS II data demonstrated significant improvements as well. We compared bacterial genome assembly metrics from the same time periods as above. Each genome was assembled both with CA7.0 and HGAP and the best assembly was selected. Of the 56 isolates sequenced on RS, 15 genomes (27%) assembled into complete genomes. The rate of genomes assembling into complete sequences on RS II increased to more than 60% (39 of 64 isolates). The RS II data also yielded increased contig N50s with an average N50 equal to 98% of genome size compared to an average N50 equal to 65% of genome size for RS.

In addition to comparing sequence and assembly metrics between RS and RS II data, we evaluated three genome assemblers (CA7.0, HGAP, and HGAP2) using a subset of 14 bacterial genomes with a range of sizes and GC%. Based upon contig count and N50, each assembler produced the best assembly for some of the isolates. Though these metrics can be limited in their utility, they provide a reasonable assessment of assembler performance on aggregate. Overall, CA7.0 produced the largest number of complete genome assemblies, while HGAP2 generated the lowest mean contig count and longest mean contig N50.

In our evaluation of assembly consensus quality, we found an average of 3 passed-filter (PF) SNPs and a median of 1 PF SNP per genome for both RS and RS II sequenced genomes. When taking genome size into account, we find just over 1 PF SNP per million bases of genome sequence. Validations of these discrepancies are underway to determine which are PacBio consensus errors and which are due to Illumina sequencing or alignment errors. However, these initial data indicate that bacterial genomes assembled using PacBio data alone generate highly accurate consensus sequence.

Ongoing studies are extending these comparisons to larger genomes and metagenomes, new sequencing chemistries and run lengths, and new assembly methods.

References

DePristo MA, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43(5):491-498. Hasan NA, et al. (2012) Genomic diversity of 2010 Haitian cholera outbreak strains. Proceedings of the National Academy of Sciences. Koren S, et al. (2012) Hybrid error correction and de novo assembly of single-molecule sequencing reads. Nat Biotechnol 30(7):693-700.

Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754-1760. Magoc T, et al. (2013) GAGE-B: An Evaluation of Genome Assemblers for Bacterial Organisms.

Bioinformatics. Miller JR, et al. (2008) Aggressive assembly of pyrosequencing reads with mates. Bioinformatics 24(24):2818-2824.

Shin SC, et al. (2013) Advantages of Single-Molecule Real-Time Sequencing in High-GC Content Genomes. PLoS ONE 8(7):e68824.

Tettelin H, et al. (2012) Genomic Insights into the Emerging Human Pathogen Mycobacterium massiliense. Journal of Bacteriology 194(19):5450.

Acknowledgements

This project has been funded in whole or part with federal funds from the National Institute Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under contract number HHSN272200900009C.



Improved Yield and Diverse Finished Bacterial Genomes using Pacific **Biosciences RS II SMRT Sequencing**

Lisa D. Sadzewicz, Naomi Sengamalay, Xinyue Liu, Sushma Nagaraj, Qi Su, Ivette Santana-Cruz, Alvaro Godinez, Luke J. Tallon Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD



Consensus Quality Evaluation

As one measure of genome consensus sequence quality, we used Illumina MiSeq 250bp PE data to align to complete genomes sequenced using PacBio data alone and assembled using one of three genome assemblers. We selected 6 genomes sequenced using RS and 6 using RS II. An average of 50x Illumina coverage was aligned to the contig consensus sequence using BWA and variants were called using GATK.

Species	GenomeSize	%GC	Platform	Assembler	Raw SNPs	PF SNPs
Ehrlichia chaffeensis	1,200,000	30%	RS	CA7.0	2	2
Ehrlichia chaffeensis	1,200,000	30%	RS	CA7.0	16	10
Bordatella holmesii	3,705,000	62%	RS	HGAP	23	0
Bordatella holmesii	3,705,000	62%	RS	HGAP	10	0
Bordatella holmesii	3,705,000	62%	RS	HGAP	5	0
Mycobacterium abscessus	4,800,000	64%	RS	CA7.0	8	5
Mean Per Genome					10.7	2.8
Median Per Genome					9.0	1.0
Mean Per Mbp Genome					3.5	0.9
Neoehrlichia lotoris	1,300,000	28%	RS II	HGAP2	8	8
Francisella tularensis	1,900,000	32%	RS II	CA7.0	0	0
Francisella tularensis	1,900,000	32%	RS II	HGAP	0	0
Staphylococcus aureus	2,850,000	33%	RS II	HGAP2	3	1
Staphylococcus aureus	2,850,000	33%	RS II	HGAP2	12	1
Mycobacterium abscessus	4,800,000	64%	RS II	CA7.0	9	9
Mean Per Genome					5.3	3.2
Median Per Genome					5.5	1.0
Mean Per Mbp Genome					2.1	1.2

