

Assemblathon Summary Report

An overview of UC Davis results from Assemblathon 1: 2010/2011

Written by Keith Bradnam with results, analysis, and other contributions from Ian Korf, Joseph Fass, Aaron Darling, Ken Yu, Vince Buffalo, Andrew Tritt, and Dawei Lin.

[Genome Center](#), UC Davis, Davis, CA, 95616

Visit <http://assemblathon.org> to download various results and also to download Perl scripts used to produce those results.

Version 0.6 - April 2011

This report is censed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).



Introduction	6
The raw data	6
Participating groups	7
Results overview	9
Basic metrics	10
<i>N50 analysis</i>	<i>12</i>
Fragment analysis	15
<i>Unpaired fragments</i>	<i>16</i>
<i>Fragments with repeats</i>	<i>17</i>
<i>Paired fragments</i>	<i>19</i>
Gene length analysis	21
Bacterial contamination	23
Principal Components Analysis	26
Appendix 1 - basic assembly metrics	29
A1_scaffolds.fa.gz	29
<i>B1_scaffolds.fa.gz</i>	<i>31</i>
<i>B2_scaffolds.fa.gz</i>	<i>32</i>
<i>C1_scaffolds.fa.gz</i>	<i>33</i>
<i>C2_scaffolds.fa.gz</i>	<i>34</i>
<i>D1_scaffolds.fa.gz</i>	<i>35</i>
<i>D2_scaffolds.fa.gz</i>	<i>36</i>
<i>D3_scaffolds.fa.gz</i>	<i>37</i>
<i>D4_scaffolds.fa.gz</i>	<i>38</i>
<i>E1_scaffolds.fa.gz</i>	<i>39</i>

<i>E2_scaffolds.fa.gz</i>	40
<i>E3_scaffolds.fa.gz</i>	41
<i>F1_scaffolds.fa.gz</i>	42
<i>F2_scaffolds.fa.gz</i>	43
<i>F3_scaffolds.fa.gz</i>	44
<i>F4_scaffolds.fa.gz</i>	45
<i>F5_scaffolds.fa.gz</i>	46
<i>G1_scaffolds.fa.gz</i>	47
<i>H1_scaffolds.fa.gz</i>	48
<i>H2_scaffolds.fa.gz</i>	49
<i>H3_scaffolds.fa.gz</i>	50
<i>H4_scaffolds.fa.gz</i>	51
<i>H5_scaffolds.fa.gz</i>	52
<i>I1_scaffolds.fa.gz</i>	53
<i>I2_scaffolds.fa.gz</i>	54
<i>J1_scaffolds.fa.gz</i>	55
<i>K1_scaffolds.fa.gz</i>	56
<i>K2_scaffolds.fa.gz</i>	57
<i>K3_scaffolds.fa.gz</i>	58
<i>L1_scaffolds.fa.gz</i>	59
<i>M1_scaffolds.fa.gz</i>	60
<i>M2_scaffolds.fa.gz</i>	61
<i>M3_scaffolds.fa.gz</i>	62
<i>M4_scaffolds.fa.gz</i>	63
<i>M5_scaffolds.fa.gz</i>	64

<i>N1_scaffolds.fa.gz</i>	65
<i>N2_scaffolds.fa.gz</i>	66
<i>N3_scaffolds.fa.gz</i>	67
<i>O1_scaffolds.fa.gz</i>	68
<i>P1_scaffolds.fa.gz</i>	69
<i>Q1_scaffolds.fa.gz</i>	70
<i>V1_scaffolds.fa.gz</i>	71
<i>V2_scaffolds.fa.gz</i>	72
<i>V3_scaffolds.fa.gz</i>	73
<i>V4_scaffolds.fa.gz</i>	74
<i>V5_scaffolds.fa.gz</i>	75
<i>V6_scaffolds.fa.gz</i>	76
<i>W10_scaffolds.fa.gz</i>	77
<i>W11_scaffolds.fa.gz</i>	78
<i>W1_scaffolds.fa.gz</i>	79
<i>W3_scaffolds.fa.gz</i>	80
<i>W5_scaffolds.fa.gz</i>	81
<i>W6_scaffolds.fa.gz</i>	82
<i>W7_scaffolds.fa.gz</i>	83
<i>W8_scaffolds.fa.gz</i>	84
<i>W9_scaffolds.fa.gz</i>	85
<i>X1_scaffolds.fa.gz</i>	86
<i>X2_scaffolds.fa.gz</i>	87
<i>X3_scaffolds.fa.gz</i>	88
<i>X4_scaffolds.fa.gz</i>	89

<i>X5_scaffolds.fa.gz</i>	90
<i>X6_scaffolds.fa.gz</i>	91
Appendix 2	92

Introduction

The Assemblathon is an international challenge concerned with improving the field of genome assembly. The first Assemblathon event of 2010/2011 invited groups to attempt to assemble synthetic Illumina reads derived from a synthetic ~110 MB diploid genome. Each group was allowed to produce up to five different genome assemblies and these were submitted to the Assemblathon organizers as a set of scaffolds for evaluation. In total we received 41 entrants from 17 different groups.

The goal of this first Assemblathon was to compare how different assemblers perform when all using exactly the same input data. A secondary, but equally important, aspect to the Assemblathon is to see whether new metrics can be used that more accurately reflect the quality and/or usefulness of genome assembly when compared to traditional assembly metrics such as N50 scaffold length.

It is hoped that the Assemblathon will become an annual event and that we can test different datasets, such as transcriptome or metagenome data, as well as 'traditional' genome data. We also would like to attempt to reflect the ever changing use of data from different next-generation sequencing technologies. Assemblathon 2 will be starting later this year (2011) and will feature real genome data from three different vertebrate species.

The raw data

Each group was provided with read data for the diploid genome of 'Species A'. This synthetic genome was constructed using the [EVOLVER](#) suite of software (Edgar, R.C., Asimeno, G., Batzoglu, S., and Sidow, A.). The raw data comprised synthetic Illumina data from four different libraries:

1. paired read data, 40x coverage, insert size 200 nt
2. paired read data, 40x coverage, insert size 300 nt
3. mate pair data, 20x coverage, insert size 3,000 nt
4. mate pair data, 20x coverage, insert size 10,000 nt

All reads were 100 nt in length and included errors as well as a small degree of sequence contamination from a bacterial genome (entrants were not told which species). Additionally, entrants had access to the genome of a 'sister' species ('Species B'). This species was produced as part of the same artificial genome simulation that produced Species A. The two genomes are separated by approximately 100 million years of evolution. The data for Species B was available as a ~110 Mbp diploid genome sequence (in FASTA format), and a set of gene annotations from that genome (in GFF format).

Participating groups

Each group could submit up to five assemblies. The first group who submitted an assembly was named 'Team A', the second group 'Team B', and so on. Individual assemblies from each team were numbered 1–5. E.g. M5 represents the fifth assembly submitted by group M. Please use the following table as a guide when viewing results in the rest of this report.

Team	Team details	Number of entries	Software tools used for assembly
A	GIS_CMB1 - Agency for Science Technology and Research, Singapore	1	paired end-Assembler
B	Phusion - Wellcome Trust Sanger Institute, UK	2	Phusion2, phrap
C	Ensembl Genomes Curtian - European Bioinformatics Institute, UK	2	SGA, BWA, Curtian, Velvet
D	SGA - Wellcome Trust Sanger Institute, UK	4	SGA
E	Borgs - Center for Research in Advanced Computing Systems (CRACS), Portugal	3	ABYSS
F	ABYSS - BC Cancer Genome Sciences Center, Canada	5	ABYSS, Anchor
G	Plant Genome Assembly Group - DOE Joint Genome Institute, USA	1	Meraculous
H	Team Symbiose - Institut de recherche en informatique et systemes aléatoires (L'IRISA), France	5	Monument
I	Terrapins - Cold Spring Harbor Laboratory, USA	2	Quake, Celera, Bambus
J	Department of Computer Science, Iowa State University, USA	1	PCAP
K	Super Dawgs - Computational Systems Biology Laboratory, University of Georgia, USA	3	Seqclean, SOAPdenovo

Team	Team details	Number of entries	Software tools used for assembly
L	PRICE - de Risi Lab, UC San Francisco, USA	1	PRICE
M	Softberry - Royal Holloway, University of London, UK	5	OligoZip
N	TGAC/TSL/Oxford - The Genome Analysis Center, Sainsbury Laboratory, Wellcome Trust Center for Human Genetics, UK	3	Coretex_con_rp
O	KAS - Department of Computer Science, University of Chicago, USA	1	Kiki
P	BGI-Shenzhen, China	1	SOAPdenovo
Q	ALLPATHS, Broad Institute, USA	1	ALLPATHS-LG

As well as the 41 submitted assemblies, we additionally generated some internal assemblies, many of which intentionally omitted some of the pairing information for mate pairs and/or paired reads. The rationale was to produce some deliberately bad assemblies which should therefore score badly by any metric which is accurately evaluating assembly quality. These assemblies are named V–X and their details are outlined below:

Assembly	Assembler	Data used in assembly
V1	Velvet	Untrimmed reads, all read libraries, all pairing info
V2	Velvet	Untrimmed reads, all read libraries, no mate pair info
V3	Velvet	Untrimmed reads, all read libraries, no mate pair or paired end info
V4	Velvet	Trimmed reads, all read libraries, all pairing info
V5	Velvet	Trimmed reads, all read libraries, no mate pair info
V6	Velvet	Trimmed reads, all read libraries, no mate pair or paired end info
W1	CLC	Untrimmed reads, all read libraries, all pairing info

Assembly	Assembler	Data used in assembly
W3	CLC	Untrimmed reads, all read libraries, no mate pair or paired end info
W5	CLC	Trimmed reads, all read libraries, no mate pair info
W6	CLC	Trimmed reads, all read libraries, no mate pair or paired end info
W7	CLC	Untrimmed reads, all read libraries, F/F orientation for mate pairs
W8	CLC	Untrimmed reads, all read libraries, F/R orientation for mate pairs
W9	CLC	Untrimmed reads, 300i paired, all others as single
W10	CLC	Untrimmed reads, 300i only, no pairing info
W11	CLC	Trimmed reads, all read libraries, F/F orientation for mate pairs
X1	ABYSS	Untrimmed reads, all read libraries, all pairing info
X2	ABYSS	Untrimmed reads, all read libraries, no mate pair info
X3	ABYSS	Untrimmed reads, all read libraries, no mate pair or paired end info
X4	ABYSS	Trimmed reads, all read libraries, all pairing info
X5	ABYSS	Trimmed reads, all read libraries, no mate pair info
X6	ABYSS	Trimmed reads, all read libraries, no mate pair or paired end info

The use of ABySS for some of our internal assemblies also meant that we would be able to compare the results of assemblies X1–X6 with those of teams E and F, who also used the ABySS assembler.

Results overview

A variety of metrics were devised to test the quality of the submitted genome assemblies. Many of these were simple statistical descriptions of each assembly, but many others made advantage of the fact that the genome sequence of Species A was actually known.

Note that for many measures, we produced results separately against each of the two diploid genome sequences of Species A and then calculated an average. There were no metrics which produced significant differences when using either haplotype 1 or 2 on its own, when compared to using the results averaged from both haplotypes.

In many cases, results were calculated at the level of scaffolds as well as contigs. It should be noted that each team only submitted one file of scaffolds per assembly, and we split scaffolds into contigs at every instance of 25 or more consecutive N characters. This was the value chosen before the Assemblathon started, though some groups may have used different values.

Basic metrics

The `assemblathon_stats.pl` Perl script¹ was used to split scaffolds into contigs and calculate a wide range of basic descriptive measures for each assembly. In a few cases, this script spotted some minor errors and inconsistencies in some assemblies (e.g. duplicate entries, use of non-unique FASTA headers) and the submitting teams were contacted so that we could correct the problems before proceeding. These results are contained in full in Appendix 1 of this report and are also available as part of a large comma-separated values (CSV) file on the website². The following table provides a list of all the measures that this script calculates, along with descriptions (unless deemed unnecessary):

Metric	Description
Number of scaffolds/contigs	Number of scaffolds is simply the number that appears in the submitted FASTA file by each team. Number of contigs is calculated by splitting each scaffold at every occurrence of 25 or more consecutive N characters. If a scaffold did not contain a run of 25 or more consecutive N characters, it was also treated as a contig.
Total size of scaffolds/contigs	Length of all sequences combined (in bp)
Total scaffold length as a proportion of known genome size	Known genome size of 'Species A' was rounded to an approximate value of 112500000 bp for this calculation. Assemblies varied in size from approximately 80–140% of this value.

¹ Available from the results page at <http://assemblathon.org>.

² This file contains all single statistic assembly metrics that we generated.

Metric	Description
Longest/shortest scaffold/contig	
Number of scaffolds/contigs Longer than 500 nt, 1K nt, 10K nt, 1M nt	Values expressed as an absolute value and also as a percentage of the total number of scaffolds/contigs.
Mean scaffold/contig length	
Median scaffold/contig length	
N50 scaffold/contig length	When summing lengths of individual scaffolds/contigs from longest to shortest, the N50 value is the length of the individual scaffold/contig that takes the sum length to 50% or more of the total assembly size. The script can also output other Nx values if desired (e.g. N30 would be the scaffold/contig length at which 30% of the total assembly size has been reached).
L50 scaffold/contig count	This refers to the number of scaffolds/contigs that are longer than the N50 length (this value has also been described elsewhere as 'N50' but we have stuck with the classical definition, and so chose to call this value L50 instead).
NG50 scaffold/contig count	This value is analogous to N50 but rather than use the assembly size as the denominator, it uses the known genome size of 'Species A' (or rather the approximate size of 112500000 bp). This value therefore allows for a comparison between assemblies that may be vary different in size.
LG50 scaffold/contig count	Analogous to L50, but based on the NG50 value (see previous entry).
N50 - NG50 scaffold/contig length difference	The absolute difference between the values of N50 and NG50.
Scaffold/contig %A, %C, %G, %T, %N	
Scaffold/contig %non-ACGTN	Added as a measure to see whether any assemblies included notable amounts of IUPAC ambiguity characters

Metric	Description
Number of scaffold/contig non-ACGTN nt	Added as a complement to the previous measure because some assemblies only include very few ambiguity characters.
Percentage of assembly in scaffolded contigs	What fraction of the total assembly size is taken up by multiple contigs that have been joined together in a scaffold (using runs of 25 or more consecutive Ns as an indicator that a scaffold contains multiple contigs).
Percentage of assembly in unscaffolded contigs	The opposite to the above measure.
Average number of contigs per scaffold	The contig count divided by the scaffold count.
Average length of break between contigs in scaffold	Using 25 or more consecutive N characters as the delimiter.

It is not possible to produce an exhaustive analysis of every one of these metrics, so the following section describes only some selected findings.

N50 analysis

N50 scaffold/contig length is one of the most commonly reported measures when describing a genome assembly. Figure 1 shows the N50 scaffold length for the best assembly from each team³. Assemblies G1 and Q1 far outperform other assemblies when using this metric and some assemblies have N50 lengths are nearly two orders of magnitude lower. The impressive results for the G1 and Q1 assembly are reflected by the longest scaffolds present in these assemblies (23 and 17 Mbp respectively).

³ This is to facilitate an easier visual comparison of the results. 'Best' in this case simply means highest N50 scaffold value.

