

Bioinformatic Analysis Report

Genome Project Solutions
1024 Promenade Street
Hercules, CA 94547

Informatics@GenomeProjectSolutions.com
<http://www.GenomeProjectSolutions.com>
877-867-0146

Project title: Chlorella Genome Sequence
Prepared for: Dr. Michael Guarnieri
Type of report: Assembly Summary Report
Date: October 22, 2011
Details: Summary Report for the complete assembly (with scaffolding)
All 10,201 scaffolds, including those <5x

Table of contents

| | |
|---|---|
| 1. Summary mapping report | 3 |
| 1.1 Summary statistics | 3 |
| 1.2 General algorithm parameters | 3 |
| 1.3 Reads parameters | 3 |
| 1.4 Distribution of read length | 4 |
| 1.5 Distribution of matched read length | 5 |
| 1.6 Distribution of non-matched read length | 6 |
| 1.7 Paired reads distance distribution | 7 |

1. Summary mapping report

1.1 Summary statistics

| | Count | Average length | Total bases |
|---------------------|-------------|----------------|----------------|
| Reads | 165,874,962 | 88.75 | 14,720,931,624 |
| Matched | 162,897,050 | 88.73 | 14,453,402,888 |
| Not matched | 2,977,912 | 89.84 | 267,528,736 |
| References | 10,201 | 3,896 | 39,746,494 |
| Reads in pairs | 158,729,290 | 199.08 | |
| Broken paired reads | 4,167,760 | 81.17 | |

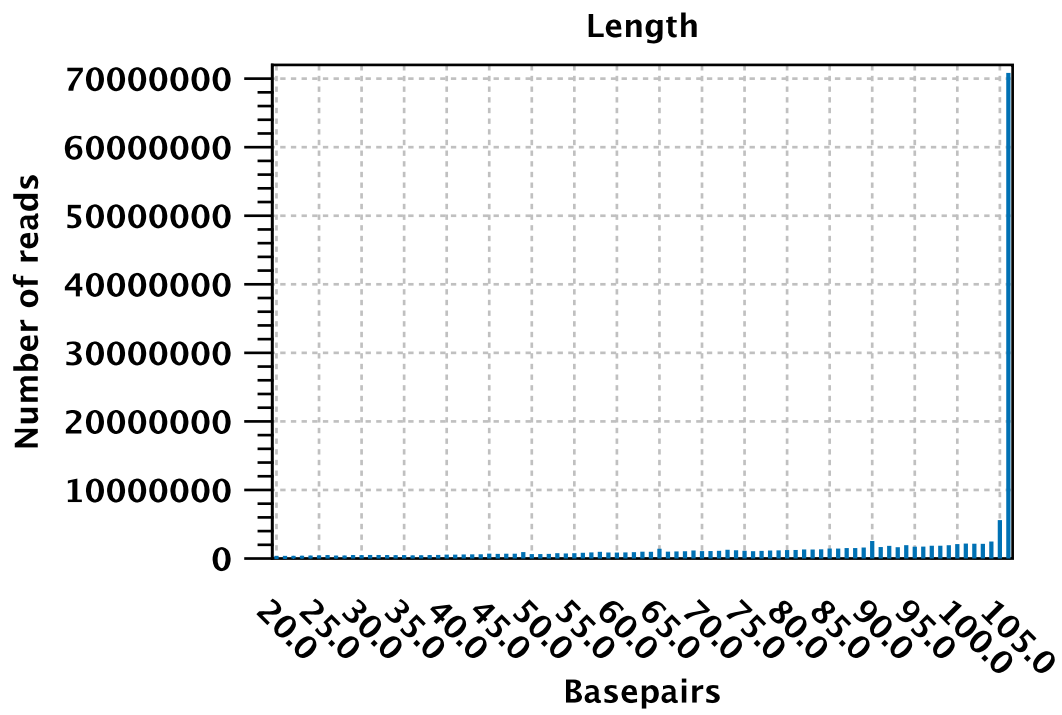
1.2 General algorithm parameters

| Parameter | Value |
|-----------------------|-------------------|
| Conflict resolution | Vote (A, C, G, T) |
| Non specific matches | random |
| Masking of references | none |