Genome Assembly & Finished Bacterial Genomes

					Median	Mean			Complete
Species	Genome Size	GC%	Genomes	Platform	Contigs	Contigs	Median N50	Mean N50	Genomes
Neorickettsia helminthoeca	900,000	42%	1	RS	1	1	894,365	894,365	1
Ehrlichia chaffeensis	1,200,000	30%	8	RS	2	2	1,183,219	1,179,717	6
Staphylococcus aureus	2,920,000	33%	1	RS	6	6	2,677,329	2,677,329	0
Bordetella holmesii	3,705,000	62%	7	RS	2	3.3	3,694,298	3,092,498	3
Acinetobacter baumannii	4,300,000	39%	6	RS	23	30	356,047	327,813	0
Mycobacteria abscessus	4,800,000	64%	12	RS	6	14	2,094,780	2,167,167	2
Vibrio parahaemolyticus	5,165,770	45%	10	RS	13	16	1,206,254	1,704,545	1
Mycobacteria (mixed)	5,350,000	64%	11	RS	3	8	2,768,387	2,800,673	2
RS SUMMARY			56		7.0	10.0	1,859,335	1,855,513	15
Ricketsiales sp.	1,500,000	35%	16	RS II	3.5	8	1,384,902	1,161,052	7
Streptococcus mitis	1,880,000	40%	1	RS II	1	1	1,881,073	1,881,073	1
Francisella tularensis	1,900,000	32%	2	RS II	1	1	1,874,734	1,874,734	2
Propionibacterium acnes	2,500,000	60%	1	RS II	1	1	2,513,768	2,513,768	1
Haemophilus parasuis	2,500,000	40%	1	RS II	1	1	2,478,759	2,478,759	1
Staphylococcus aureus	2,850,000	33%	36	RS II	2	5.7	2,727,458	2,341,993	23
Salegentibacter mishustinae	3,800,000	37%	1	RS II	1	1	3,791,564	3,791,564	1
Clostridium sp.	4,360,000	31%	2	RS II	1	1	4,373,190	4,373,190	2
Mycobacteria abscessus	4,800,000	64%	3	RS II	2	5	4,818,226	4,381,566	1
Parabacteriodes sp.	7,000,000	43%	1	RS II	5	5	6,847,904	6,847,904	0
RS II SUMMARY			64		1.9	3.0	3,269,158	3,164,560	39



Genome Assembler Comparison

Using 14 genomes with a range of sizes and GC%, and sequenced using both RS and RS II, we evaluated three genome assemblers (CA7.0, HGAP, and HGAP2) for their ability to generate high-quality, complete genomes.