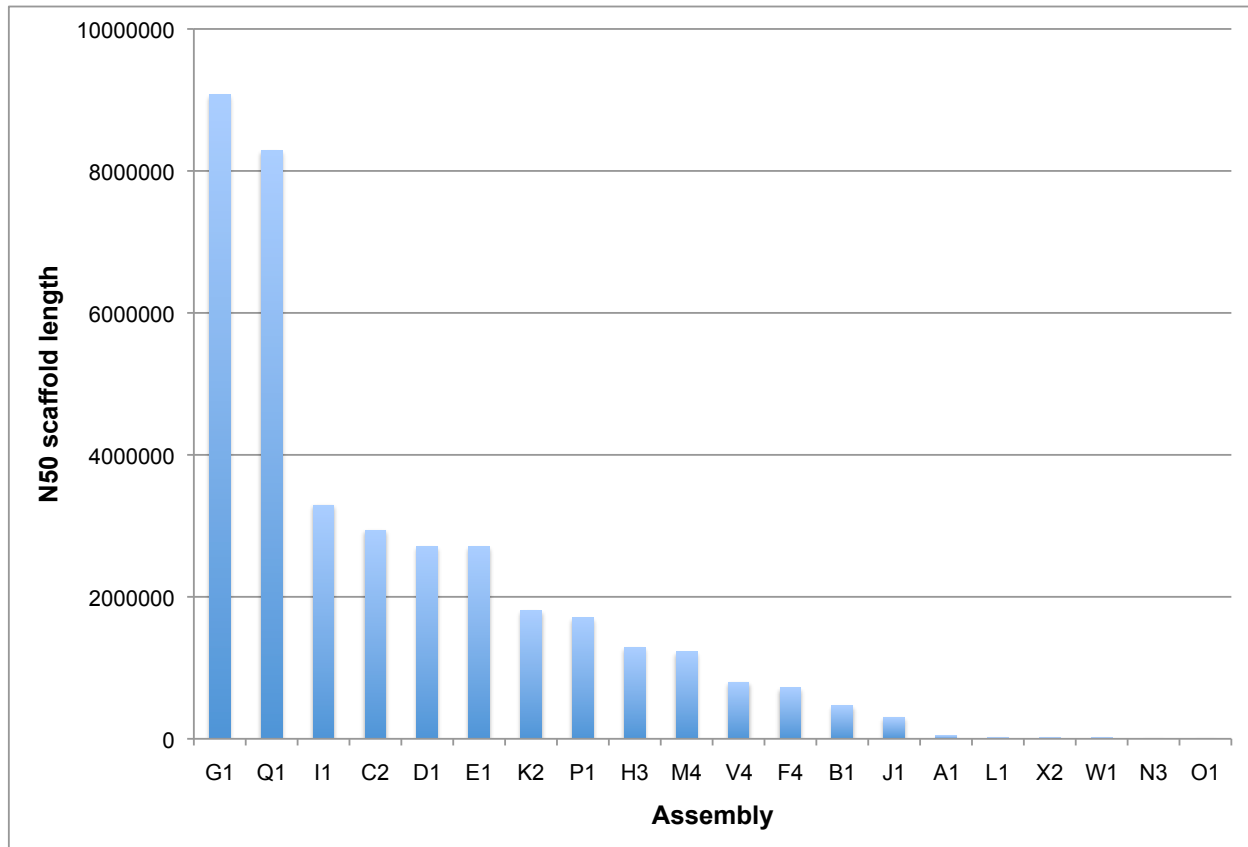


Figure 1. N50 scaffold lengths for best assemblies from each participating team.

It can be misleading to compare N50 values between different assemblies due to the varying sizes of each assembly. E.g. at 170 Mbp, assembly M3 is almost twice the size of assembly O1 (90 Mbp). This is why we developed the NG50 measure which uses the known genome size of Species A in the calculation. This permits a fairer comparison between assemblies and also allows us to see whether any assemblies score perform better when assessed by their N50 scaffold length than their NG50 length. This information is shown in Figure 2 which ranks assemblies based on their difference between N50 and NG50 measures.

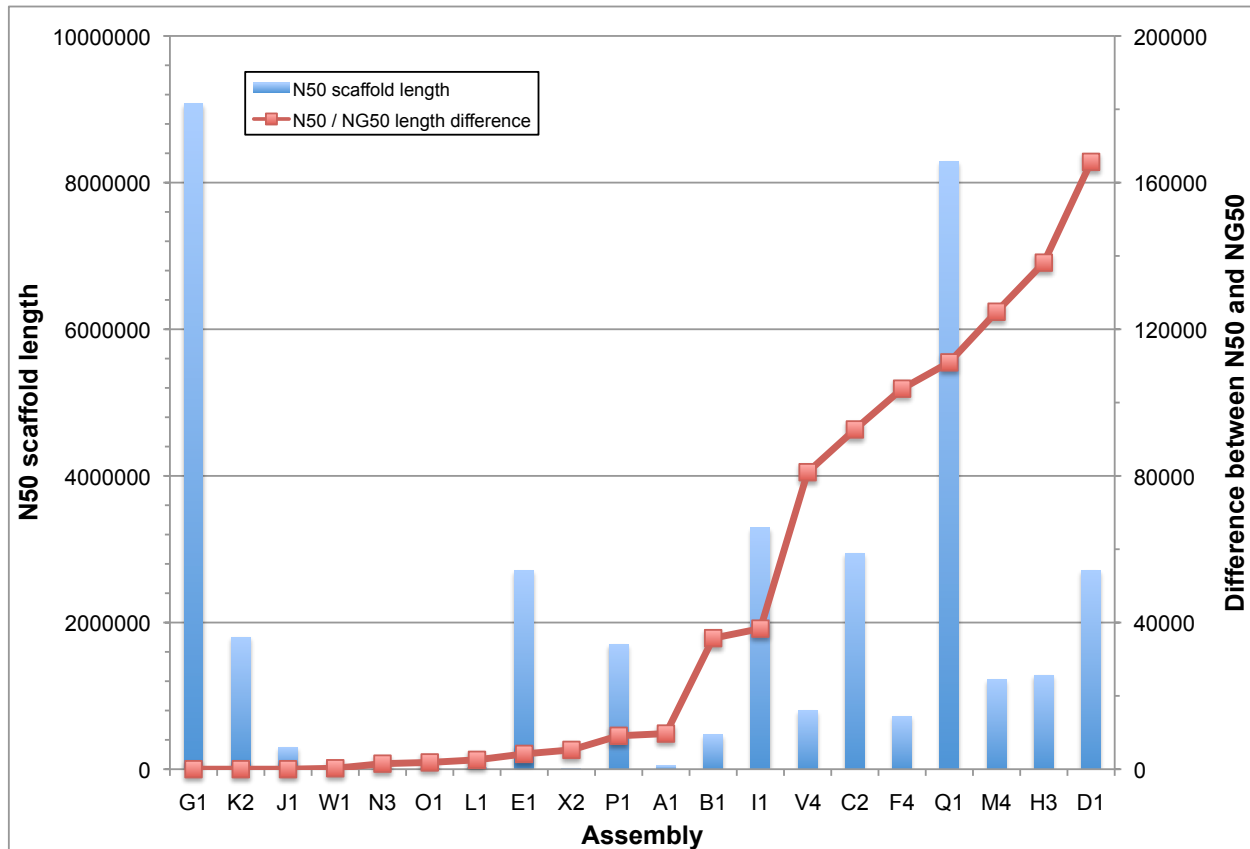


Figure 2. N50 vs NG50. Best assemblies from each team are shown in order of increasing difference between their N50 and NG50 scaffold lengths. The absolute difference is shown by the red line.

For some assemblies with many, very long scaffolds (e.g. G1), it makes no difference as to whether you use N50 or NG50, however with other assemblies the difference becomes more noticeable. As shown in Figure 2, the D1 assembly shows a difference of over 160,000 nt between the two measures. Not shown in Figure 2 are the results for the I2 assembly which is ranked 10th overall based on its N50 scaffold length, but is ranked 4th when considering its NG50 length, which is almost 800,000 nt longer.

Whereas the use of N50 or NG50 metrics gives you a snapshot of how assemblies are performing, it can be far more instructive to look at other Nx measures. Values such as NG60, NG70 etc. reveal more detail about the tail end of the scaffold length distribution. It is possible to calculate all NGx values (e.g. NG1 through to NG99) and then plot these. This is what is shown in Figure 3.

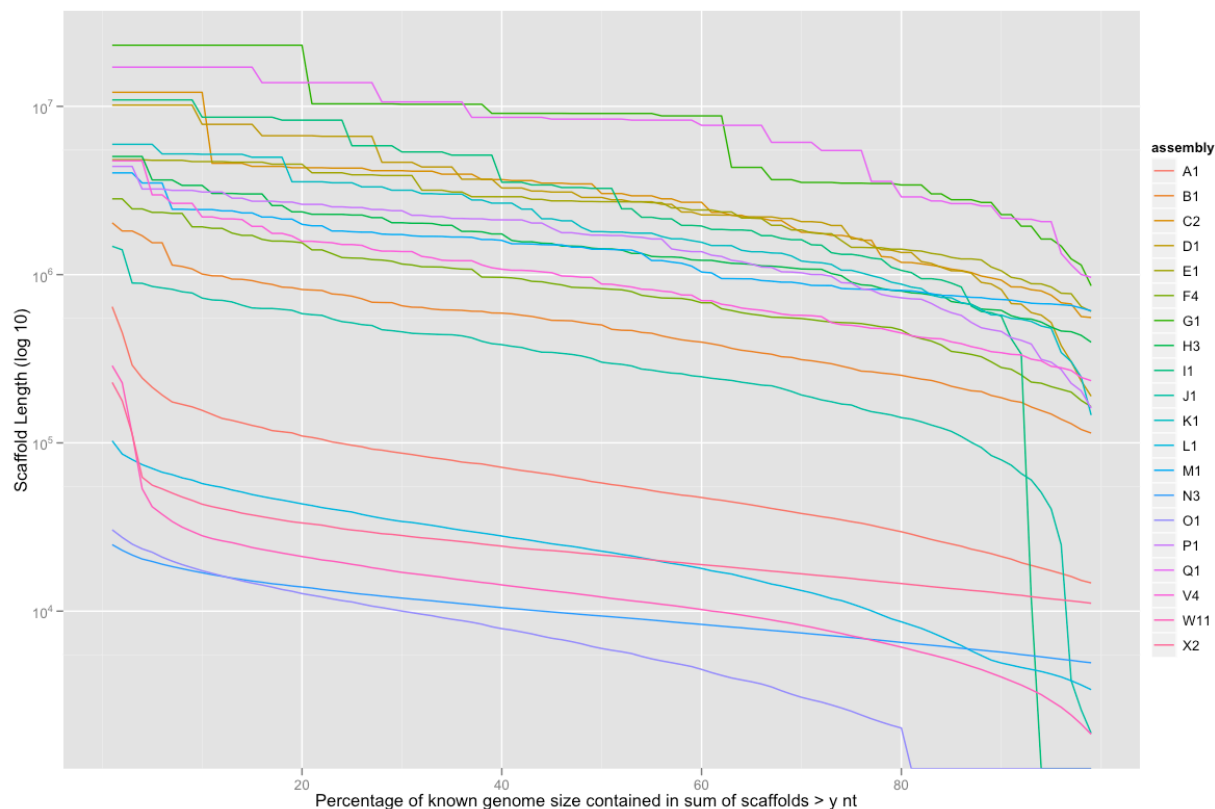


Figure 3. Nx plot of best assemblies from each team. Y-axis shows scaffold length plotted on a log scale.

This 'NGx plot' reveals much more information about the dynamics of scaffold lengths across all assemblies. The long horizontal lines shown in the top left of the graph for assemblies G1 and Q1 are because of the very long scaffolds that are present in these assemblies. Where a line falls to a zero scaffold length (e.g. assembly O1), this is because that assembly contains less sequence than is present in the known genome of Species A. This figure also reveals that while assembly G1 has the highest NG50 value, it performs slightly less well than assembly Q1 at other values (e.g. NG70). If calculated, the area under the curve for each assembly would also serve as a useful metric.

Fragment analysis

We developed a novel approach to characterize the quality and completeness of the genome assemblies by using information from the known genome. Specifically, we wrote a script that first extracts randomly chosen fragments from the known genome of Species A, and then attempts to see whether those fragments are present in each genome assembly. Several variations on this approach were used and these will be outlined in the following pages. However, all approaches rely on using BLAST and we classify a fragment as being present if at least 95% of its length is present in an assembly with at least 95% identity. In all cases, we run this procedure separately using

random fragments from both haplotypes of the Species A genome, and then we take an average. Finally, in all experiments results are presented for 1,000 randomly selected fragments.

Unpaired fragments

In our first variation of this approach, we used a Perl script (blastoff_speciesA.pl) to choose fragments from random positions in the known genome. We started by choosing 1,000 fragments at size 100 nt, and repeatedly doubled the fragment size until we reached 25,600 nt. The results of this analysis can be seen in Figure 4.

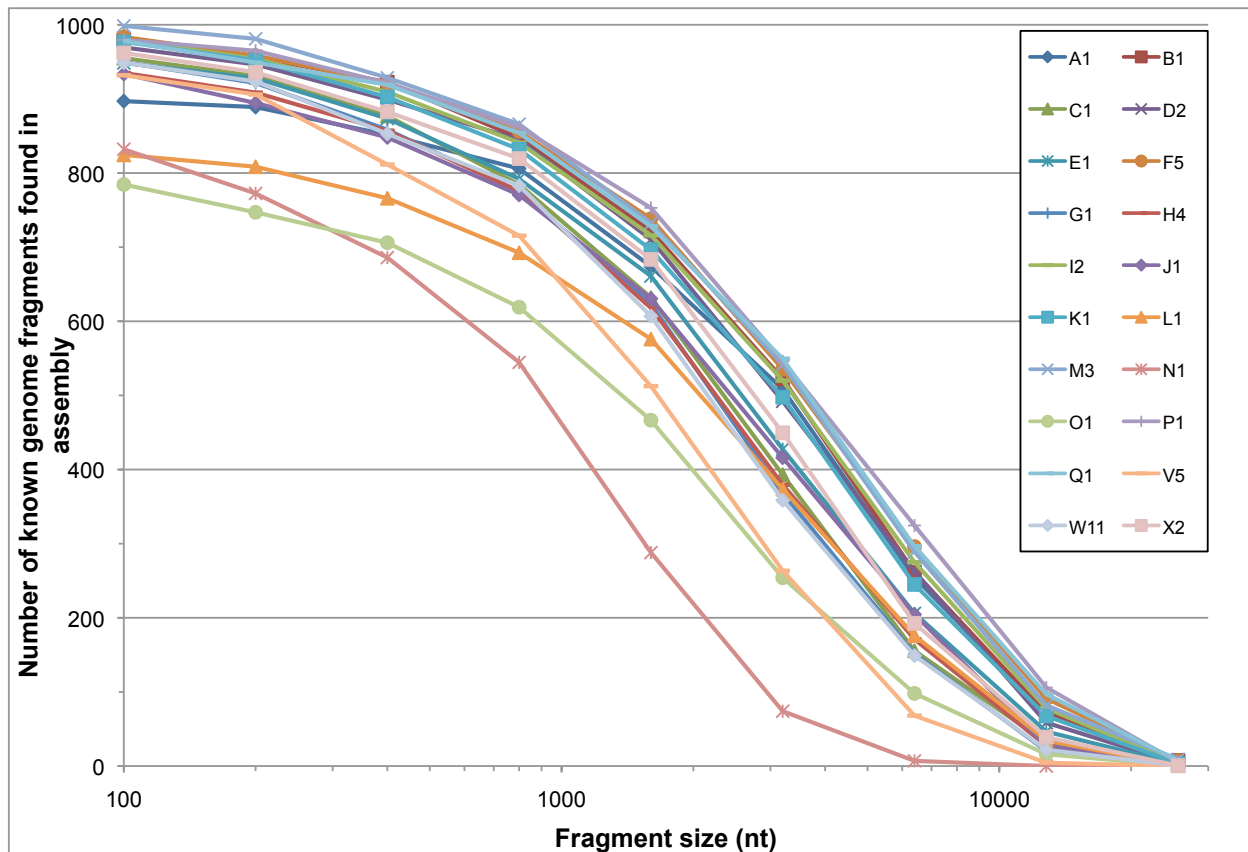


Figure 4. Fragment analysis chart for best assemblies from each team. X-axis is plotted on a logarithmic scale. Each line represents the averaged values from using both haplotypes of the Species A genome.

Even when using fragments as short as 100 nt, there was no assembly that contained all fragments (best = M3, 999 fragments detected; worst = N2, 763 fragments detected). As fragments become longer there is a predictable decline in the numbers of fragments that could still be detected in each assembly. Even at the longest fragment size (25,600 nt), some assemblies contained a small number of fragments (best = M3, 9 fragments detected). We used these results to calculate a Fragment Completeness Index (FCI) which produces a single result for each assembly. These results are shown in Figure 5.

