

# Assembly Report

## Experiment Details

ExperimentID:	Experiment 85
Sequencing Coverage:	21
Number of Partitions:	5
Allowed Mismatch Count in (Base):	0
Length of Sliding Window:	41

## Reads Dataset Details

Total Reads Count:	2205
Total Dataset Size in (Base):	440865
Valid Read Count:	2205
Rejected Read Count:	0
Maximum Read Length:	201
Minimum Read Length:	162

## Reference Genome Details

Length of Reference Genome (Base):	20160
Number of K-mers:	19960
Length of K-mer	326

## Assembly Details

Number of Contigs:	43
Contig N50:	1076
Contig N90:	926
Number of Scaffold	1
Scaffold N50:	20160
Scaffold N90:	20160
Mis-assembly Count:	0
TotalAssembly Size:	20160

## Repeat Details

Total Reads Count(Non-Repeat) :	421
Total Reads Count(Repeat):	360
Retetitive Read Count based on (Partitions Identifier):	358
Retetitive Read Count (Entire Read Frequency Identifier):	5
Retetitive Read Count (Contained Read Frequency Identifier):	3
Total Repetitive Sequences Count	5727
Total Unique Repetitive Sequences Count	7

Total Repeat Size (Base)			8400			
Starting	Ending	Repetitive Sequences	Repeat Count	Length	Positions	
1	20160	CGGTCGAAAA	8	350		
		ACTGCTGGCA				
		GTGGGGCATT				
		ACCTCGAATCT				
		ACCGTCGATAT				
		TGCTGAGTCC				
		ACCCGCCGTA				
		TTGCGGCAAG				
		TCGTATTCCGG				
		CTGATCACATG				
		GTGCTGATGG				
		CAGGTTTCACC				
		GCCGGTAATG				
		AAAAAGGCGA				
		ACTGGTGGTG				
		CTTGGACGCA				
		ACGGTTCCGA				
		CTACTCTGCTG				
		CGGTGCTGGC				
		TGCCTGTTTAC				
		GCGCCGATTG				
		TTGCGAGATTT				
		GGACGGACGT				
		TGACGGGGTC				
		TATACCTGCGA				
		CCCGCGTCAG				
		GTGCCCCGATG				
		CGAGGTTGTT				
		GAAGTCGATG				
		TCCTACCAGG				
		AAGCGATGGA				
		GCTTTCCTACT				
TCGGCGCTCT						
AGGTCAGGCC						

1	20160	GATTCATTTCGG	8	350	
		GATGGTCTGT			
		GTGGATTAAAA			
		AAAGAGTGTCT			
		GATAGCAGCTT			
		CTGAACTGGTT			
		ACCTGCCGTG			
		AGTAAATTTAAA			
		ATTTTATTGAC			
		TTAGGTCACCTA			
		AATACTTTAAC			
		CAATATAGGCA			
		TAGCGCACAG			
		ACAGATAAAAA			
		TTACAGAGTAC			
ACAACATCCTC					
AAAGCCTACC					
GGTGACAGTG					
CGGGCTTTTTT					
TTCGACCAAAG					
GTAACGAGGT					
AACAACCATGC					
GAGTGTTGAA					
GTCAGGAGAT					
CCTAAAGGCC					
TGTACCCGTTA					
CCTAGCCAGTT					
GGCATTTAAC					
GTATACGGTAC					
CTAGGCATGTA					
CGTAATCGTAG					
CCTTAGCAATC					
TCCAGTCC					

Start Positio n	End Positio n
871	1220
4144	4493
6008	6357
9464	9813
11173	11522
13704	14053
16623	16972
19422	19771

1	20160	GCTTTTCATTC	8	350	<table><tr><th>Start Position</th><th>End Position</th></tr><tr><td>1</td><td>350</td></tr><tr><td>1316</td><td>1665</td></tr><tr><td>2021</td><td>2370</td></tr><tr><td>5041</td><td>5390</td></tr><tr><td>8441</td><td>8790</td></tr><tr><td>13123</td><td>13472</td></tr><tr><td>15543</td><td>15892</td></tr><tr><td>18132</td><td>18481</td></tr></table>	Start Position	End Position	1	350	1316	1665	2021	2370	5041	5390	8441	8790	13123	13472	15543	15892	18132	18481
		Start Position				End Position																	
		1				350																	
		1316				1665																	
		2021				2370																	
		5041				5390																	
		8441				8790																	
		13123				13472																	
		15543				15892																	
		18132				18481																	
		TGACTGCAAC																					
		GGGCAATATG																					
		TCTCTGTGTGG																					
		ATTAAAAAAG																					
		AGTGTCTGATA																					
		GCAGCTTCTG																					
AACTGGTTACC																							
TGCCGTGAGT																							
AAATTAAAATT																							
TTATTGACTTA																							
GGTCACTAAAT																							
ACTTTAACCAA																							
TATAGGCATAG																							
CGCACAGACA																							
GATAAAAATTA																							
CAGAGTACAC																							
AACATCCATGA																							
AACGCATTAGC																							
ACCACCATTAC																							
CACCACCATCA																							
CCATTACCACA																							
GGTAACGGTG																							
CGGGCTGACG																							
CGTACAGGAA																							
ACACAGAAAAA																							
AGCCCGCACCC																							
TGACAGTGCG																							
GGCTTTTTTTTT																							
CGACCAAAGG																							
TAACGAGGTAA																							
CAACCATGCG																							
AGTGTTGAAGT																							

### SRGD Performance Metrics

Initial Reads Count:	2205
Reads Count after Removing Duplication:	781

Reads Count after Overlapping:	577
Initial Dataset Size in (Base):	440865
Dataset Siset after Removing Duplication:	178158
Dataset Size after Overlapping:	121186

Overlapping Metrics

Repeat Identification Time:	00h:00m:14s:000ms
Overlapping Time:	00h:03m:42s:000ms
Reads Alignment Time:	00h:01m:30s:000ms
Total Hybrid Assembly Time:	00h:05m:12s:000ms
Repeat Annotation Time:	00h:26m:42s:000ms

Overlapping Length	O(N)2 Time Complexity	Hit Index Count	Overlapping Matched Count	RI
196	177241	58	1	0
192	176400	65	1	0
181	175561	61	1	0
176	174724	433	195	0
151	49729	14	6	0