Mean Median Mean Net Mean Nso Median Nso Median Nso Median Nso Median Nso Median Nso Median Nso Median Nso Complete Genomes CA7.0 18.1 1 11.2 1 2,081,210 1,640,793 1,892,823 1,538,3150 8 GAP 5.1 2 3.9 2 2,554,490 2,601,204 1,923,204 1,797,944 6 HGAP 2.6 1.5 2.1 1 3,116,890 2,607,251 2,491,518 2,181,180 7 species Assembler Platfor Corrected CCK Contig Contig base NS Median 1,177,627 1,1		- -			_									
Total Contigs Large Contigs Large Contigs Mean N50 Mean N		Mean	Media	n N	Mean	Media	an		Modia	n		Mad	ian (Complete
Contigs Local Local <thlocal< th=""> Local <thlocal< th=""></thlocal<></thlocal<>		Total	Total	L	.arge	Large	e N	Aean N50		''' M	ean N9	5 No		
CA7.0 18.1 1 11.2 10.10		Contigs	Contia	s Co	ontigs	Conti	gs		N50			N9.	5 (enomes
HGAP Lot Lot <thlot< th=""> <thlot< th=""></thlot<></thlot<>	CA7.0	18.1	1		11 2	1		2 081 210	1 640 7	793 1	892 82	3 1 5 3 8	150	8
HGAP 3.1 2 5.3 2 2,734,490 2,011,204 1,923,204 1,757,9744 0 HGAP 2.6 1.5 2.1 1 3,116,890 2,001,204 1,923,204 1,757,9744 0 species Assembler Platform Corrected Coverage Cortige Contige Large Contige Large Contig Large Contig Bases Mac Contig Bases r/debin chaffeerasis Liberty HGAP RS 30.7 2.12 1,177,827 1,177,827 1,178,167 1,178,167 1,188,167 1,194,281 1,194,281 1,194,281 1,194,281 1,194,281 1,194,281 1,194,281 1,194,281 1,194,281 1,194,281 1,194,281 1,194,281 1,194,2		то.т с 1	- -		20	- 2		2,001,210	2,040,7	001	072,02	4 1 707	044	6
HGAP2 2.6 1.5 2.1 1 3,116,890 2,607,251 2,491,518 2,181,180 7 Species Assembler Platform Corrected Coverage GC% Contigs Contigs N50 N55 Large Contig Enclose chaptersis Liberry Large Contig Bases Min Contig Bases Max Contig Bases Erichic chaptersis Liberry C67,0 R5 20.2 1 1,127,627 1 1,127,627 1 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,167 1,138,163 </td <td>ПСАР</td> <td>5.1</td> <td>Ζ</td> <td></td> <td>5.5</td> <td>Z</td> <td>•</td> <td>2,554,490</td> <td>2,011,2</td> <td>204 1,</td> <td>925,204</td> <td>+ 1,/9/</td> <td>,944</td> <td>0</td>	ПСАР	5.1	Ζ		5.5	Z	•	2,554,490	2,011,2	204 1,	925,204	+ 1,/9/	,944	0
Species Assemble Platform Correctage GC/s Contig Contig NSD Large Large Mins Mins <th< th=""><th>HGAP2</th><th>2.6</th><th>1.5</th><th></th><th>2.1</th><th>1</th><th></th><th>3,116,890</th><th colspan="2">2,607,251 2</th><th>491,51</th><th colspan="2">2,181,180</th><th>7</th></th<>	HGAP2	2.6	1.5		2.1	1		3,116,890	2,607,251 2		491,51	2,181,180		7
Species Assembler Peatform Corrected Correct GCA10 RS Contige Large Contige Large Large Contige Large Large Contige Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large Large<														
Species Assembler CAT/0 Pittorm Coverage Cortig Coverage Contig Coverage Coverage Coverage <thcoverage< th=""> <thcoverage< th=""> Cov</thcoverage<></thcoverage<>					Corrected						Largo	Large Contig	Min Contia	Max Contig
CA2.0 HS So 2 30.26 1 1,177,627 1,181,167 <th< th=""><th>Specie</th><th>25</th><th>Assembler</th><th>Platform</th><th>Coverage</th><th>GC%</th><th>Contigs</th><th>Contig Bases</th><th>N50</th><th>N95</th><th>Contigs*</th><th>Bases</th><th>Bases</th><th>Bases</th></th<>	Specie	25	Assembler	Platform	Coverage	GC%	Contigs	Contig Bases	N50	N95	Contigs*	Bases	Bases	Bases
Erichia chaffeensis Lberty H6AP R5 30.9 30.22 1 1,183,167 1,183,167 1,183,167 1,183,167 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,185,149 1,194,301<			CA7.0	RS	26.2	30.26	1	1,177,627	1,177,627	1,177,627	1	1,177,627	1,177,627	1,177,627
H6AP2 R5 D07 30.28 1 1,135,649 1,135,649 1,135,649 1,135,649 1,135,649 1,135,649 1,135,649 1,135,649 1,135,649 1,135,640 1,135,740 1,134,756 1,134,756 1,134,756 1,134,756 1,134,757 1,246,712 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073 1,381,073	Erlichia chaffeer	nsis Liberty	HGAP	RS	30.9	30.22	1	1,183,167	1,183,167	1,183,167	1	1,183,167	1,183,167	1,183,167