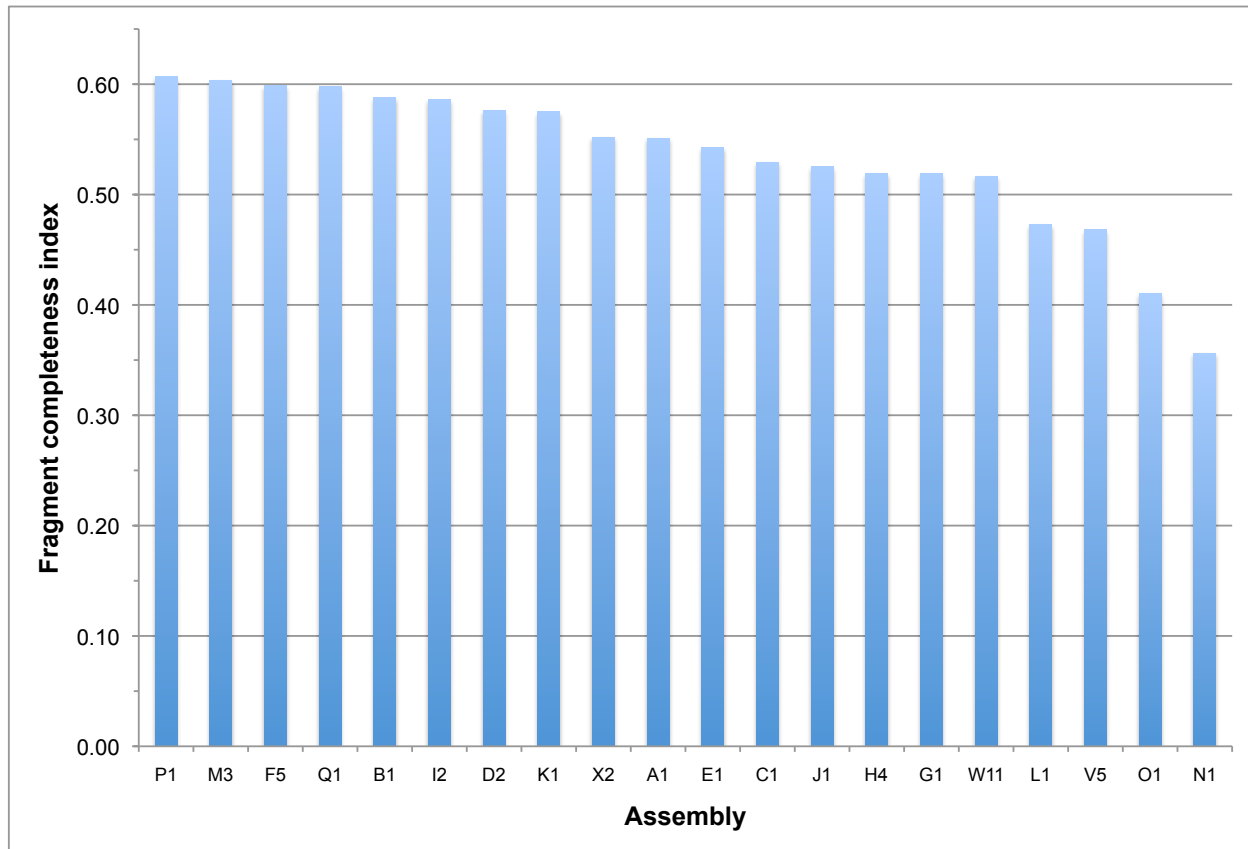


Figure 5. Fragment completeness index (FCI) for best assemblies from each team. FCI is simply calculated as the sum of all fragments detected in each assembly at all fragment sizes, divided by the maximum number that could be detected (9,000).

The results varied from about 32% of all fragments present in assembly N2 to 61% present in assembly P1. The G1 and Q1 assemblies that had both stood out on the basis of their N50 values (Figure 1) showed less agreement when using this metric. Assembly Q1 remains in the top 5 assemblies ranked using the FCI, whereas the G1 assembly places in the bottom third of all assemblies. This initial comparison of just two different assembly metrics suggests that the use of different metrics might help paint a more complete picture of genome assembly completeness and quality.

Fragments with repeats

One of the obvious challenges for genome assemblers is the problem of dealing with repeats that can exist in genomes. Typically, large tandem repeat arrays are not present in most genome assemblies and dispersed repeats such as transposons can lead to misassemblies of contigs. The Species A genome contains a variety of repeats, from homopolymer runs, tandem repeats, and transposons. We used another Perl script (`blastoff_repeats_nonrepeats.pl`) in order to generate two new pools of random sequence fragments from the known genome. The first pool consisted of fragments that overlapped a known repeat. At one extreme, the overlap could be as little as 1 bp; at the

other extreme, the entire fragment could consist of repeat sequence. The second pool of fragments consisted of sequences that did not overlap repeats at all. The goal of this new script was to calculate how many of the 1,000 repeat- and non-repeat containing fragments were present in each assembly. The difference between the results for these two sets of fragments sheds light on how well each assembly deals with repeats. Figure 6 shows results from this approach for two selected assemblies.

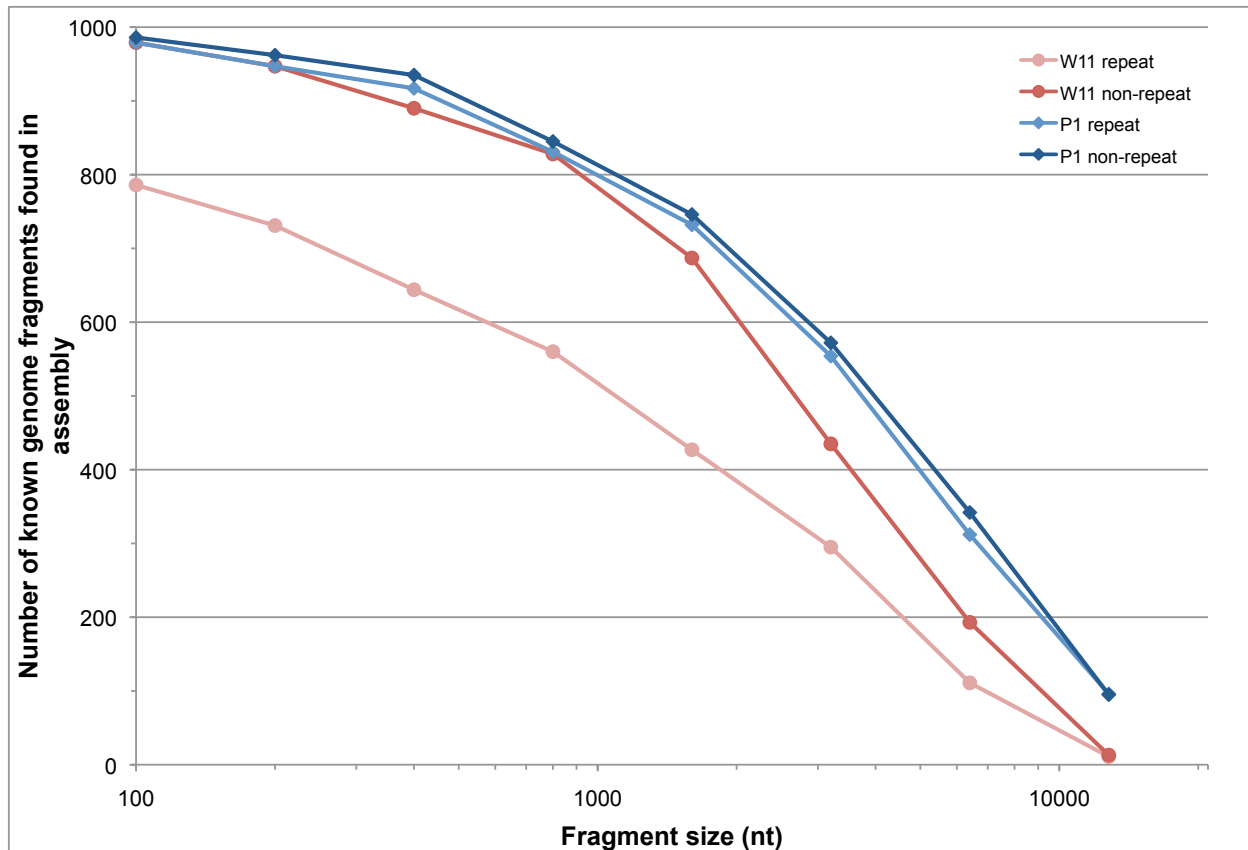


Figure 6. Fragment repeat analysis chart. X-axis is plotted on a logarithmic scale. Each line represents the averaged values from using both haplotypes of the Species A genome. Blue lines refer to P1 assembly, red line refers to W11 assembly.

Using this new metric reveals that assembly P1 is among the best assemblies; fragments from the known genome that overlap repeats are almost as likely to be included in the assembly as fragments that doesn't overlap repeats. In contrast, the W11 test assembly is very poor at dealing with repeats and there is a large difference between the results for repeat- and non-repeat fragments. We calculated a new statistic Assembler Repeat Tolerance (ART) as the sum of the repeat fragment values divided by the sum of the non-repeat fragment values. The results for this new metric are shown in Figure 7.

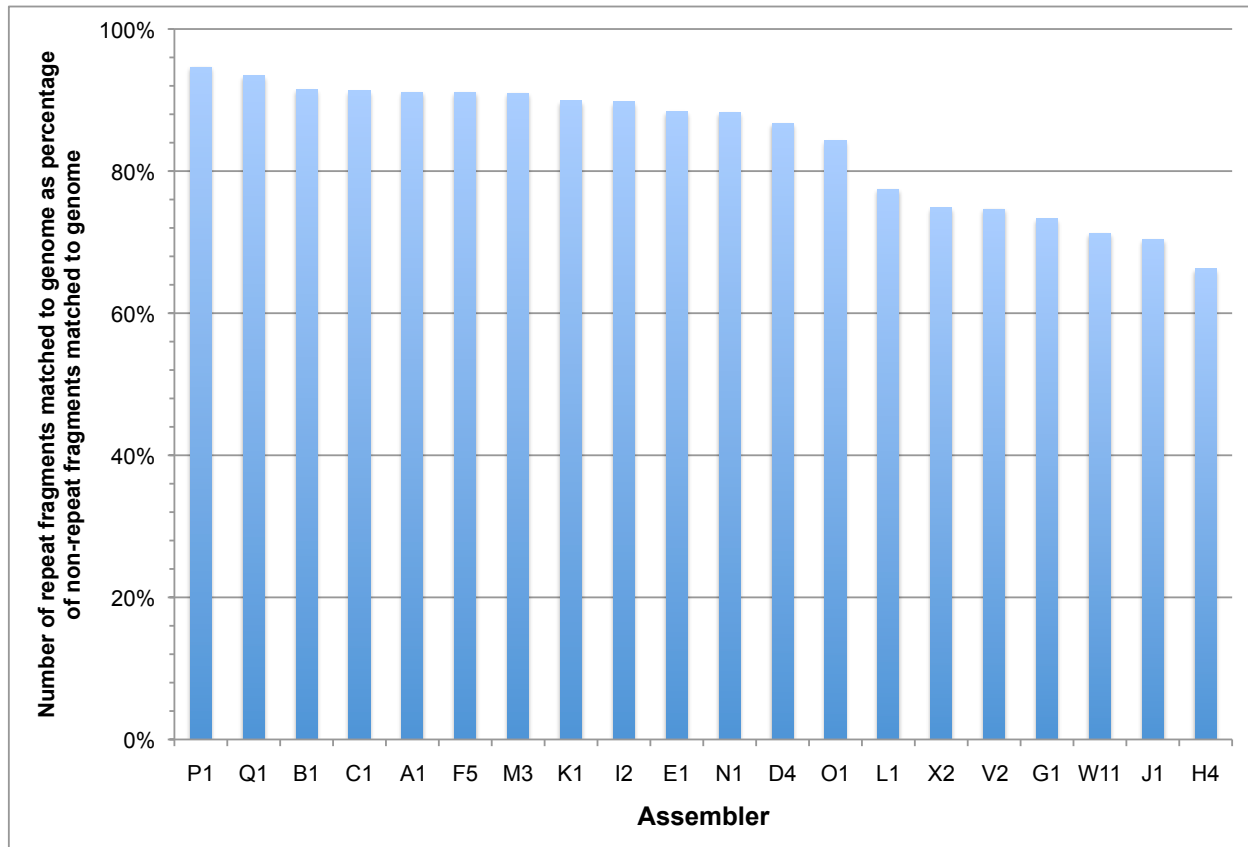


Figure 7. Assembler Repeat Tolerance chart. Results are only shown for the best assembly from each team. A value of 100% would indicate that repeat-containing fragments can be found in an assembly at the same level as non-repeat containing fragments.

These results show that the P1 and Q1 assemblies both deal very well at assembling sequences containing repeats. A number of assemblies performed at a very similar level, and a few assemblies were less successful in the way they dealt with repeat sequences.

Paired fragments

The next approach that we developed used *pairs* of fragments from the known genome that were separated by a specific length. Rather than just ask whether a single fragment from the known genome is present in an assembly, we can instead ask whether a pair of fragments both exist, and whether the distance between those fragments is preserved in the assembly. This approach reveals more information about the longer range dynamics of the assembled contigs and scaffolds. In this analysis, the fragments were always kept at 100 nt, but the distances between those fragments started at 100 nt and was repeatedly doubled. The full criteria to satisfy a 'match' are as follows:

1. Both paired fragments must exist in the assembly at 95% identity with 95% of the length of each fragment present
2. Both fragments must match the same scaffold sequence
3. Both fragments must be in the same orientation
4. The distance between the start of each pair of fragments in the assembly must be between 95–105% of the distance between the fragments in the known genome

This approach helps distinguish between many of the assemblies (Figure 8). At long distances, with paired fragments 102,400 bp apart, assembly Q1 shows only a very slight decline in matching the fragment pairs.

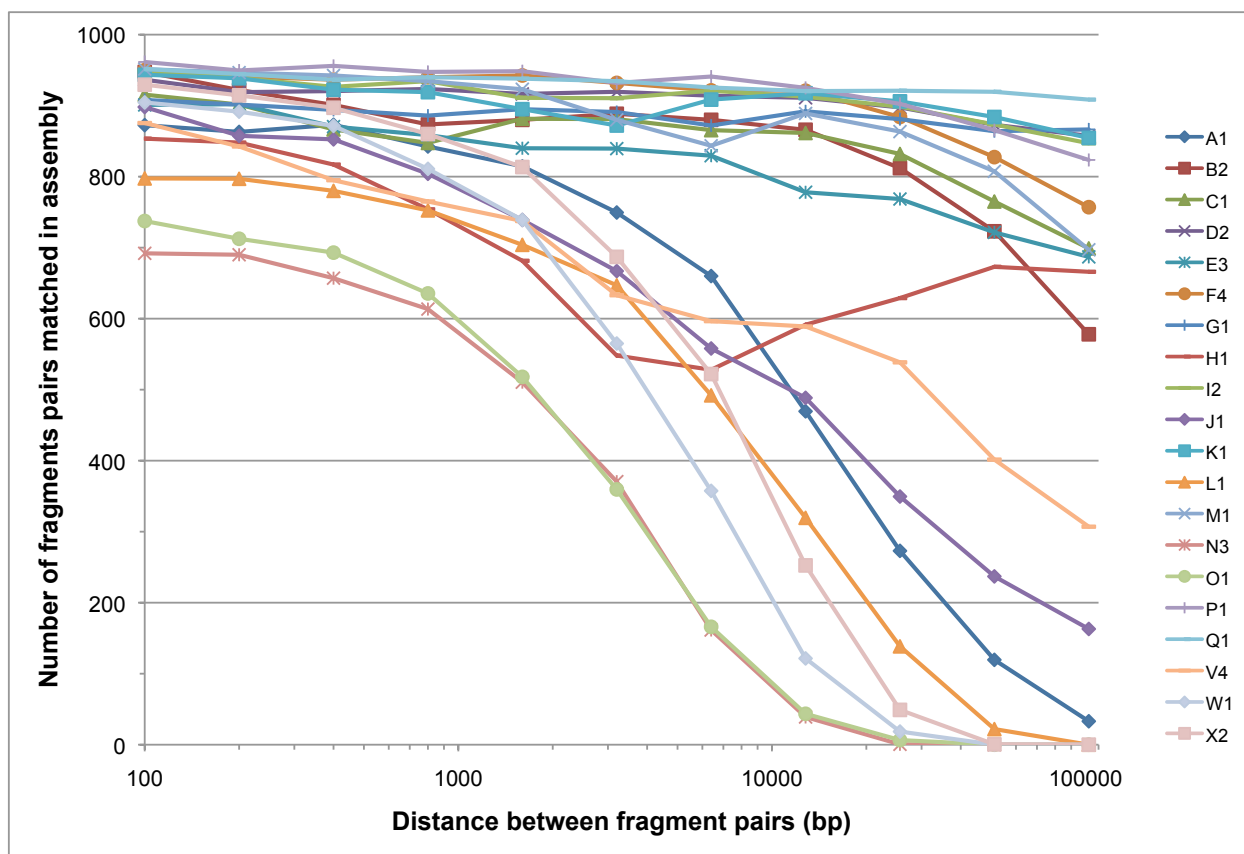


Figure 8. Paired fragment repeat analysis chart. X-axis is plotted on a logarithmic scale. Each line represents the averaged values from using both haplotypes of the Species A genome.