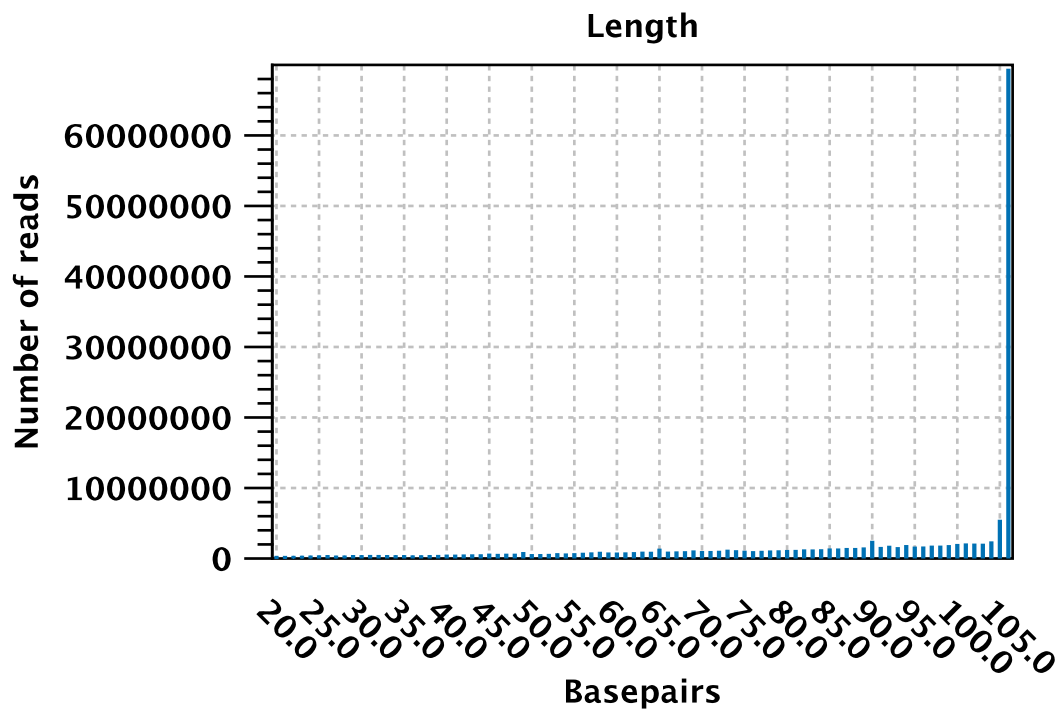
1.3 Reads parameters

| Reads | Length | Type | Parameters |
|-------------------------------|--------|--------|----------------------------|
| s_7_sequence trimmed (paired) | Long | Paired | Length=0.8, Similarity=0.9 |

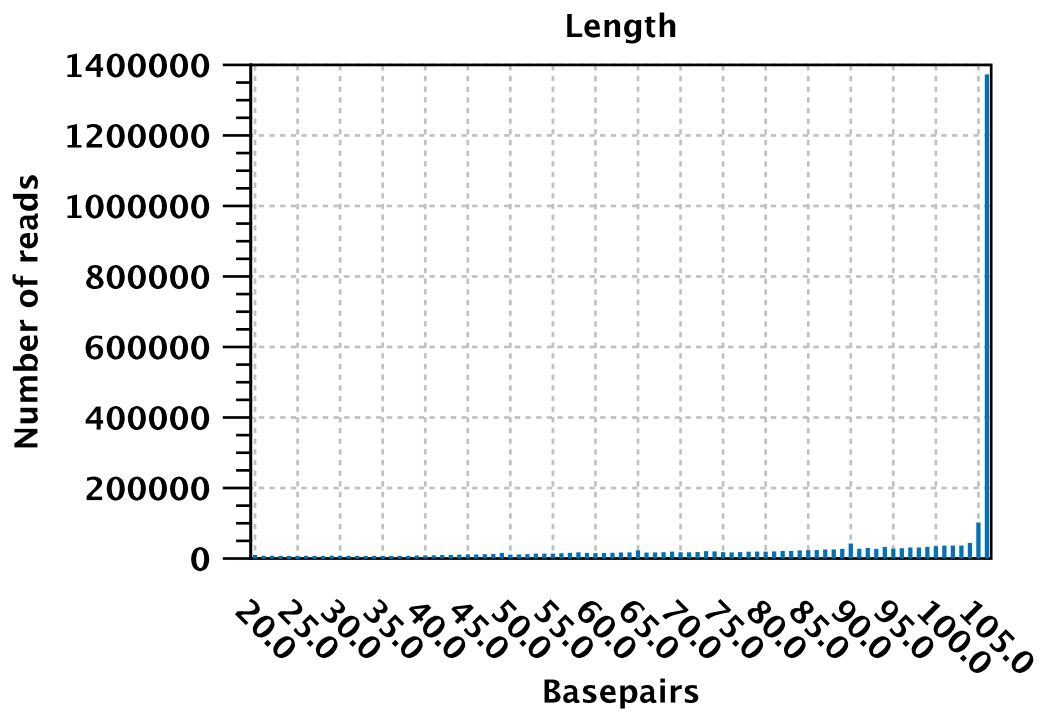
1.4 Distribution of read length



1.5 Distribution of matched read length



1.6 Distribution of non-matched read length



1.7 Paired reads distance distribution