Erichio choffeensis Jax CAV.0 Ris 25.8 302 1 1.194.001 1.194.801 1.184.801 1.184.801 1.184.801 1.184.801 1.184.801 1.184.801 1.881.401 1.881.401 1.881.401 1.881.401 1.881.401 1.881.401 1.881.401 1.881.401 1.881.401<			HGAP2	RS	20.7	30.28	1	1,185,549	1,185,549	1,185,549	1	1,185,549	1,185,549	1,185,549
Indum Onlygeniss and Hear 2 Hear 2 RS 3 3.5.7 30.50 1 1, 194,353 1, 194,353 1 1, 194,353 1,341,353 1,341,353 1,	Erlichia chaffa	oncia lov		RS	25.8	30.29	1	1,194,801	1,194,801	1,194,801	1	1,194,801	1,194,801	1,194,801
CA7.0 R51 36.2 1 <th1< td=""><td>Enicina chajje</td><td>ensis Jax</td><td>ПОАР НСАР2</td><td>RS</td><td>20.0</td><td>30.30</td><td>1 1</td><td>1,194,555</td><td>1,194,555</td><td>1,194,555</td><td>1 1</td><td>1,194,555</td><td>1,194,555</td><td>1,194,555</td></th1<>	Enicina chajje	ensis Jax	ПОАР НСАР2	RS	20.0	30.30	1 1	1,194,555	1,194,555	1,194,555	1 1	1,194,555	1,194,555	1,194,555
Streptococcus mitis HGAP RSI 25.1 40.45 1 1.881,073			CA7 0	RS II	36.3	40.46	1	1 881 498	1 881 498	1 881 498	1	1 881 498	1 881 498	1 881 498
HGAP2 RS II 21.6 40.46 1 1.883,601	Streptococci	us mitis	HGAP	RS II	25.1	40.45	1	1.881.073	1.881.073	1.881.073	1	1.881.073	1.881.073	1.881.073
CA7.0 R5 II 19.9 39.69 1 2.480.212 2.480.212 1 2.480.212 2.505.633 2.507.260 2.507.260 2.507.260 2.507.850 2.707.878 2.706.878 2.706.878 2.706.878 2.706.878 2.706.878 2.706.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.707.878 2.70	01.0000000		HGAP2	RS II	21.6	40.46	- 1	1.883.601	1.883.601	1.883.601	1	1.883.601	1.883.601	1.883.601
Haemophilus parasuis HGAP R5 II 15.8 39.77 2 2.502,045 2.486,298 2.486,298 2.486,298 2.486,298 2.486,298 2.486,298 2.486,298 2.486,298 2.478,759 <td></td> <td></td> <td>CA7.0</td> <td>RS II</td> <td>19.9</td> <td>39.69</td> <td>1</td> <td>2,480,212</td> <td>2,480,212</td> <td>2,480,212</td> <td>1</td> <td>2,480,212</td> <td>2,480,212</td> <td>2,480,212</td>			CA7.0	RS II	19.9	39.69	1	2,480,212	2,480,212	2,480,212	1	2,480,212	2,480,212	2,480,212
HGAP2 RS II 16.0 39.72 1 2,478,759 2,478,759 2,478,759 1 2,478,759 2,478,759 2,478,759 1 2,507,260 2,507,260 1 2,507,260 2,507,260 2,507,260 1 2,507,260 1,508 2,507,260 2,507,260 1,508,32 2,700,787 2,700,787 1,604,69 1,164,163 2,707,878 2,707,8	Haemophilus	parasuis	HGAP	RS II	15.8	39.77	2	2,502,045	2,486,298	2,486,298	2	2,502,045	15,747	2,486,298
Propionibacterium acnes CA7.0 RS II 22.00 60.01 1 2,507,260 2,507,20	,		HGAP2	RS II	16.0	39.72	1	2,478,759	2,478,759	2,478,759	1	2,478,759	2,478,759	2,478,759
Propionibacterium acces HGAP RS II 30.5 60.04 1 2,13,768 2,11,768 1 2,513,768 2,520,672 2,570,735 2,708,783 2,708,783 2,708,783 2,708,783 2,708,783 2,708,783 2,780,783 2,821,485 2,721,582 1,516 2,812,135 <			CA7.0	RS II	22.0	60.01	1	2,507,260	2,507,260	2,507,260	1	2,507,260	2,507,260	2,507,260
HGAP2 RS II 17.0 60.01 1 2,506,623 2,506,623 1 2,506,623 2,506,623 2,70,678 1,828 2,700,788 2,705,850 2,705,850 2,70,738 1,250,735 1,2584 2,705,850 Staphylococcus aureus 241 HGAP RS II 15.6 32.77 2 2,739,831 2,707,878 2 2,739,831 3,1953 2,707,878 2 2,739,831 3,1953 2,707,878 2 2,739,831 3,1953 2,707,878 2 2,739,831 3,1953 2,707,878 2 2,739,831 3,1953 2,707,878 2 2,739,831 3,1953 2,707,878 2 2,739,831 3,1953 2,707,878 2 2,739,831 3,1953 2,707,878 2 2,739,831 1,766,70 1,164,165 3,282,148 2,822,148 1,282,148 2,822,148 2,822,148 2,822,148 2,822,148 2,822,148 2,822,148 2,822,148 2,822,148 2,822,148 2,823,048 2,823,048 2,823,048 2,823,048 2,823,048	Propionibacter	ium acnes	HGAP	RS II	30.5	60.04	1	2,513,768	2,513,768	2,513,768	1	2,513,768	2,513,768	2,513,768
Staphylococcus aureus 243 HGAP RS II 15.8 32.81 2 2720,678 2,705,850 2 2,720,787 3 2,750,735 2,708,639 3 2,750,735 2,708,639 3 2,750,735 2,708,639 3 2,750,735 2,708,787 2 2,739,831 3,1953 2,708,850 Staphylococcus aureus 211 HGAP RS II 21.0 32.82 1 2,823,743 2,822,436 1 2,820,436 1			HGAP2	RS II	17.0	60.01	1	2,506,623	2,506,623	2,506,623	1	2,506,623	2,506,623	2,506,623