Assembly H1 was the only assembly which showed an unusual pattern of initially containing fewer matched fragments with increasing distance, only to then see an increase at longer distances. It is not clear why this might be the case, and if the distance between fragment pairs is increased even further, then the number of matches continues to decrease once again (data not shown). We used the results of this analysis to calculate a single statistic that we call Paired Fragment Completeness Index (PFCI).

As shown in Figure 9, the PFCI metric more readily distinguishes between assemblies than the simpler FCI metric on which it is based (Figure 5).

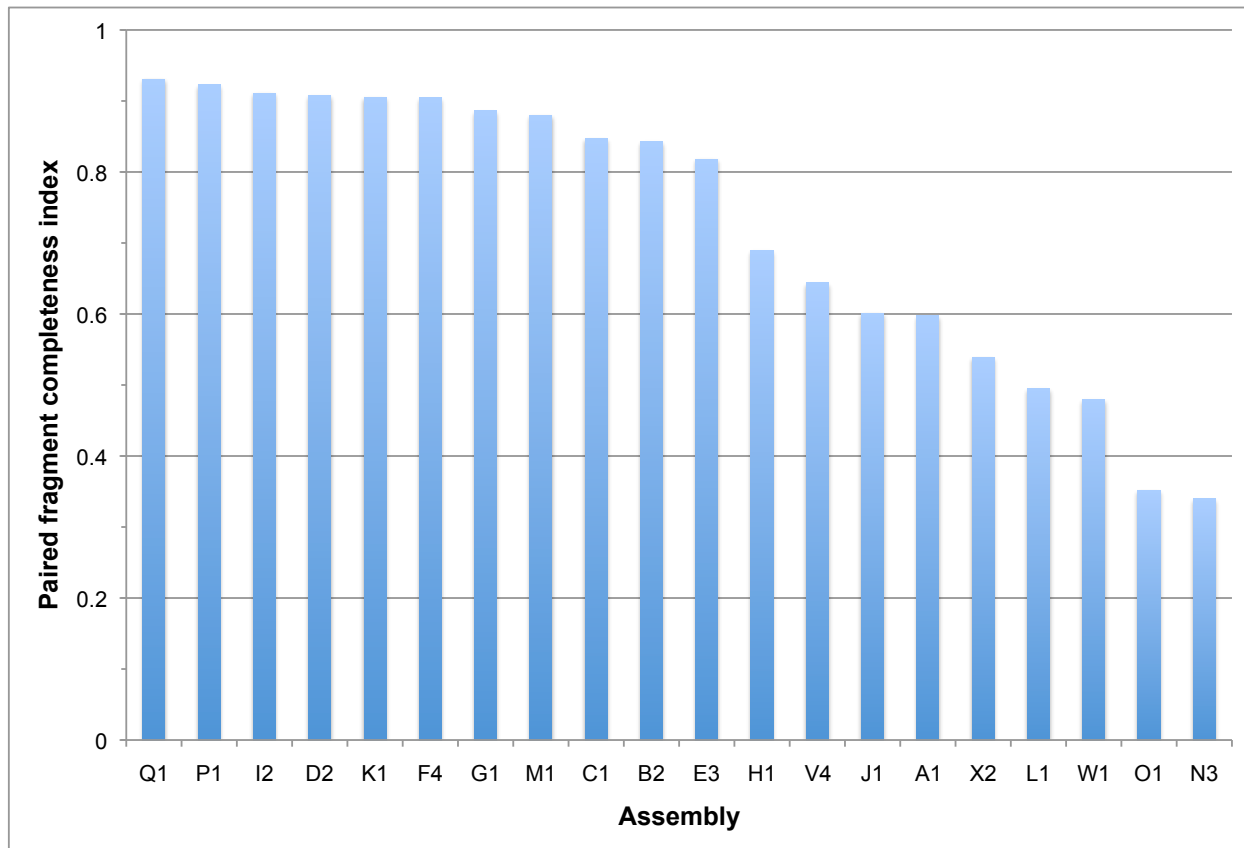


Figure 9. Paired fragment completeness index (PFCI) for best assemblies from each team. PFCI is simply calculated as the sum of all fragments detected in each assembly at all fragment sizes, divided by the maximum number that could be detected (11,000).

Gene length analysis

The genome of Species A contains 176 protein-coding genes and the coordinate details of all exons for these genes are known. This enabled us to simply ask how many genes are present in each assembly. To do this we used BLAST to search the full length transcripts of each gene (start codon to stop codon, inclusive of all introns if present), against each assembly. To count as a match, a gene was counted as being present in an assembly if it exhibited at least 95% identity with at least 95% of the transcript length present in a single scaffold. Because the genome of Species A contains many genes with only 1–3 exons, and a few very long, multi-exon, genes⁴, we decided to first focus on the total length of matching transcripts, rather than the total number of genes (Figure 10).

⁴ 30 genes are single exon, and about a third of all genes have 3 exons or less.

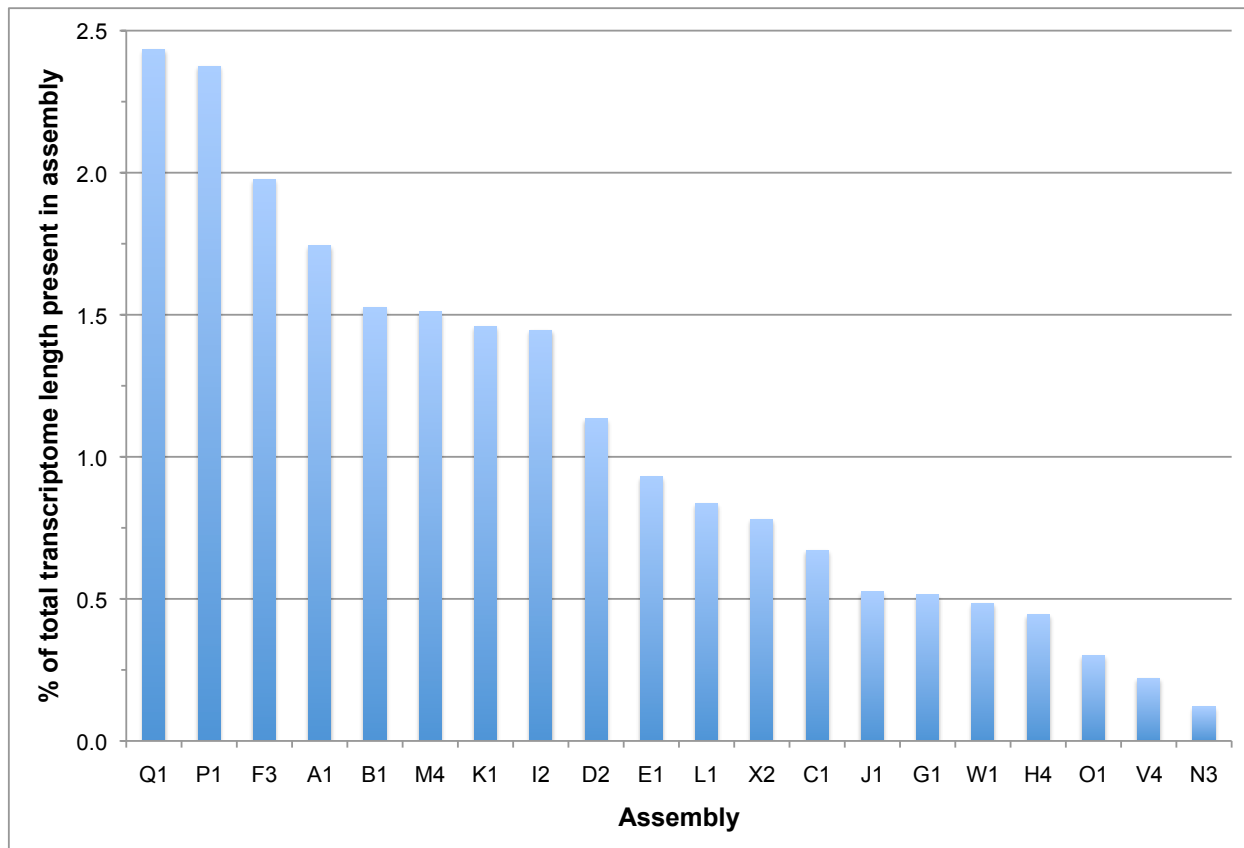


Figure 10. Gene coverage in each assembly. The total length of mapped transcripts is shown as the percentage of the maximum possible length. Results only shown for best assembly from each team.

The gene length analysis reveals a great disparity in the amount of the Species A transcriptome that was present in each assembly. Perhaps surprisingly, even the best assembly (Q1) only captured approximately 2.5% of the maximum gene length possible. This assembly contains 44 or 45 genes (depending on which haplotype you use) whereas other assemblies contained many more genes, but captured less of the total gene space. E.g. assembly L1 contains 55/57 genes but only has 1.75% of the total gene space. The relationship between number of genes and total amount of gene space is shown in Figure 11.

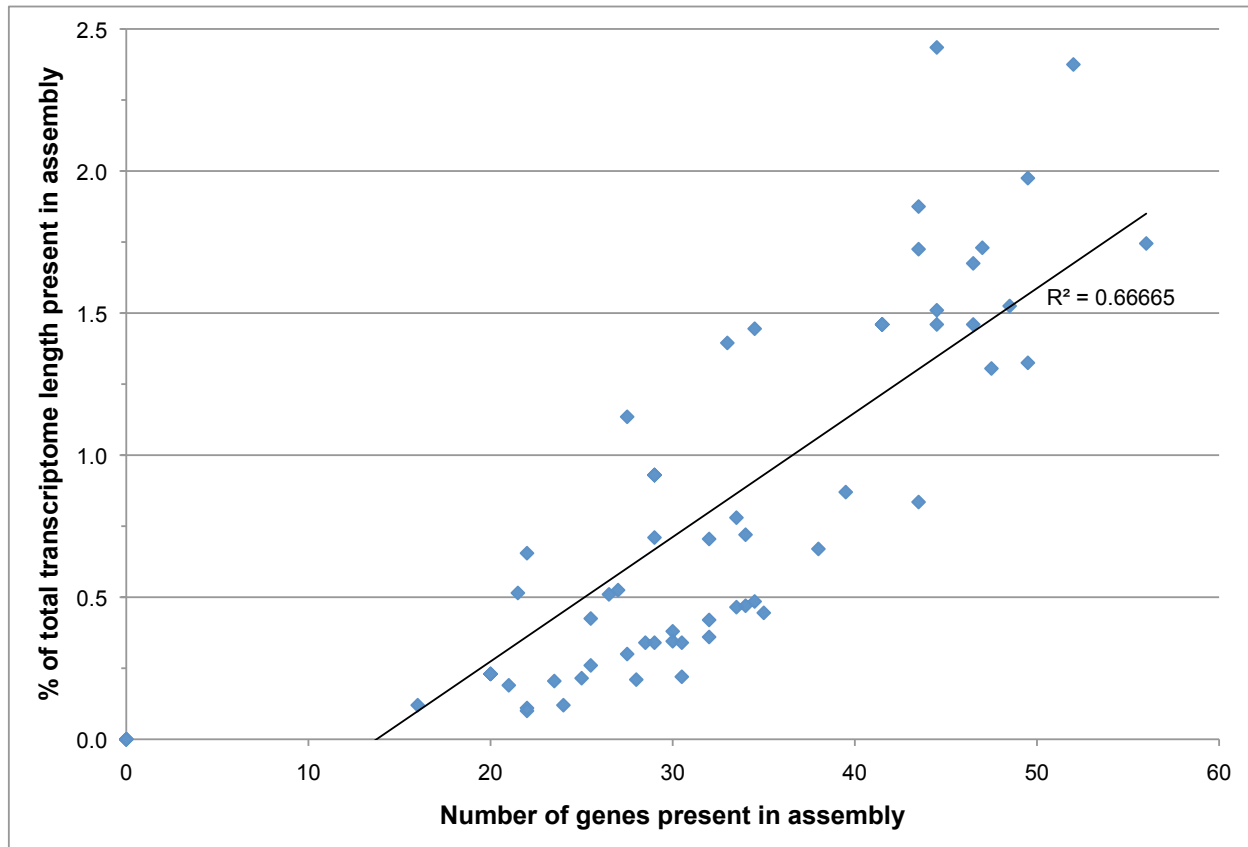


Figure 11. Number of genes vs gene coverage in each assembly. The total length of mapped transcripts is shown as the percentage of the maximum possible length.

We also used a similar approach to look at how many individual exons are present in each assembly. However, the dynamics of the exon length distribution also means that it is misleading to simply count the total number of matched exons. The longest exon in the Species A genome is over 5,000 nt, which is longer than a quarter of all of the genes. Some assemblies which captured more of the gene space did not necessarily capture more of the exon space. E.g. assembly Q1 has about 2.5% of the gene space but only 15.9% of the exon space. Conversely, assembly W10 captured only 0.4% of the gene space but represented over 65% of the exon space.

Bacterial contamination

One piece of information that was provided to participants was that "*all libraries will contain some bacterial contamination*"⁵. The amount and nature of the contamination was not specified, but we were curious to see whether groups would attempt to remove

⁵ This information was included in a README file provided alongside the reads that participants downloaded.

any such contamination as part of their assembly pipeline. We were also curious as to how successful any such attempts at contamination removal would be.

The raw reads that were provided to participants were contaminated with synthetic reads derived from the *Escherichia coli* genome sequence⁶. Approximately 5% of all reads were derived from this bacterial genome, and the high coverage of the libraries meant that there should be sufficient information to separately assembly the *E. coli* genome.

Using BLAST, we investigated what fraction of each assembly constituted *E. coli* sequence. Conversely, we also looked to see what fraction of the *E. coli* genome sequence was present in each assembly. Seven teams submitted assemblies that were free, or almost free, of any contamination (Figure 12). Assemblies from all other teams consisted of approximately 2.75–5.5% *E. coli* sequence.

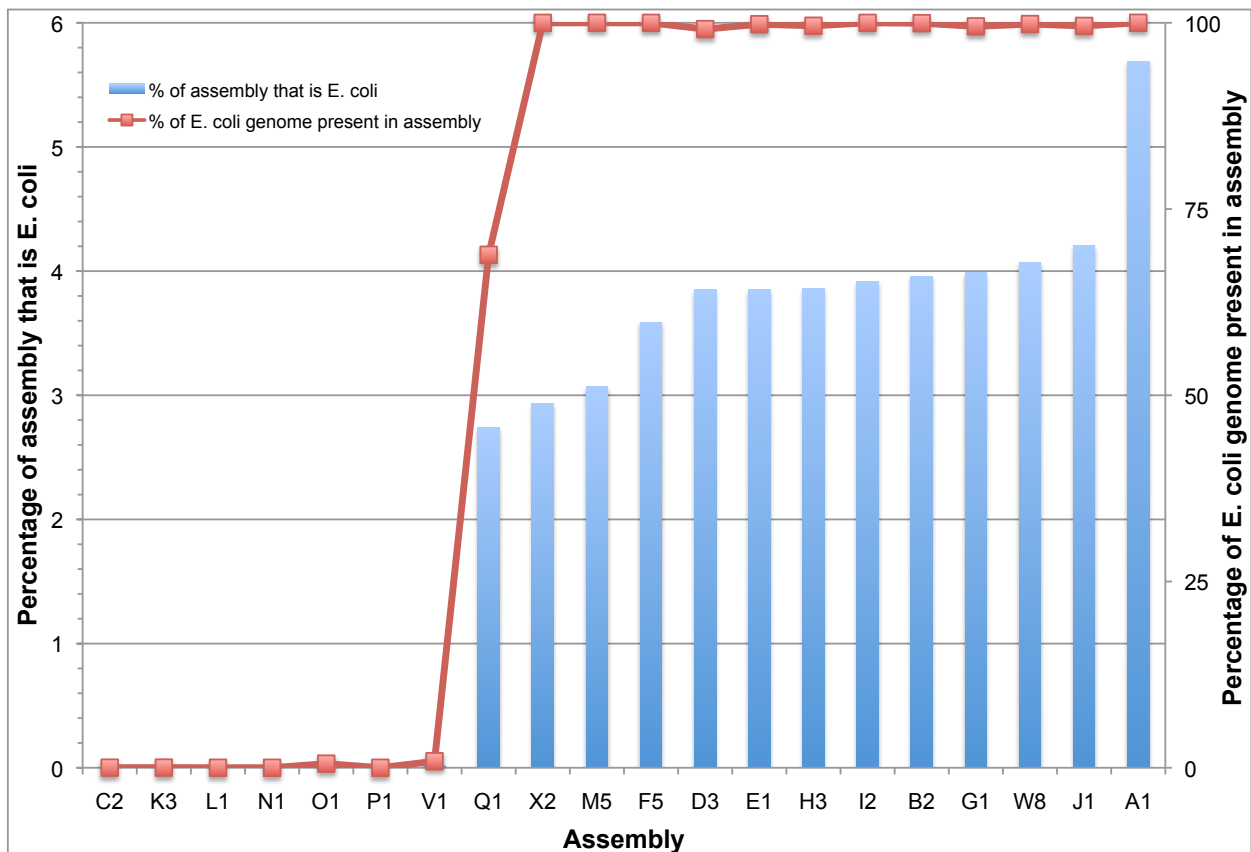
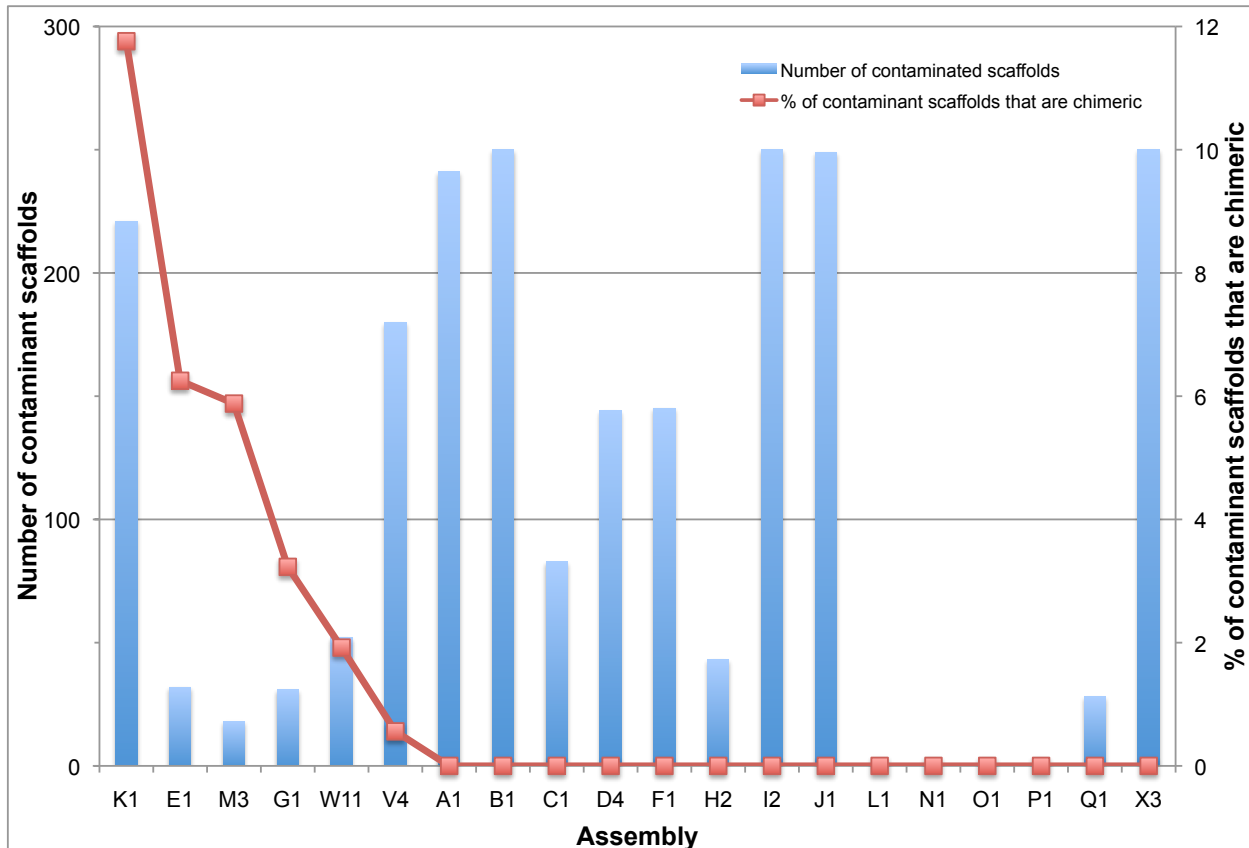


Figure 12. *E. coli* contamination. Series in blue shows percentage of assemblies that is *E. coli* sequence; red line shows the percentage of the *E. coli* genome that is present in each assembly. Only the best assemblies from each team are shown.

⁶ Using the sequence of the K12 strain.

Most assemblies that had contamination resulted in a single copy of the complete *E. coli* genome sequence present in a set of scaffolds. Assembly A1 was unusual in having approximately 1.5 copies of the *E. coli* genome present. Another unusual results was seen in assembly Q1 which had most, but not all, of the *E. coli* sequence present; this implies that some contamination removal occurred but such removal was limited. The only assemblies that were completely free of any contamination were L1, N1–N3, and P1⁷.

A more important consideration of whether there is contamination or not, is how that contamination manifests itself within the scaffolds. Ideally, any contaminant sequence should be self-contained and not mixed in with other 'Species A' sequence. We looked to see whether we could detect 'chimeric' scaffolds which had a mixture of contaminant and non-contaminant sequence⁸. Most assemblies that had contamination did not include any chimeric scaffolds (Figure 13). However, assemblies from six teams had a small proportion of chimeric scaffolds (up to 12% of contaminant scaffolds for assembly K1).



⁷ Many other assemblies were essentially free of contamination but still contained a few hundred to a few thousand bp of *E. coli* sequence.

⁸ Chimeric scaffolds were those which had both types of sequence present (*E. coli* and Species A), and with each type accounting for more than 5% of the entire scaffold.

Figure 13. Chimeric *E. coli* contamination. Series in blue shows number of scaffolds in each assembly which have any degree of *E. coli* contamination; red line shows the percentage of those scaffolds that are chimeric. Only the best assemblies from each team are shown.

Principal Components Analysis

In using so many different assembly metrics there is the possibility that several, unrelated, metrics will correlate in how well they explain variation across the set of assemblies. We investigated whether principal components analysis (PCA) could be used to whether certain metrics explain different facets of the variation that is present. One problem with using PCA is in the selection of which metrics to use. Many metrics that we have generated are clearly not independent of each other (e.g. NG50 scaffold length correlates with N50 scaffold length). Figure 14 shows one PCA plot based on nine different measures which are largely derived from independent analyses.

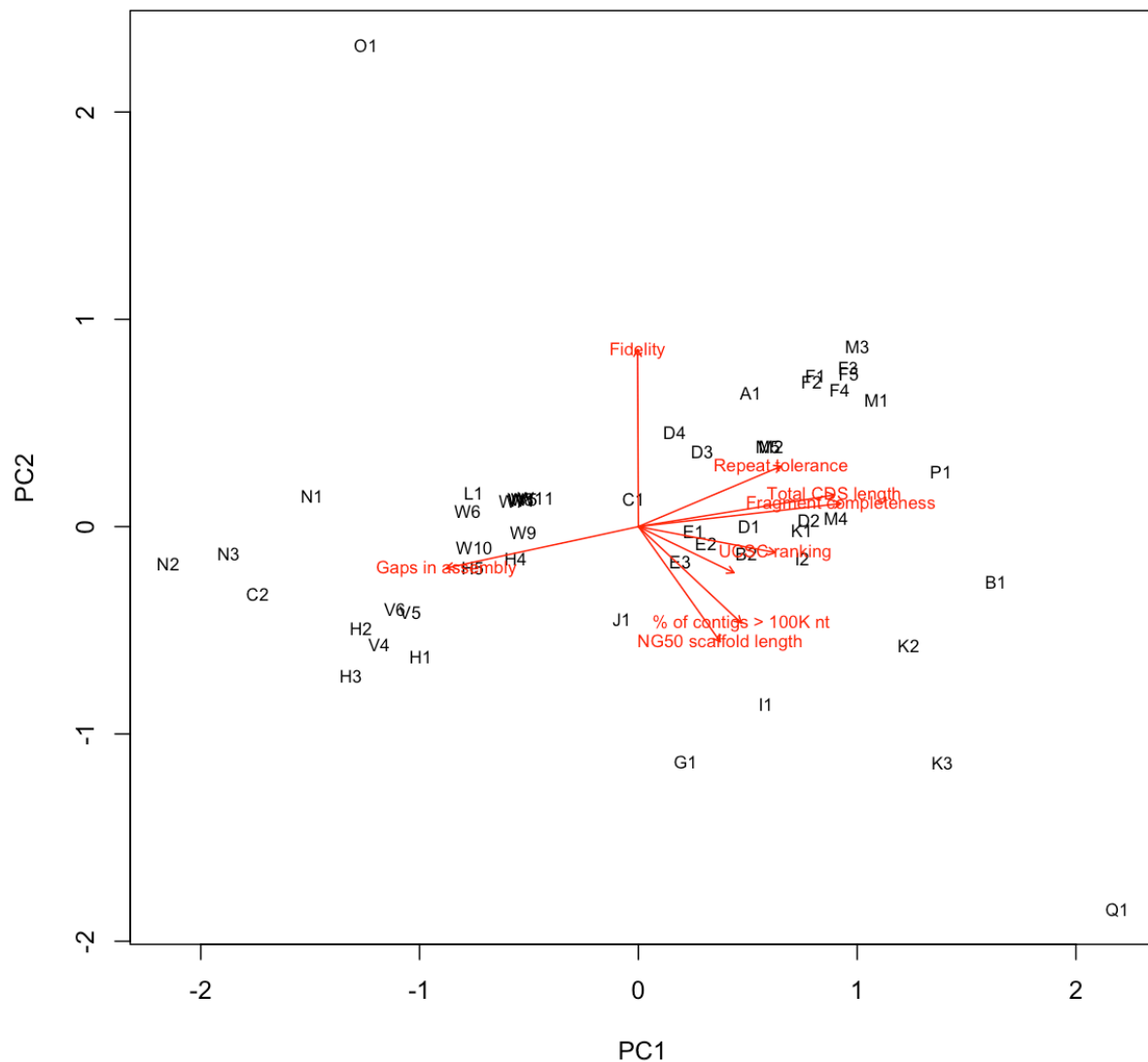


Figure 14. Principal Components Analysis plot for all assemblies, based on nine different metrics.

This analysis shows that a single principle component (PC1) can be used to order many of the assemblies. At one end of the axis defined by PC1 are assemblies which have longer scaffolds and contigs, which have more gene space represented, and which are more inclusive of repeat rich sequence. Assemblies at the other end of this axis are also defined by having more gaps (when compared to the reference). Overall, the position of the assemblies along this axis is in broad agreement with many of the metrics outlined earlier in this report, with assemblies by teams Q, P, B, M, and K ranking highly. A secondary trend (PC2) that accounts for a smaller amount of variance across the nine metrics, largely separates out assemblies on the basis of their 'Fidelity'. This metric – to

be explained in more detail elsewhere – measures the accuracy of those portions of each assembly which align to the known genome sequence. As seen in Figure 14, assembly O1 ranks very highly by this measure. It makes sense that this metric is largely independent of the other eight metrics. This is because an assembly can have very short contigs, be missing large amounts of gene space etc., but still have very accurate sequence.

Appendix 1 - basic assembly metrics

A1_scaffolds.fa.gz

Number of scaffolds	6229	
Total size of scaffolds	133356337	
Total scaffold length as percentage of known genome size	118.5%	
Longest scaffold	769386	
Shortest scaffold	450	
Number of scaffolds > 500 nt	6212	99.7%
Number of scaffolds > 1K nt	5924	95.1%
Number of scaffolds > 10K nt	3147	50.5%
Number of scaffolds > 100K nt	175	2.8%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	21409	
Median scaffold size	10170	
N50 scaffold length	48089	
L50 scaffold count	769	
NG50 scaffold length	57802	
LG50 scaffold count	570	
N50 scaffold - NG50 scaffold length difference	9713	
scaffold %A	29.57	
scaffold %C	20.26	
scaffold %G	20.28	
scaffold %T	29.63	
scaffold %N	0.26	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	65.5%	
Percentage of assembly in unscaffolded contigs	34.5%	
Average number of contigs per scaffold	1.6	
Average length of break (>25 Ns) between contigs in scaffold	56	
Number of contigs	9741	
Number of contigs in scaffolds	5464	
Number of contigs not in scaffolds	4277	
Total size of contigs	133005137	
Longest contig	494128	
Shortest contig	450	
Number of contigs > 500 nt	9721	99.8%
Number of contigs > 1K nt	9121	93.6%
Number of contigs > 10K nt	4217	43.3%
Number of contigs > 100K nt	40	0.4%
Number of contigs > 1M nt	0	0.0%
Mean contig size	13654	
Median contig size	8008	
N50 contig length	25383	
L50 contig count	1511	
NG50 contig length	29971	
LG50 contig count	1139	
N50 contig - NG50 contig length difference	4588	
contig %A	29.65	
contig %C	20.31	
contig %G	20.33	
contig %T	29.71	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

B1_scaffolds.fa.gz

Number of scaffolds	1449	
Total size of scaffolds	117920931	
Total scaffold length as percentage of known genome size	104.8%	
Longest scaffold	2032281	
Shortest scaffold	201	
Number of scaffolds > 500 nt	1228	84.7%
Number of scaffolds > 1K nt	711	49.1%
Number of scaffolds > 10K nt	375	25.9%
Number of scaffolds > 100K nt	308	21.3%
Number of scaffolds > 1M nt	8	0.6%
Mean scaffold size	81381	
Median scaffold size	980	
N50 scaffold length	466811	
L50 scaffold count	81	
NG50 scaffold length	502551	
LG50 scaffold count	75	
N50 scaffold - NG50 scaffold length difference	35740	
scaffold %A	29.70	
scaffold %C	20.20	
scaffold %G	20.21	
scaffold %T	29.64	
scaffold %N	0.26	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	97.4%	
Percentage of assembly in unscaffolded contigs	2.6%	
Average number of contigs per scaffold	3.1	
Average length of break (>25 Ns) between contigs in scaffold	210	
Number of contigs	4502	
Number of contigs in scaffolds	3398	
Number of contigs not in scaffolds	1104	
Total size of contigs	117615631	
Longest contig	347708	
Shortest contig	201	
Number of contigs > 500 nt	4281	95.1%
Number of contigs > 1K nt	3764	83.6%
Number of contigs > 10K nt	2270	50.4%
Number of contigs > 100K nt	240	5.3%
Number of contigs > 1M nt	0	0.0%
Mean contig size	26125	
Median contig size	10493	
N50 contig length	66967	
L50 contig count	548	
NG50 contig length	70244	
LG50 contig count	511	
N50 contig - NG50 contig length difference	3277	
contig %A	29.77	
contig %C	20.25	
contig %G	20.26	
contig %T	29.72	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