Staphylococcus aureus 243 HGAP RS II 16.0 3.277 3 2,750,735 2,708,639 3 2,750,735 12,584 2,708,639 Staphylococcus aureus 221 RGAP RS II 15.6 32.77 2 2,739,831 2,707,878 2 2,739,831 31,953 2,707,878 Staphylococcus aureus 221 HGAP RS II 20.0 32.81 7 2,830,704 1,046,604 12,521 5 2,815,159 6,147 1,164,165 Staphylococcus aureus 211 HGAP2 RS II 13.0 32.84 1 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,807 2,833,087			CA7.0	RS II	15.8	32.81	2	2,720,678	2,705,850	2,705,850	2	2,720,678	14,828	2,705,850
HGAP2 RS II 15.6 32.77 2 2,73,93.81 2,707,878 2 2,739,831 31,953 2,707,878 Staphylococcus aureus 221 HGAP RS II 21.7 32.81 7 2,832,436 1,282,436 1,282,436 1,282,436 1,282,436 1,282,436 2,822,436 1,282,436 2,822,436 1,282,436 2,822,436 2,822,436 2,822,436 2,822,436 2,822,436 2,822,138 2,822,138 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 1,936,920 1,593 2,820,118 1 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 1,804 2,823,087 2,823,087 1,804 1,428,493 Mycobacterium abscessus 248 HGAP RS	Staphylococcus	aureus 243	HGAP	RS II	16.0	32.77	3	2,750,735	2,708,639	2,708,639	3	2,750,735	12,584	2,708,639
CA7.0 RS II 21.7 32.81 7 2,830,704 1,048,604 129,521 5 2,815,159 6,147 1,164,163 Staphylococcus aureus 221 HGAP RS II 20.0 32.82 1 2,822,436 2,820,118 1 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,118 2,820,108 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087 2,823,087			HGAP2	RS II	15.6	32.77	2	2,739,831	2,707,878	2,707,878	2	2,739,831	31,953	2,707,878
Staphylococcus aureus 221 HGAP RS II 20.0 32.82 1 2,822,436 2,822,0118 2,822,0118 2,820,118 2,823,087 2,830,			CA7.0	RS II	21.7	32.81	7	2,830,704	1,048,604	129,521	5	2,815,159	6,147	1,164,165
HGAP2 RS II 19.5 32.83 3 2,842,228 2,421,582 407,680 3 2,842,228 1,2,966 2,421,582 Staphylococcus aureus 241 HGAP RS II 15.9 32.84 1 2,820,118 2,823,087 <td< td=""><td>Staphylococcus</td><td>aureus 221</td><td>HGAP</td><td>RS II</td><td>20.0</td><td>32.82</td><td>1</td><td>2,822,436</td><td>2,822,436</td><td>2,822,436</td><td>1</td><td>2,822,436</td><td>2,822,436</td><td>2,822,436</td></td<>	Staphylococcus	aureus 221	HGAP	RS II	20.0	32.82	1	2,822,436	2,822,436	2,822,436	1	2,822,436	2,822,436	2,822,436
CA7.0 RS II 15.9 32.84 1 2,820,118 2,820,118 1 2,820,118 16,642 2 2,823,087			HGAP2	RS II	19.5	32.83	3	2,842,228	2,421,582	407,680	3	2,842,228	12,966	2,421,582
Staphylococcus aureus 241 HGAP RS II 13.4 32.88 2 2,840,510 2,823,666 2,823,666 2,823,087 1 2,823,087 2,833,04			CA7.0	RS II	15.9	32.84	1	2,820,118	2,820,118	2,820,118	1	2,820,118	2,820,118	2,820,118
HGAP2 RS II 15.0 32.82 1 2,823,087 2,823,087 1 2,823,087 3,214,483 Mycobacterium abscessus 248 HGAP RS 18.5 64.20 4 4,807,730 1,400,088 1,056,005 3 4,802,889 4,841 2,346,796 Mycobacterium abscessus 244 HGAP RS 12.3 64.17 1 5,030,544 5,030,544 1 5,030,5454 1 5,030,545	Staphylococcus	aureus 241	HGAP	RS II	13.4	32.88	2	2,840,510	2,823,566	2,823,566	2	2,840,510	16,944	2,823,566
Mycobacterium abscessus 2258 HGAP RS 8.6 64.22 6 4,798,287 3,214,483 185,224 14 4,785,270 18,019 1,128,493 Mycobacterium abscessus 2258 HGAP RS 8.6 64.22 6 4,798,287 3,214,483 185,294 5 4,795,950 2,337 3,214,483 Mycobacterium abscessus 2446 HGAP2 RS 15.4 64.20 4 4,807,730 1,400,088 1,056,005 3 4,802,889 4,841 2,346,796 Mycobacterium abscessus 2446 HGAP RS 12.3 64.20 4 4,832,858 3,187,532 1,630,014 2 4,817,546 5,720 3,187,532 Mycobacterium abscessus 21 HGAP RS 19.1 64.17 1 5,030,544			HGAP2	RS II	15.0	32.82	1	2,823,087	2,823,087	2,823,087	1	2,823,087	2,823,087	2,823,087