B2_scaffolds.fa.gz

Number of scaffolds	2048	
Total size of scaffolds	119840251	
Total scaffold length as percentage of known genome size	106.5%	
Longest scaffold	1922525	
Shortest scaffold	200	
Number of scaffolds > 500 nt	1261	61.6%
Number of scaffolds > 1K nt	823	40.2%
Number of scaffolds > 10K nt	440	21.5%
Number of scaffolds > 100K nt	314	15.3%
Number of scaffolds > 1M nt	10	0.5%
Mean scaffold size	58516	
Median scaffold size	731	
N50 scaffold length	436591	
L50 scaffold count	84	
NG50 scaffold length	469540	
LG50 scaffold count	76	
N50 scaffold - NG50 scaffold length difference	32949	
scaffold %A	29.47	
scaffold %C	20.03	
scaffold %G	20.02	
scaffold %T	29.43	
scaffold %N	1.05	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	86.6%	
Percentage of assembly in unscaffolded contigs	13.4%	
Average number of contigs per scaffold	1.5	
Average length of break (>25 Ns) between contigs in scaffold	613	
Number of contigs	3040	
Number of contigs in scaffolds	1308	
Number of contigs not in scaffolds	1732	
Total size of contigs	118583921	
Longest contig	1110065	
Shortest contig	15	
Number of contigs > 500 nt	2144	70.5%
Number of contigs > 1K nt	1665	54.8%
Number of contigs > 10K nt	1124	37.0%
Number of contigs > 100K nt	406	13.4%
Number of contigs > 1M nt	1	0.0%
Mean contig size	39008	
Median contig size	1456	
N50 contig length	153374	
L50 contig count	232	
NG50 contig length	160054	
LG50 contig count	213	
N50 contig - NG50 contig length difference	6680	
contig %A	29.78	
contig %C	20.24	
contig %G	20.23	
contig %T	29.74	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

C1_scaffolds.fa.gz

Number of scaffolds	5249	
Total size of scaffolds	114193167	
Total scaffold length as percentage of known genome size	101.5%	
Longest scaffold	2660971	
Shortest scaffold	105	
Number of scaffolds > 500 nt	871	16.6%
Number of scaffolds > 1K nt	278	5.3%
Number of scaffolds > 10K nt	198	3.8%
Number of scaffolds > 100K nt	171	3.3%
Number of scaffolds > 1M nt	36	0.7%
Mean scaffold size	21755	
Median scaffold size	172	
N50 scaffold length	958921	
L50 scaffold count	41	
NG50 scaffold length	959728	
LG50 scaffold count	40	
N50 scaffold - NG50 scaffold length difference	807	
scaffold %A	29.14	
scaffold %C	19.48	
scaffold %G	19.46	
scaffold %T	29.13	
scaffold %N	2.79	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	98.8%	
Percentage of assembly in unscaffolded contigs	1.2%	
Average number of contigs per scaffold	3.9	
Average length of break (>25 Ns) between contigs in scaffold	607	
Number of contigs	20229	
Number of contigs in scaffolds	15186	
Number of contigs not in scaffolds	5043	
Total size of contigs	111027307	
Longest contig	70936	
Shortest contig	105	
Number of contigs > 500 nt	15679	77.5%
Number of contigs > 1K nt	14338	70.9%
Number of contigs > 10K nt	3598	17.8%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	5489	
Median contig size	3318	
N50 contig length	10819	
L50 contig count	3169	
NG50 contig length	10700	
LG50 contig count	3237	
N50 contig - NG50 contig length difference	119	
contig %A	29.97	
contig %C	20.03	
contig %G	20.01	
contig %T	29.96	
contig %N	0.02	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

C2_scaffolds.fa.gz

Number of scaffolds	6114	
Total size of scaffolds	119411083	
Total scaffold length as percentage of known genome size	106.1%	
Longest scaffold	12125519	
Shortest scaffold	135	
Number of scaffolds > 500 nt	1022	16.7%
Number of scaffolds > 1K nt	226	3.7%
Number of scaffolds > 10K nt	76	1.2%
Number of scaffolds > 100K nt	67	1.1%
Number of scaffolds > 1M nt	36	0.6%
Mean scaffold size	19531	
Median scaffold size	309	
N50 scaffold length	2939915	
L50 scaffold count	14	
NG50 scaffold length	3032585	
LG50 scaffold count	13	
N50 scaffold - NG50 scaffold length difference	92670	
scaffold %A	26.24	
scaffold %C	17.50	
scaffold %G	17.52	
scaffold %T	26.27	
scaffold %N	12.46	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	98.4%	
Percentage of assembly in unscaffolded contigs	1.6%	
Average number of contigs per scaffold	8.7	
Average length of break (>25 Ns) between contigs in scaffold	2434	
Number of contigs	53273	
Number of contigs in scaffolds	47558	
Number of contigs not in scaffolds	5715	
Total size of contigs	104645412	
Longest contig	19554	
Shortest contig	126	
Number of contigs > 500 nt	42501	79.8%
Number of contigs > 1K nt	33710	63.3%
Number of contigs > 10K nt	184	0.3%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	1964	
Median contig size	1500	
N50 contig length	3003	
L50 contig count	11146	
NG50 contig length	2798	
LG50 contig count	12501	
N50 contig - NG50 contig length difference	205	
contig %A	29.95	
contig %C	19.97	
contig %G	20.00	
contig %T	29.97	
contig %N	0.11	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

D1_scaffolds.fa.gz

Number of scaffolds	1433	
Total size of scaffolds	119175323	
Total scaffold length as percentage of known genome size	105.9%	
Longest scaffold	10172515	
Shortest scaffold	500	
Number of scaffolds > 500 nt	1430	99.8%
Number of scaffolds > 1K nt	533	37.2%
Number of scaffolds > 10K nt	180	12.6%
Number of scaffolds > 100K nt	76	5.3%
Number of scaffolds > 1M nt	34	2.4%
Mean scaffold size	83165	
Median scaffold size	860	
N50 scaffold length	2710485	
L50 scaffold count	13	
NG50 scaffold length	2876145	
LG50 scaffold count	11	
N50 scaffold - NG50 scaffold length difference	165660	
scaffold %A	29.20	
scaffold %C	19.82	
scaffold %G	19.85	
scaffold %T	29.19	
scaffold %N	1.94	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	93.8%	
Percentage of assembly in unscaffolded contigs	6.2%	
Average number of contigs per scaffold	7.7	
Average length of break (>25 Ns) between contigs in scaffold	1615	
Number of contigs	11067	
Number of contigs in scaffolds	9716	
Number of contigs not in scaffolds	1351	
Total size of contigs	116860825	
Longest contig	207957	
Shortest contig	500	
Number of contigs > 500 nt	11064	100.0%
Number of contigs > 1K nt	10148	91.7%
Number of contigs > 10K nt	4147	37.5%
Number of contigs > 100K nt	14	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	10559	
Median contig size	7012	
N50 contig length	17882	
L50 contig count	1969	
NG50 contig length	18481	
LG50 contig count	1849	
N50 contig - NG50 contig length difference	599	
contig %A	29.78	
contig %C	20.21	
contig %G	20.24	
contig %T	29.77	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

D2_scaffolds.fa.gz

Number of scaffolds	1381	
Total size of scaffolds	118740053	
Total scaffold length as percentage of known genome size	105.5%	
Longest scaffold	10167372	
Shortest scaffold	500	
Number of scaffolds > 500 nt	1378	99.8%
Number of scaffolds > 1K nt	502	36.4%
Number of scaffolds > 10K nt	166	12.0%
Number of scaffolds > 100K nt	76	5.5%
Number of scaffolds > 1M nt	34	2.5%
Mean scaffold size	85981	
Median scaffold size	854	
N50 scaffold length	2709433	
L50 scaffold count	13	
NG50 scaffold length	2874844	
LG50 scaffold count	11	
N50 scaffold - NG50 scaffold length difference	165411	
scaffold %A	29.23	
scaffold %C	19.84	
scaffold %G	19.87	
scaffold %T	29.22	
scaffold %N	1.85	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	94.0%	
Percentage of assembly in unscaffolded contigs	6.0%	
Average number of contigs per scaffold	5.7	
Average length of break (>25 Ns) between contigs in scaffold	1586	
Number of contigs	7904	
Number of contigs in scaffolds	6600	
Number of contigs not in scaffolds	1304	
Total size of contigs	116548634	
Longest contig	207957	
Shortest contig	500	
Number of contigs > 500 nt	7901	100.0%
Number of contigs > 1K nt	7021	88.8%
Number of contigs > 10K nt	3858	48.8%
Number of contigs > 100K nt	26	0.3%
Number of contigs > 1M nt	0	0.0%
Mean contig size	14746	
Median contig size	9622	
N50 contig length	26304	
L50 contig count	1346	
NG50 contig length	27089	
LG50 contig count	1270	
N50 contig - NG50 contig length difference	785	
contig %A	29.78	
contig %C	20.21	
contig %G	20.24	
contig %T	29.77	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

D3_scaffolds.fa.gz

Number of scaffolds	2429	
Total size of scaffolds	122066780	
Total scaffold length as percentage of known genome size	108.5%	
Longest scaffold	2094112	
Shortest scaffold	500	
Number of scaffolds > 500 nt	2426	99.9%
Number of scaffolds > 1K nt	1529	62.9%
Number of scaffolds > 10K nt	836	34.4%
Number of scaffolds > 100K nt	290	11.9%
Number of scaffolds > 1M nt	7	0.3%
Mean scaffold size	50254	
Median scaffold size	3132	
N50 scaffold length	391554	
L50 scaffold count	96	
NG50 scaffold length	425643	
LG50 scaffold count	84	
N50 scaffold - NG50 scaffold length difference	34089	
scaffold %A	29.37	
scaffold %C	19.93	
scaffold %G	19.93	
scaffold %T	29.35	
scaffold %N	1.42	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	87.1%	
Percentage of assembly in unscaffolded contigs	12.9%	
Average number of contigs per scaffold	4.7	
Average length of break (>25 Ns) between contigs in scaffold	715	
Number of contigs	11310	
Number of contigs in scaffolds	9282	
Number of contigs not in scaffolds	2028	
Total size of contigs	120328105	
Longest contig	207957	
Shortest contig	500	
Number of contigs > 500 nt	11307	100.0%
Number of contigs > 1K nt	10394	91.9%
Number of contigs > 10K nt	4291	37.9%
Number of contigs > 100K nt	14	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	10639	
Median contig size	7078	
N50 contig length	18021	
L50 contig count	2026	
NG50 contig length	19085	
LG50 contig count	1814	
N50 contig - NG50 contig length difference	1064	
contig %A	29.79	
contig %C	20.21	
contig %G	20.22	
contig %T	29.77	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

D4_scaffolds.fa.gz

Number of scaffolds	23037	
Total size of scaffolds	118787544	
Total scaffold length as percentage of known genome size	105.6%	
Longest scaffold	207957	
Shortest scaffold	200	
Number of scaffolds > 500 nt	10850	47.1%
Number of scaffolds > 1K nt	9952	43.2%
Number of scaffolds > 10K nt	4119	17.9%
Number of scaffolds > 100K nt	14	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	5156	
Median scaffold size	386	
N50 scaffold length	17519	
L50 scaffold count	2030	
NG50 scaffold length	18380	
LG50 scaffold count	1855	
N50 scaffold - NG50 scaffold length difference	861	
scaffold %A	29.78	
scaffold %C	20.23	
scaffold %G	20.24	
scaffold %T	29.76	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in unscaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	23037	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	23037	
Total size of contigs	118787544	
Longest contig	207957	
Shortest contig	200	
Number of contigs > 500 nt	10850	47.1%
Number of contigs > 1K nt	9952	43.2%
Number of contigs > 10K nt	4119	17.9%
Number of contigs > 100K nt	14	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	5156	
Median contig size	386	
N50 contig length	17519	
L50 contig count	2030	
NG50 contig length	18380	
LG50 contig count	1855	
N50 contig - NG50 contig length difference	861	
contig %A	29.78	
contig %C	20.23	
contig %G	20.24	
contig %T	29.76	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

E1_scaffolds.fa.gz

Number of scaffolds	6166	
Total size of scaffolds	122138492	
Total scaffold length as percentage of known genome size	108.6%	
Longest scaffold	4822698	
Shortest scaffold	200	
Number of scaffolds > 500 nt	602	9.8%
Number of scaffolds > 1K nt	293	4.8%
Number of scaffolds > 10K nt	172	2.8%
Number of scaffolds > 100K nt	75	1.2%
Number of scaffolds > 1M nt	42	0.7%
Mean scaffold size	19808	
Median scaffold size	215	
N50 scaffold length	2708554	
L50 scaffold count	17	
NG50 scaffold length	2712723	
LG50 scaffold count	16	
N50 scaffold - NG50 scaffold length difference	4169	
scaffold %A	27.99	
scaffold %C	19.05	
scaffold %G	19.06	
scaffold %T	28.06	
scaffold %N	5.83	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	98.7%	
Percentage of assembly in unscaffolded contigs	1.3%	
Average number of contigs per scaffold	2.8	
Average length of break (>25 Ns) between contigs in scaffold	1155	
Number of contigs	17341	
Number of contigs in scaffolds	11411	
Number of contigs not in scaffolds	5930	
Total size of contigs	115021309	
Longest contig	251891	
Shortest contig	57	
Number of contigs > 500 nt	10976	63.3%
Number of contigs > 1K nt	10175	58.7%
Number of contigs > 10K nt	4109	23.7%
Number of contigs > 100K nt	17	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	6633	
Median contig size	2510	
N50 contig length	16897	
L50 contig count	2073	
NG50 contig length	17259	
LG50 contig count	1999	
N50 contig - NG50 contig length difference	362	
contig %A	29.72	
contig %C	20.23	
contig %G	20.24	
contig %T	29.80	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

E2_scaffolds.fa.gz

Number of scaffolds	82	
Total size of scaffolds	119530310	
Total scaffold length as percentage of known genome size	106.2%	
Longest scaffold	4822698	
Shortest scaffold	16258	
Number of scaffolds > 500 nt	82	100.0%
Number of scaffolds > 1K nt	82	100.0%
Number of scaffolds > 10K nt	82	100.0%
Number of scaffolds > 100K nt	75	91.5%
Number of scaffolds > 1M nt	42	51.2%
Mean scaffold size	1457687	
Median scaffold size	1054628	
N50 scaffold length	2708554	
L50 scaffold count	17	
NG50 scaffold length	2712723	
LG50 scaffold count	16	
N50 scaffold - NG50 scaffold length difference	4169	
scaffold %A	28.19	
scaffold %C	19.19	
scaffold %G	19.20	
scaffold %T	28.27	
scaffold %N	5.17	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	100.0%	
Percentage of assembly in unscaffolded contigs	0.0%	
Average number of contigs per scaffold	135.3	
Average length of break (>25 Ns) between contigs in scaffold	75310	
Number of contigs	11093	
Number of contigs in scaffolds	11093	
Number of contigs not in scaffolds	0	
Total size of contigs	113361086	
Longest contig	251891	
Shortest contig	57	
Number of contigs > 500 nt	10504	94.7%
Number of contigs > 1K nt	10079	90.9%
Number of contigs > 10K nt	4109	37.0%
Number of contigs > 100K nt	17	0.2%
Number of contigs > 1M nt	0	0.0%
Mean contig size	10219	
Median contig size	6919	
N50 contig length	17129	
L50 contig count	2024	
NG50 contig length	17259	
LG50 contig count	1999	
N50 contig - NG50 contig length difference	130	
contig %A	29.72	
contig %C	20.23	
contig %G	20.24	
contig %T	29.80	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

E3_scaffolds.fa.gz

Number of scaffolds	1937	
Total size of scaffolds	113447607	
Total scaffold length as percentage of known genome size	100.8%	
Longest scaffold	4689666	
Shortest scaffold	200	
Number of scaffolds > 500 nt	256	13.2%
Number of scaffolds > 1K nt	146	7.5%
Number of scaffolds > 10K nt	73	3.8%
Number of scaffolds > 100K nt	73	3.8%
Number of scaffolds > 1M nt	41	2.1%
Mean scaffold size	58569	
Median scaffold size	217	
N50 scaffold length	2575286	
L50 scaffold count	17	
NG50 scaffold length	2575286	
LG50 scaffold count	17	
N50 scaffold - NG50 scaffold length difference	0	
scaffold %A	29.74	
scaffold %C	20.24	
scaffold %G	20.23	
scaffold %T	29.79	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in unscaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	1937	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	1937	
Total size of contigs	113447607	
Longest contig	4689666	
Shortest contig	200	
Number of contigs > 500 nt	256	13.2%
Number of contigs > 1K nt	146	7.5%
Number of contigs > 10K nt	73	3.8%
Number of contigs > 100K nt	73	3.8%
Number of contigs > 1M nt	41	2.1%
Mean contig size	58569	
Median contig size	217	
N50 contig length	2575286	
L50 contig count	17	
NG50 contig length	2575286	
LG50 contig count	17	
N50 contig - NG50 contig length difference	0	
contig %A	29.74	
contig %C	20.24	
contig %G	20.23	
contig %T	29.79	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

F1_scaffolds.fa.gz

Number of scaffolds	44949	
Total size of scaffolds	129788571	
Total scaffold length as percentage of known genome size	115.4%	
Longest scaffold	721724	
Shortest scaffold	67	
Number of scaffolds > 500 nt	5369	11.9%
Number of scaffolds > 1K nt	5064	11.3%
Number of scaffolds > 10K nt	3501	7.8%
Number of scaffolds > 100K nt	73	0.2%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	2887	
Median scaffold size	131	
N50 scaffold length	37403	
L50 scaffold count	1046	
NG50 scaffold length	42952	
LG50 scaffold count	831	
N50 scaffold - NG50 scaffold length difference	5549	
scaffold %A	29.76	
scaffold %C	20.22	
scaffold %G	20.21	
scaffold %T	29.78	
scaffold %N	0.03	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	2663	
Percentage of assembly in scaffolded contigs	15.1%	
Percentage of assembly in unscaffolded contigs	84.9%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	45487	
Number of contigs in scaffolds	1035	
Number of contigs not in scaffolds	44452	
Total size of contigs	129765436	
Longest contig	721724	
Shortest contig	67	
Number of contigs > 500 nt	5813	12.8%
Number of contigs > 1K nt	5492	12.1%
Number of contigs > 10K nt	3662	8.1%
Number of contigs > 100K nt	58	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	2853	
Median contig size	131	
N50 contig length	34247	
L50 contig count	1138	
NG50 contig length	39308	
LG50 contig count	903	
N50 contig - NG50 contig length difference	5061	
contig %A	29.76	
contig %C	20.23	
contig %G	20.21	
contig %T	29.79	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	2663	

F2_scaffolds.fa.gz

Number of scaffolds	33063	
Total size of scaffolds	128848474	
Total scaffold length as percentage of known genome size	114.5%	
Longest scaffold	721724	
Shortest scaffold	67	
Number of scaffolds > 500 nt	5369	16.2%
Number of scaffolds > 1K nt	5064	15.3%
Number of scaffolds > 10K nt	3501	10.6%
Number of scaffolds > 100K nt	73	0.2%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	3897	
Median scaffold size	133	
N50 scaffold length	37733	
L50 scaffold count	1034	
NG50 scaffold length	42952	
LG50 scaffold count	831	
N50 scaffold - NG50 scaffold length difference	5219	
scaffold %A	29.76	
scaffold %C	20.23	
scaffold %G	20.21	
scaffold %T	29.78	
scaffold %N	0.02	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	2497	
Percentage of assembly in scaffolded contigs	10.6%	
Percentage of assembly in unscaffolded contigs	89.4%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	33437	
Number of contigs in scaffolds	726	
Number of contigs not in scaffolds	32711	
Total size of contigs	128832348	
Longest contig	721724	
Shortest contig	67	
Number of contigs > 500 nt	5662	16.9%
Number of contigs > 1K nt	5345	16.0%
Number of contigs > 10K nt	3613	10.8%
Number of contigs > 100K nt	63	0.2%
Number of contigs > 1M nt	0	0.0%
Mean contig size	3853	
Median contig size	133	
N50 contig length	35510	
L50 contig count	1094	
NG50 contig length	40359	
LG50 contig count	879	
N50 contig - NG50 contig length difference	4849	
contig %A	29.76	
contig %C	20.23	
contig %G	20.21	
contig %T	29.78	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	2497	

F3_scaffolds.fa.gz

Number of scaffolds	28678	
Total size of scaffolds	13100911	
Total scaffold length as percentage of known genome size	116.4%	
Longest scaffold	721864	
Shortest scaffold	67	
Number of scaffolds > 500 nt	3825	13.3%
Number of scaffolds > 1K nt	3597	12.5%
Number of scaffolds > 10K nt	2796	9.7%
Number of scaffolds > 100K nt	188	0.7%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	4568	
Median scaffold size	133	
N50 scaffold length	54764	
L50 scaffold count	741	
NG50 scaffold length	62065	
LG50 scaffold count	582	
N50 scaffold - NG50 scaffold length difference	7301	
scaffold %A	29.77	
scaffold %C	20.21	
scaffold %G	20.20	
scaffold %T	29.77	
scaffold %N	0.04	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	3969	
Percentage of assembly in scaffolded contigs	23.5%	
Percentage of assembly in unscaffolded contigs	76.5%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	1	
Number of contigs	29300	
Number of contigs in scaffolds	1168	
Number of contigs not in scaffolds	28132	
Total size of contigs	130964214	
Longest contig	721864	
Shortest contig	67	
Number of contigs > 500 nt	4348	14.8%
Number of contigs > 1K nt	4102	14.0%
Number of contigs > 10K nt	3072	10.5%
Number of contigs > 100K nt	140	0.5%
Number of contigs > 1M nt	0	0.0%
Mean contig size	4470	
Median contig size	133	
N50 contig length	48200	
L50 contig count	843	
NG50 contig length	54850	
LG50 contig count	664	
N50 contig - NG50 contig length difference	6650	
contig %A	29.78	
contig %C	20.22	
contig %G	20.21	
contig %T	29.78	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	3969	

F4_scaffolds.fa.gz

Number of scaffolds	28681	
Total size of scaffolds	130232384	
Total scaffold length as percentage of known genome size	115.8%	
Longest scaffold	2820201	
Shortest scaffold	67	
Number of scaffolds > 500 nt	1052	3.7%
Number of scaffolds > 1K nt	748	2.6%
Number of scaffolds > 10K nt	480	1.7%
Number of scaffolds > 100K nt	222	0.8%
Number of scaffolds > 1M nt	28	0.1%
Mean scaffold size	4541	
Median scaffold size	133	
N50 scaffold length	721724	
L50 scaffold count	55	
NG50 scaffold length	825511	
LG50 scaffold count	44	
N50 scaffold - NG50 scaffold length difference	103787	
scaffold %A	29.46	
scaffold %C	20.00	
scaffold %G	20.01	
scaffold %T	29.45	
scaffold %N	1.08	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	2465	
Percentage of assembly in scaffolded contigs	89.4%	
Percentage of assembly in unscaffolded contigs	10.6%	
Average number of contigs per scaffold	1.1	
Average length of break (>25 Ns) between contigs in scaffold	49	
Number of contigs	32134	
Number of contigs in scaffolds	3728	
Number of contigs not in scaffolds	28406	
Total size of contigs	128846436	
Longest contig	721724	
Shortest contig	67	
Number of contigs > 500 nt	4451	13.9%
Number of contigs > 1K nt	4139	12.9%
Number of contigs > 10K nt	3030	9.4%
Number of contigs > 100K nt	135	0.4%
Number of contigs > 1M nt	0	0.0%
Mean contig size	4010	
Median contig size	133	
N50 contig length	47737	
L50 contig count	817	
NG50 contig length	53825	
LG50 contig count	656	
N50 contig - NG50 contig length difference	6088	
contig %A	29.78	
contig %C	20.22	
contig %G	20.22	
contig %T	29.77	
contig %N	0.02	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	2465	

F5_scaffolds.fa.gz

Number of scaffolds	25812	
Total size of scaffolds	132224524	
Total scaffold length as percentage of known genome size	117.5%	
Longest scaffold	4232700	
Shortest scaffold	67	
Number of scaffolds > 500 nt	916	3.5%
Number of scaffolds > 1K nt	689	2.7%
Number of scaffolds > 10K nt	513	2.0%
Number of scaffolds > 100K nt	275	1.1%
Number of scaffolds > 1M nt	19	0.1%
Mean scaffold size	5123	
Median scaffold size	133	
N50 scaffold length	551979	
L50 scaffold count	68	
NG50 scaffold length	677662	
LG50 scaffold count	52	
N50 scaffold - NG50 scaffold length difference	125683	
scaffold %A	29.53	
scaffold %C	20.03	
scaffold %G	20.02	
scaffold %T	29.48	
scaffold %N	0.95	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	3952	
Percentage of assembly in scaffolded contigs	91.2%	
Percentage of assembly in unscaffolded contigs	8.8%	
Average number of contigs per scaffold	1.1	
Average length of break (>25 Ns) between contigs in scaffold	48	
Number of contigs	28683	
Number of contigs in scaffolds	3201	
Number of contigs not in scaffolds	25482	
Total size of contigs	130989538	
Longest contig	721871	
Shortest contig	67	
Number of contigs > 500 nt	3714	12.9%
Number of contigs > 1K nt	3473	12.1%
Number of contigs > 10K nt	2726	9.5%
Number of contigs > 100K nt	190	0.7%
Number of contigs > 1M nt	0	0.0%
Mean contig size	4567	
Median contig size	133	
N50 contig length	56660	
L50 contig count	717	
NG50 contig length	65108	
LG50 contig count	564	
N50 contig - NG50 contig length difference	8448	
contig %A	29.80	
contig %C	20.22	
contig %G	20.21	
contig %T	29.76	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	3952	

G1_scaffolds.fa.gz

Number of scaffolds	966	
Total size of scaffolds	117479181	
Total scaffold length as percentage of known genome size	104.4%	
Longest scaffold	23079946	
Shortest scaffold	150	
Number of scaffolds > 500 nt	364	37.7%
Number of scaffolds > 1K nt	183	18.9%
Number of scaffolds > 10K nt	41	4.2%
Number of scaffolds > 100K nt	34	3.5%
Number of scaffolds > 1M nt	21	2.2%
Mean scaffold size	121614	
Median scaffold size	333	
N50 scaffold length	9073174	
L50 scaffold count	5	
NG50 scaffold length	9073174	
LG50 scaffold count	5	
N50 scaffold - NG50 scaffold length difference	0	
scaffold %A	29.10	
scaffold %C	19.75	
scaffold %G	19.74	
scaffold %T	29.03	
scaffold %N	2.38	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	99.6%	
Percentage of assembly in unscaffolded contigs	0.4%	
Average number of contigs per scaffold	15.3	
Average length of break (>25 Ns) between contigs in scaffold	2897	
Number of contigs	14817	
Number of contigs in scaffolds	14090	
Number of contigs not in scaffolds	727	
Total size of contigs	114744669	
Longest contig	722469	
Shortest contig	100	
Number of contigs > 500 nt	12154	82.0%
Number of contigs > 1K nt	11111	75.0%
Number of contigs > 10K nt	4019	27.1%
Number of contigs > 100K nt	18	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	7744	
Median contig size	4452	
N50 contig length	15585	
L50 contig count	2176	
NG50 contig length	15913	
LG50 contig count	2105	
N50 contig - NG50 contig length difference	328	
contig %A	29.79	
contig %C	20.22	
contig %G	20.21	
contig %T	29.72	
contig %N	0.06	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

H1_scaffolds.fa.gz

Number of scaffolds	3924	
Total size of scaffolds	124075451	
Total scaffold length as percentage of known genome size	110.3%	
Longest scaffold	4989023	
Shortest scaffold	100	
Number of scaffolds > 500 nt	2846	72.5%
Number of scaffolds > 1K nt	1968	50.2%
Number of scaffolds > 10K nt	158	4.0%
Number of scaffolds > 100K nt	128	3.3%
Number of scaffolds > 1M nt	47	1.2%
Mean scaffold size	31620	
Median scaffold size	1001	
N50 scaffold length	1271958	
L50 scaffold count	30	
NG50 scaffold length	1406792	
LG50 scaffold count	26	
N50 scaffold - NG50 scaffold length difference	134834	
scaffold %A	26.73	
scaffold %C	18.21	
scaffold %G	18.23	
scaffold %T	26.66	
scaffold %N	10.17	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	97.5%	
Percentage of assembly in unscaffolded contigs	2.5%	
Average number of contigs per scaffold	5.0	
Average length of break (>25 Ns) between contigs in scaffold	3214	
Number of contigs	19499	
Number of contigs in scaffolds	16117	
Number of contigs not in scaffolds	3382	
Total size of contigs	111473322	
Longest contig	722170	
Shortest contig	0	
Number of contigs > 500 nt	16786	86.1%
Number of contigs > 1K nt	15582	79.9%
Number of contigs > 10K nt	3440	17.6%
Number of contigs > 100K nt	16	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	5717	
Median contig size	3208	
N50 contig length	11024	
L50 contig count	2911	
NG50 contig length	10915	
LG50 contig count	2958	
N50 contig - NG50 contig length difference	109	
contig %A	29.75	
contig %C	20.27	
contig %G	20.29	
contig %T	29.68	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

H2_scaffolds.fa.gz

Number of scaffolds	22181	
Total size of scaffolds	110280191	
Total scaffold length as percentage of known genome size	98.0%	
Longest scaffold	541160	
Shortest scaffold	100	
Number of scaffolds > 500 nt	21099	95.1%
Number of scaffolds > 1K nt	20123	90.7%
Number of scaffolds > 10K nt	2486	11.2%
Number of scaffolds > 100K nt	19	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	4972	
Median scaffold size	3417	
N50 scaffold length	7698	
L50 scaffold count	4248	
NG50 scaffold length	7538	
LG50 scaffold count	4394	
N50 scaffold - NG50 scaffold length difference	160	
scaffold %A	29.71	
scaffold %C	20.21	
scaffold %G	20.24	
scaffold %T	29.56	
scaffold %N	0.28	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	22.2%	
Percentage of assembly in unscaffolded contigs	77.8%	
Average number of contigs per scaffold	1.2	
Average length of break (>25 Ns) between contigs in scaffold	13	
Number of contigs	26333	
Number of contigs in scaffolds	7735	
Number of contigs not in scaffolds	18598	
Total size of contigs	109987924	
Longest contig	541160	
Shortest contig	0	
Number of contigs > 500 nt	22939	87.1%
Number of contigs > 1K nt	21448	81.4%
Number of contigs > 10K nt	2189	8.3%
Number of contigs > 100K nt	19	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	4177	
Median contig size	2774	
N50 contig length	6980	
L50 contig count	4624	
NG50 contig length	6828	
LG50 contig count	4806	
N50 contig - NG50 contig length difference	152	
contig %A	29.79	
contig %C	20.26	
contig %G	20.30	
contig %T	29.64	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