Mycobacterium abscessus 2258 HGAP RS 8.6 64.22 6 4,795,87 3,214,483 185,294 5 4,795,950 2,337 3,214,483 HGAP2 RS 18.5 64.21 3 4,813,404 3,407,250 1,399,956 2 4,807,206 6,198 3,407,250 Mycobacterium abscessus 2446 HGAP RS 12.3 64.20 4 4,807,730 1,400,088 1,056,005 3 4,802,889 4,841 2,346,796 Mycobacterium abscessus 2446 HGAP RS 12.3 64.20 4 4,832,858 3,187,532 1,630,014 2 4,817,546 5,720 3,187,532 HGAP2 RS 19.1 64.17 1 5,030,544 5,030,544 1 5,030,544 5,030,544 5,024,321 163,172 4,888,065 Mycobacterium abscessus 21 HGAP RS 8.9 64.17 2 5,024,231 5,024,321 1 5,022,131 6,358 5,024,321 6,352,194 5,252,194	N dura ha atani wa ak	2250	CA7.0	RS	12.3	64.21	14	4,786,270	/32,562	116,622	14	4,786,270	18,019	1,128,493
Mycobacterium abscessus 2446 HGAP RS 18.5 64.21 3 4,813,404 5,407,250 1,399,956 2 4,807,006 6,198 3,407,250 Mycobacterium abscessus 2446 HGAP RS 15.4 64.20 4 4,832,858 3,187,532 1,630,014 2 4,807,250 3 4,821,889 4,841,766 5,720 3,187,532 HGAP2 RS 19.1 64.19 6 4,843,766 1,466,456 948,720 3 4,821,413 6,186 2,408,967 Mycobacterium abscessus 21 HGAP RS 2.4.4 64.17 1 5,030,544 5,030,544 1 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,030,544 5,024,237 136,172 4,888,065 2 5,024,237 136,172 4,888,065 2 5,024,237 136,172 4,888,065 2 5,024,237 136,172 4,888,065 2 5,024,321 5,024,321 </td <td>Nycobacterium ab</td> <td>oscessus 2258</td> <td>HGAP</td> <td>KS DC</td> <td>8.6</td> <td>64.22</td> <td>ь С</td> <td>4,798,287</td> <td>3,214,483</td> <td>185,294</td> <td>5</td> <td>4,795,950</td> <td>2,337</td> <td>3,214,483</td>	Nycobacterium ab	oscessus 2258	HGAP	KS DC	8.6	64.22	ь С	4,798,287	3,214,483	185,294	5	4,795,950	2,337	3,214,483
Mycobacterium abscessus 2446 HGAP RS 12.3 64.20 4 4,807,730 1,400,088 1,050,005 3 4,002,889 4,841 2,340,730 Mycobacterium abscessus 2446 HGAP2 RS 12.3 64.19 6 4,843,376 1,466,456 948,720 3 4,802,889 5,720 3,317,532 Mycobacterium abscessus 21 HGAP2 RS 19.1 64.17 1 5,030,544 5,024,231 16,517 13,6172 4,888,065 2,52,194 15,252,194 15,252,194 15,252,194 15,252,194 15,252,194 15,252,19			HGAP2	KS DC	18.5	64.21	3	4,813,404	3,407,250	1,399,950	2	4,807,206	6,198	3,407,250
Mycobacterium abscessus 2440 IIGAP IIGAP <th< td=""><td>Mucobactorium at</td><td>DECREEUE DAAE</td><td></td><td>DC K2</td><td>10.4 10.2</td><td>04.2U</td><td>4 л</td><td>4,0U1,13U 1 Q27 OE0</td><td>1,400,088 2 1 27 500</td><td>1 620 014</td><td>5 7</td><td>4,0UZ,889 1 817 EVC</td><td>4,041 5 700</td><td>2,340,/90 2 107 E22</td></th<>	Mucobactorium at	DECREEUE DAAE		DC K2	10.4 10.2	04.2U	4 л	4,0U1,13U 1 Q27 OE0	1,400,088 2 1 27 500	1 620 014	5 7	4,0UZ,889 1 817 EVC	4,041 5 700	2,340,/90 2 107 E22
Mycobacterium abscessus 21 IGM 2 IGM 2 <thigm 2<="" th=""> IGM 2 <thi< td=""><td>wycobucienum al</td><td>JSCC35US 2440</td><td>HGADO</td><td>BC LD</td><td>12.5 10 1</td><td>61 10</td><td>4 6</td><td>4,002,000 1 Q11 276</td><td>2,107,252 1 //66 //56</td><td>1,030,014 9/8 720</td><td>2</td><td>4,017,040 <u>1</u> 801 112</td><td>5,720 6 196</td><td>2,101,222 7,102 0A7</td></thi<></thigm>	wycobucienum al	JSCC35US 2440	HGADO	BC LD	12.5 10 1	61 10	4 6	4,002,000 1 Q11 276	2,107,252 1 //66 //56	1,030,014 9/8 720	2	4,017,040 <u>1</u> 801 112	5,720 6 196	2,101,222 7,102 0A7
Mycobacterium abscessus 21 HGAP RS 8.9 64.17 2 5,030,344 3,030,344 1 5,030,344 1 5,030,344 1 5,030,344 1 5,030,344 5,024,321 1 5,024,321 1 5,024,321 1 5,024,321 1 5,024,321 1 5,024,321 1 6,04,371 1 5,252,194 5,252,194 5,252,194 5,252,207 5,252,207 5,252,207 5,252,207 5,252,207 5,252,207				RC	24 /	64.13	1	4,044,370 5 020 5 <i>11</i>	5 020 544	5 020 544	ى 1	5 024,145	5 020 544	5 030 511
Mycobacterium abscessus 11 HGAP2 RS 23.2 64.17 2 5,024,257 4,000,005 2 5,024,321 1,01,72 4,000,005 2 6,005,005 2 5,024,321 1,01,72 4,000,005 2 5,024,321 1 5,024,321 6,358 5,024,321 6,435 7,025,194 1,60,474 7,02,511 1,162,445 742,287 8 5,301,331 4,815 1,660,474 Mycobacterium kansasii HGAP RS II 15.9 66.03 1	Mycohacterium	abscessus 21	HGAP	RS	24.4 8 9	64 17	2	5 024 227	4 888 065	4 888 065	2	5 024 227	136 172	4 888 065
Mycobacterium abscessus 1518 CA7.0 RS 19.9 64.03 1 5,252,194 5,252,194 1 5,252,194 1,60,474 Mycobacterium kansasii HGAP RS II 15.9 66.03 197 7,162,511 101,842 22,158 107 6,615,200 1,613 303,398 303,398 303,398 303,131 4,436,441 220,171 6 6,726,86	,coouccentant o		HGAP2	RS	23.2	64.17	2	5.030.679	5.024.321	5.024.321	-	5.024.321	6,358	5.024.321
Mycobacterium abscessus 1518 HGAP RS 14.6 64.03 22 5,394,060 1,162,445 742,287 8 5,301,331 4,815 1,660,474 HGAP2 RS 22.0 64.04 1 5,252,207 5,252,207 1 5,252,207 5,252			CA7.0	RS	19.9	64.03	1	5.252.194	5,252.194	5,252.194	1	5,252.194	5.252.194	5.252.194
HGAP2 RS 22.0 64.04 1 5,252,207 5,252,207 1 5,252,207 1,613 303,398 13,433,207 13,433,207 14,436,441 220,171 6 6,264,699 10,181 1,343,42,407 14,436,441 220,171 6 6,964,1,437 13,425 5,177,124 <th< td=""><td>Mycobacterium ab</td><td>oscessus 1518</td><td>HGAP</td><td>RS</td><td>14.6</td><td>64.03</td><td>22</td><td>5.394.060</td><td>1,162.445</td><td>742.287</td><td>- 8</td><td>5,301.331</td><td>4.815</td><td>1,660.474</td></th<>	Mycobacterium ab	oscessus 1518	HGAP	RS	14.6	64.03	22	5.394.060	1,162.445	742.287	- 8	5,301.331	4.815	1,660.474
Mycobacterium kansasii CA7.0 RS II 15.9 66.03 197 7,162,511 101,842 22,158 107 6,615,200 1,613 303,398 Mycobacterium kansasii HGAP RS II 14.9 66.11 21 6,801,699 519,935 151,103 21 6,801,699 10,181 1,343,207 HGAP2 RS II 19.1 66.14 7 6,733,948 4,436,441 220,171 6 6,726,861 7,087 4,436,441 Parabacteroides sp. HGAP RS II 15.2 43.44 21 6,922,820 803,737 125,116 18 6,895,544 8,221 2,901,215 HGAP RS II 16.1 43.43 5 6,941,437 5,177,124 1,714,814 5 6,941,437 13,425 5,177,124 HGAP2 RS II 15.8 43.44 6 6,937,625 6,847,904 5 6,928,550 9,075 6,847,904	,		HGAP2	RS	22.0	64.04	1	5,252,207	5,252,207	5,252,207	1	5,252,207	5,252,207	5,252,207
Mycobacterium kansasii HGAP RS II 14.9 66.11 21 6,801,699 519,935 151,103 21 6,801,699 10,181 1,343,207 HGAP2 RS II 19.1 66.14 7 6,733,948 4,436,441 220,171 6 6,726,861 7,087 4,436,441 CA7.0 RS II 15.2 43.44 21 6,922,820 803,737 125,116 18 6,895,544 8,221 2,901,215 Parabacteroides sp. HGAP RS II 16.1 43.43 5 6,941,437 5,177,124 1,714,814 5 6,941,437 13,425 5,177,124 HGAP2 RS II 15.8 43.44 6 6,937,625 6,847,904 5 6,928,550 9,075 6,847,904			CA7.0	RS II	15.9	66.03	197	7,162,511	101,842	22,158	107	6,615,200	1,613	303,398
HGAP2RS II19.166.1476,733,9484,436,441220,17166,726,8617,0874,436,441CA7.0RS II15.243.44216,922,820803,737125,116186,895,5448,2212,901,215Parabacteroides sp.HGAPRS II16.143.4356,941,4375,177,1241,714,81456,941,43713,4255,177,124HGAP2RS II15.843.4466,937,6256,847,9046,847,90456,928,5509,0756,847,904	Mycobacteriur	m kansasii	HGAP	RS II	14.9	66.11	21	6,801,699	519,935	151,103	21	6,801,699	10,181	1,343,207
CA7.0 RS II 15.2 43.44 21 6,922,820 803,737 125,116 18 6,895,544 8,221 2,901,215 Parabacteroides sp. HGAP RS II 16.1 43.43 5 6,941,437 5,177,124 1,714,814 5 6,941,437 13,425 5,177,124 HGAP2 RS II 15.8 43.44 6 6,937,625 6,847,904 6,847,904 5 6,928,550 9,075 6,847,904			HGAP2	RS II	19.1	66.14	7	6,733,948	4,436,441	220,171	6	6,726,861	7,087	4,436,441
Parabacteroides sp. HGAP RS II 16.1 43.43 5 6,941,437 5,177,124 1,714,814 5 6,941,437 13,425 5,177,124 HGAP2 RS II 15.8 43.44 6 6,937,625 6,847,904 6,847,904 5 6,928,550 9,075 6,847,904			CA7.0	RS II	15.2	43.44	21	6,922,820	803,737	125,116	18	6,895,544	8,221	2,901,215
HGAP2 RS II 15.8 43.44 6 6,937,625 6,847,904 6,847,904 5 6,928,550 9,075 6,847,904	Parabactero	<i>ides</i> sp.	HGAP	RS II	16.1	43.43	5	6,941,437	5,177,124	1,714,814	5	6,941,437	13,425	5,177,124
			HGAP2	RS II	15.8	43.44	6	6,937,625	6,847,904	6,847,904	5	6,928,550	9,075	6,847,904

				_	:								
Mean		Media	n N	/lean	Media	an		Madia	n		Mad	ian (`omplete
	Total	Total	L	arge	Large	e M	Mean N50		"' M	ean N9	5 No		ompiete
	Contigs	Contig	s Co	ontigs	Conti	gs		N50			N9.	5 6	ienomes
	18 1	1	<u>, </u>	11 2	<u>5</u>		2 081 210	1 6/0 7	793 1	207 27	2 1 5 2 8	150	8
	то.т г 1	- -		2.0	1			1,0+0,7	$\begin{array}{c} 33 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 $	002,02	, 1 707	,130	0
HGAP	5.1	Z		3.9	Z		2,554,490	2,611,2	204 I,	923,204	4 1,/9/	,944	6
HGAP2	2.6	1.5		2.1	1		3,116,890	2,607,251 2		491,51	2,181,180		7
Species		Assembler	Platform	Corrected Coverage	GC%	Contigs	Contig Bases	N50	N95	Large Contigs*	Large Contig Bases	Min Contig Bases	Max Contig Bases
		CA7.0	RS	26.2	30.26	1	1,177,627	1,177,627	1,177,627	1	1,177,627	1,177,627	1,177,627
Erlichia chaffeensis	is Liberty	HGAP	RS	30.9	30.22	1	1,183,167	1,183,167	1,183,167	1	1,183,167	1,183,167	1,183,167
		HGAP2	RS	20.7	30.28	1	1,185,549	1,185,549	1,185,549	1	1,185,549	1,185,549	1,185,549
Frlichia chaffeen	nsis lax	HGAP	RS	25.8 35.7	30.29	1	1,194,601	1,194,601	1,194,601	1	1,194,601 1 194 333	1,194,601	1,194,801 1 194 333
	1515 54	HGAP2	RS	20.0	30.30	1	1.194.796	1.194.796	1.194.796	1	1.194.796	1.194.796	1,194,796
		CA7.0	RS II	36.3	40.46	1	1,881,498	1,881,498	1,881,498	1	1,881,498	1,881,498	1,881,498
Streptococcus i	mitis	HGAP	RS II	25.1	40.45	1	1,881,073	1,881,073	1,881,073	1	1,881,073	1,881,073	1,881,073
		HGAP2	RS II	21.6	40.46	1	1,883,601	1,883,601	1,883,601	1	1,883,601	1,883,601	1,883,601
		CA7.0	RS II	19.9	39.69	1	2,480,212	2,480,212	2,480,212	1	2,480,212	2,480,212	2,480,212
Haemophilus pa	arasuis	HGAP	RS II	15.8	39.77	2	2,502,045	2,486,298	2,486,298	2	2,502,045	15,747	2,486,298
		HGAP2	RS II	16.0	39.72	1	2,478,759	2,478,759	2,478,759	1	2,478,759	2,478,759	2,478,759
		CA7.0	RS II	22.0	60.01	1	2,507,260	2,507,260	2,507,260	1	2,507,260	2,507,260	2,507,260
Propionibacteriun	m acnes	HGAP	RS II	30.5	60.04	1	2,513,768	2,513,768	2,513,768	1	2,513,768	2,513,768	2,513,768
		HGAP2	RS II	17.0	60.01	1	2,506,623	2,506,623	2,506,623	1	2,506,623	2,506,623	2,506,623
		CA7.0	RS II	15.8	32.81	2	2,720,678	2,705,850	2,705,850	2	2,720,678	14,828	2,705,850
Staphylococcus au	ireus 243	HGAP	RS II	16.0	32.77	3	2,750,735	2,708,639	2,708,639	3	2,750,735	12,584	2,708,639
		HGAP2	RS II	15.6	32.77	2	2,739,831	2,707,878	2,707,878	2	2,739,831	31,953	2,707,878
Charle Jacobia		CA7.0	RS II	21.7	32.81	7	2,830,704	1,048,604	129,521	5	2,815,159	6,147	1,164,165
Staphylococcus au	ireus 221	HGAP	RS II	20.0	32.82	1	2,822,436	2,822,436	2,822,436	1	2,822,436	2,822,436	2,822,436
		HGAPZ		19.5	32.83	3	2,842,228	2,421,582	407,680	3	2,842,228	12,966	2,421,582
Stanbulococcus au	1 vrous 241			15.9	32.84 27.00	1	2,820,118	2,820,118	2,820,118	1	2,820,118	2,820,118	2,820,118
Stupilylococcus uu	1eus 241	HGAP2	RS II	15.4	32.00	2	2,840,310	2,823,500	2,823,300	2 1	2,840,310	2 823 087	2,823,300
		CA7.0	RS	12.3	64 21	14	4 786 270	732 562	116 622	14	4 786 270	18 019	1 128 493
Mycobacterium absc	cessus 2258	HGAP	RS	8.6	64.22	6	4,798,287	3.214.483	185.294	5	4,795,950	2.337	3.214.483
		HGAP2	RS	18.5	64.21	3	4.813.404	3.407.250	1.399.956	2	4.807.206	6.198	3.407.250
		CA7.0	RS	15.4	64.20	4	4.807.730	1.400.088	1.056.005	3	4.802.889	4.841	2,346,796
Mycobacterium absc	cessus 2446	HGAP	RS	12.3	64.20	4	4,832,858	3,187,532	1,630,014	2	4,817,546	5,720	3,187,532
		HGAP2	RS	19.1	64.19	6	4,844,376	1,466,456	948,720	3	4,824,143	6,186	2,408,967
		CA7.0	RS	24.4	64.17	1	5,030,544	5,030,544	5,030,544	1	5,030,544	5,030,544	5,030,544
Mycobacterium abs	scessus 21	HGAP	RS	8.9	64.17	2	5,024,237	4,888,065	4,888,065	2	5,024,237	136,172	4,888,065
		HGAP2	RS	23.2	64.17	2	5,030,679	5,024,321	5,024,321	1	5,024,321	6,358	5,024,321
		CA7.0	RS	19.9	64.03	1	5,252,194	5,252,194	5,252,194	1	5,252,194	5,252,194	5,252,194
Mycobacterium absc	cessus 1518	HGAP	RS	14.6	64.03	22	5,394,060	1,162,445	742,287	8	5,301,331	4,815	1,660,474
		HGAP2	RS	22.0	64.04	1	5,252,207	5,252,207	5,252,207	1	5,252,207	5,252,207	5,252,207
		CA7.0	RS II	15.9	66.03	197	7,162,511	101,842	22,158	107	6,615,200	1,613	303,398
Mycobacterium k	kansasii	HGAP	RS II	14.9	66.11	21	6,801,699	519,935	151,103	21	6,801,699	10,181	1,343,207
		HGAP2	RS II	19.1	66.14	7	6,733,948	4,436,441	220,171	6	6,726,861	7,087	4,436,441
		CA7.0	RS II	15.2	43.44	21	6,922,820	803,737	125,116	18	6,895,544	8,221	2,901,215
Parabacteroide	es sp.	HGAP	KS II	16.1	43.43	5	6,941,437	5,1//,124	1,/14,814	5	6,941,437	13,425	5,1//,124
		нбар5	KS II	15.8	43.44	6	6,937,625	७,847,904	b,847,904	5	o,928,550	9,075	b,847,904



Neorickettsia helminthoeca Ehrlichia chaffeensis Staphylococcus aureus Bordetella holmesii Acinetobacter baumannii Mycobacteria abscessus Vibrio parahaemolyticus Mycobacteria (mixed) **Ricketsiales** sp Streptococcus mitis Francisella tularensis Propionibacterium acnes Haemophilus parasuis Staphylococcus aureus Salegentibacter mishustinae Clostridium sp. Mycobacteria abscessus Parabacteriodes sp