H3_scaffolds.fa.gz

Number of scaffolds	3924	
Total size of scaffolds	125050796	
Total scaffold length as percentage of known genome size	111.2%	
Longest scaffold	5037481	
Shortest scaffold	100	
Number of scaffolds > 500 nt	2846	72.5%
Number of scaffolds > 1K nt	1968	50.2%
Number of scaffolds > 10K nt	158	4.0%
Number of scaffolds > 100K nt	128	3.3%
Number of scaffolds > 1M nt	47	1.2%
Mean scaffold size	31868	
Median scaffold size	1001	
N50 scaffold length	1283620	
L50 scaffold count	30	
NG50 scaffold length	1421787	
LG50 scaffold count	26	
N50 scaffold - NG50 scaffold length difference	138167	
scaffold %A	26.17	
scaffold %C	17.82	
scaffold %G	17.85	
scaffold %T	26.10	
scaffold %N	12.06	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	97.6%	
Percentage of assembly in unscaffolded contigs	2.4%	
Average number of contigs per scaffold	6.6	
Average length of break (>25 Ns) between contigs in scaffold	3842	
Number of contigs	25762	
Number of contigs in scaffolds	22388	
Number of contigs not in scaffolds	3374	
Total size of contigs	109987145	
Longest contig	541160	
Shortest contig	0	
Number of contigs > 500 nt	22679	88.0%
Number of contigs > 1K nt	21204	82.3%
Number of contigs > 10K nt	2217	8.6%
Number of contigs > 100K nt	20	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	4269	
Median contig size	2880	
N50 contig length	7057	
L50 contig count	4586	
NG50 contig length	6900	
LG50 contig count	4766	
N50 contig - NG50 contig length difference	157	
contig %A	29.75	
contig %C	20.27	
contig %G	20.29	
contig %T	29.68	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

H4_scaffolds.fa.gz

Number of scaffolds	8517	
Total size of scaffolds	114286718	
Total scaffold length as percentage of known genome size	101.6%	
Longest scaffold	726869	
Shortest scaffold	100	
Number of scaffolds > 500 nt	8333	97.8%
Number of scaffolds > 1K nt	8189	96.1%
Number of scaffolds > 10K nt	3587	42.1%
Number of scaffolds > 100K nt	22	0.3%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	13419	
Median scaffold size	7812	
N50 scaffold length	24903	
L50 scaffold count	1329	
NG50 scaffold length	25191	
LG50 scaffold count	1293	
N50 scaffold - NG50 scaffold length difference	288	
scaffold %A	29.68	
scaffold %C	20.22	
scaffold %G	20.23	
scaffold %T	29.60	
scaffold %N	0.26	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	53.1%	
Percentage of assembly in unscaffolded contigs	46.9%	
Average number of contigs per scaffold	1.5	
Average length of break (>25 Ns) between contigs in scaffold	34	
Number of contigs	12516	
Number of contigs in scaffolds	6668	
Number of contigs not in scaffolds	5848	
Total size of contigs	114004040	
Longest contig	587037	
Shortest contig	0	
Number of contigs > 500 nt	11419	91.2%
Number of contigs > 1K nt	10975	87.7%
Number of contigs > 10K nt	3784	30.2%
Number of contigs > 100K nt	15	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	9109	
Median contig size	5346	
N50 contig length	17117	
L50 contig count	1905	
NG50 contig length	17411	
LG50 contig count	1862	
N50 contig - NG50 contig length difference	294	
contig %A	29.76	
contig %C	20.27	
contig %G	20.28	
contig %T	29.68	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

H5_scaffolds.fa.gz

Number of scaffolds	12904	
Total size of scaffolds	112341484	
Total scaffold length as percentage of known genome size	99.9%	
Longest scaffold	1143876	
Shortest scaffold	100	
Number of scaffolds > 500 nt	12467	96.6%
Number of scaffolds > 1K nt	12230	94.8%
Number of scaffolds > 10K nt	3850	29.8%
Number of scaffolds > 100K nt	14	0.1%
Number of scaffolds > 1M nt	1	0.0%
Mean scaffold size	8706	
Median scaffold size	5526	
N50 scaffold length	14999	
L50 scaffold count	2248	
NG50 scaffold length	14964	
LG50 scaffold count	2254	
N50 scaffold - NG50 scaffold length difference	35	
scaffold %A	29.68	
scaffold %C	20.23	
scaffold %G	20.23	
scaffold %T	29.62	
scaffold %N	0.24	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	35.2%	
Percentage of assembly in unscaffolded contigs	64.8%	
Average number of contigs per scaffold	1.3	
Average length of break (>25 Ns) between contigs in scaffold	21	
Number of contigs	16402	
Number of contigs in scaffolds	6313	
Number of contigs not in scaffolds	10089	
Total size of contigs	112082306	
Longest contig	831935	
Shortest contig	0	
Number of contigs > 500 nt	14943	91.1%
Number of contigs > 1K nt	14348	87.5%
Number of contigs > 10K nt	3571	21.8%
Number of contigs > 100K nt	15	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	6833	
Median contig size	4210	
N50 contig length	12130	
L50 contig count	2696	
NG50 contig length	12097	
LG50 contig count	2713	
N50 contig - NG50 contig length difference	33	
contig %A	29.75	
contig %C	20.28	
contig %G	20.27	
contig %T	29.69	
contig %N	0.01	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

I1_scaffolds.fa.gz

Number of scaffolds	128	
Total size of scaffolds	104674317	
Total scaffold length as percentage of known genome size	93.0%	
Longest scaffold	10924052	
Shortest scaffold	77	
Number of scaffolds > 500 nt	73	57.0%
Number of scaffolds > 1K nt	64	50.0%
Number of scaffolds > 10K nt	56	43.8%
Number of scaffolds > 100K nt	54	42.2%
Number of scaffolds > 1M nt	31	24.2%
Mean scaffold size	817768	
Median scaffold size	1097	
N50 scaffold length	3293136	
L50 scaffold count	9	
NG50 scaffold length	3254796	
LG50 scaffold count	10	
N50 scaffold - NG50 scaffold length difference	38340	
scaffold %A	29.50	
scaffold %C	20.11	
scaffold %G	20.10	
scaffold %T	29.51	
scaffold %N	0.78	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	100.0%	
Percentage of assembly in unscaffolded contigs	0.0%	
Average number of contigs per scaffold	14.0	
Average length of break (>25 Ns) between contigs in scaffold	6416	
Number of contigs	1798	
Number of contigs in scaffolds	1726	
Number of contigs not in scaffolds	72	
Total size of contigs	103898208	
Longest contig	1442666	
Shortest contig	64	
Number of contigs > 500 nt	1457	81.0%
Number of contigs > 1K nt	1351	75.1%
Number of contigs > 10K nt	1071	59.6%
Number of contigs > 100K nt	368	20.5%
Number of contigs > 1M nt	1	0.1%
Mean contig size	57785	
Median contig size	20278	
N50 contig length	151121	
L50 contig count	200	
NG50 contig length	135806	
LG50 contig count	230	
N50 contig - NG50 contig length difference	15315	
contig %A	29.72	
contig %C	20.26	
contig %G	20.25	
contig %T	29.73	
contig %N	0.04	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

I2_scaffolds.fa.gz

Number of scaffolds	35661	
Total size of scaffolds	121259411	
Total scaffold length as percentage of known genome size	107.8%	
Longest scaffold	8283751	
Shortest scaffold	64	
Number of scaffolds > 500 nt	548	1.5%
Number of scaffolds > 1K nt	77	0.2%
Number of scaffolds > 10K nt	65	0.2%
Number of scaffolds > 100K nt	63	0.2%
Number of scaffolds > 1M nt	37	0.1%
Mean scaffold size	3400	
Median scaffold size	89	
N50 scaffold length	2461784	
L50 scaffold count	12	
NG50 scaffold length	3254796	
LG50 scaffold count	11	
N50 scaffold - NG50 scaffold length difference	793012	
scaffold %A	29.55	
scaffold %C	20.11	
scaffold %G	20.11	
scaffold %T	29.56	
scaffold %N	0.67	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	96.6%	
Percentage of assembly in unscaffolded contigs	3.4%	
Average number of contigs per scaffold	1.1	
Average length of break (>25 Ns) between contigs in scaffold	22	
Number of contigs	37571	
Number of contigs in scaffolds	1975	
Number of contigs not in scaffolds	35596	
Total size of contigs	120501333	
Longest contig	1442666	
Shortest contig	64	
Number of contigs > 500 nt	2137	5.7%
Number of contigs > 1K nt	1552	4.1%
Number of contigs > 10K nt	1223	3.3%
Number of contigs > 100K nt	413	1.1%
Number of contigs > 1M nt	1	0.0%
Mean contig size	3207	
Median contig size	90	
N50 contig length	139666	
L50 contig count	245	
NG50 contig length	150434	
LG50 contig count	218	
N50 contig - NG50 contig length difference	10768	
contig %A	29.74	
contig %C	20.24	
contig %G	20.23	
contig %T	29.75	
contig %N	0.04	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

J1_scaffolds.fa.gz

Number of scaffolds	4791	
Total size of scaffolds	113043196	
Total scaffold length as percentage of known genome size	100.5%	
Longest scaffold	1474238	
Shortest scaffold	200	
Number of scaffolds > 500 nt	2505	52.3%
Number of scaffolds > 1K nt	1963	41.0%
Number of scaffolds > 10K nt	537	11.2%
Number of scaffolds > 100K nt	336	7.0%
Number of scaffolds > 1M nt	2	0.0%
Mean scaffold size	23595	
Median scaffold size	556	
N50 scaffold length	301691	
L50 scaffold count	111	
NG50 scaffold length	301691	
LG50 scaffold count	111	
N50 scaffold - NG50 scaffold length difference	0	
scaffold %A	29.75	
scaffold %C	20.22	
scaffold %G	20.21	
scaffold %T	29.69	
scaffold %N	0.14	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in unscaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	32	
Number of contigs	4791	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	4791	
Total size of contigs	113043196	
Longest contig	1474238	
Shortest contig	200	
Number of contigs > 500 nt	2505	52.3%
Number of contigs > 1K nt	1963	41.0%
Number of contigs > 10K nt	537	11.2%
Number of contigs > 100K nt	336	7.0%
Number of contigs > 1M nt	2	0.0%
Mean contig size	23595	
Median contig size	556	
N50 contig length	301691	
L50 contig count	111	
NG50 contig length	301691	
LG50 contig count	111	
N50 contig - NG50 contig length difference	0	
contig %A	29.75	
contig %C	20.22	
contig %G	20.21	
contig %T	29.69	
contig %N	0.14	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

K1_scaffolds.fa.gz

Number of scaffolds	14856	
Total size of scaffolds	115872926	
Total scaffold length as percentage of known genome size	103.0%	
Longest scaffold	5949022	
Shortest scaffold	100	
Number of scaffolds > 500 nt	766	5.2%
Number of scaffolds > 1K nt	496	3.3%
Number of scaffolds > 10K nt	120	0.8%
Number of scaffolds > 100K nt	90	0.6%
Number of scaffolds > 1M nt	38	0.3%
Mean scaffold size	7800	
Median scaffold size	129	
N50 scaffold length	1789861	
L50 scaffold count	18	
NG50 scaffold length	1801023	
LG50 scaffold count	17	
N50 scaffold - NG50 scaffold length difference	11162	
scaffold %A	29.68	
scaffold %C	19.88	
scaffold %G	19.87	
scaffold %T	29.76	
scaffold %N	0.82	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	96.8%	
Percentage of assembly in unscaffolded contigs	3.2%	
Average number of contigs per scaffold	1.1	
Average length of break (>25 Ns) between contigs in scaffold	63	
Number of contigs	15689	
Number of contigs in scaffolds	960	
Number of contigs not in scaffolds	14729	
Total size of contigs	114925837	
Longest contig	788375	
Shortest contig	100	
Number of contigs > 500 nt	1562	10.0%
Number of contigs > 1K nt	1263	8.1%
Number of contigs > 10K nt	797	5.1%
Number of contigs > 100K nt	407	2.6%
Number of contigs > 1M nt	0	0.0%
Mean contig size	7325	
Median contig size	131	
N50 contig length	209662	
L50 contig count	169	
NG50 contig length	214562	
LG50 contig count	163	
N50 contig - NG50 contig length difference	4900	
contig %A	29.92	
contig %C	20.04	
contig %G	20.04	
contig %T	30.00	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

K2_scaffolds.fa.gz

Number of scaffolds	2043	
Total size of scaffolds	113224906	
Total scaffold length as percentage of known genome size	100.6%	
Longest scaffold	5949022	
Shortest scaffold	100	
Number of scaffolds > 500 nt	338	16.5%
Number of scaffolds > 1K nt	244	11.9%
Number of scaffolds > 10K nt	103	5.0%
Number of scaffolds > 100K nt	90	4.4%
Number of scaffolds > 1M nt	38	1.9%
Mean scaffold size	55421	
Median scaffold size	132	
N50 scaffold length	1801023	
L50 scaffold count	17	
NG50 scaffold length	1801023	
LG50 scaffold count	17	
N50 scaffold - NG50 scaffold length difference	0	
scaffold %A	29.73	
scaffold %C	19.89	
scaffold %G	19.88	
scaffold %T	29.81	
scaffold %N	0.69	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	98.7%	
Percentage of assembly in unscaffolded contigs	1.3%	
Average number of contigs per scaffold	1.4	
Average length of break (>25 Ns) between contigs in scaffold	380	
Number of contigs	2796	
Number of contigs in scaffolds	848	
Number of contigs not in scaffolds	1948	
Total size of contigs	112447027	
Longest contig	788375	
Shortest contig	100	
Number of contigs > 500 nt	1089	38.9%
Number of contigs > 1K nt	986	35.3%
Number of contigs > 10K nt	791	28.3%
Number of contigs > 100K nt	407	14.6%
Number of contigs > 1M nt	0	0.0%
Mean contig size	40217	
Median contig size	159	
N50 contig length	214562	
L50 contig count	163	
NG50 contig length	214562	
LG50 contig count	163	
N50 contig - NG50 contig length difference	0	
contig %A	29.93	
contig %C	20.03	
contig %G	20.02	
contig %T	30.02	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

K3_scaffolds.fa.gz

Number of scaffolds	173	
Total size of scaffolds	112588332	
Total scaffold length as percentage of known genome size	100.1%	
Longest scaffold	5949022	
Shortest scaffold	100	
Number of scaffolds > 500 nt	135	78.0%
Number of scaffolds > 1K nt	126	72.8%
Number of scaffolds > 10K nt	100	57.8%
Number of scaffolds > 100K nt	90	52.0%
Number of scaffolds > 1M nt	38	22.0%
Mean scaffold size	650800	
Median scaffold size	131182	
N50 scaffold length	1801023	
L50 scaffold count	17	
NG50 scaffold length	1801023	
LG50 scaffold count	17	
N50 scaffold - NG50 scaffold length difference	0	
scaffold %A	29.73	
scaffold %C	19.88	
scaffold %G	19.87	
scaffold %T	29.82	
scaffold %N	0.69	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	99.2%	
Percentage of assembly in unscaffolded contigs	0.8%	
Average number of contigs per scaffold	5.4	
Average length of break (>25 Ns) between contigs in scaffold	4498	
Number of contigs	926	
Number of contigs in scaffolds	848	
Number of contigs not in scaffolds	78	
Total size of contigs	111810453	
Longest contig	788375	
Shortest contig	100	
Number of contigs > 500 nt	886	95.7%
Number of contigs > 1K nt	868	93.7%
Number of contigs > 10K nt	788	85.1%
Number of contigs > 100K nt	407	44.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	120746	
Median contig size	81841	
N50 contig length	216393	
L50 contig count	162	
NG50 contig length	214562	
LG50 contig count	163	
N50 contig - NG50 contig length difference	1831	
contig %A	29.94	
contig %C	20.02	
contig %G	20.01	
contig %T	30.03	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

L1_scaffolds.fa.gz

Number of scaffolds	14822	
Total size of scaffolds	124310199	
Total scaffold length as percentage of known genome size	110.5%	
Longest scaffold	128347	
Shortest scaffold	404	
Number of scaffolds > 500 nt	14820	100.0%
Number of scaffolds > 1K nt	14799	99.8%
Number of scaffolds > 10K nt	3457	23.3%
Number of scaffolds > 100K nt	11	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	8387	
Median scaffold size	3810	
N50 scaffold length	20165	
L50 scaffold count	1762	
NG50 scaffold length	22716	
LG50 scaffold count	1487	
N50 scaffold - NG50 scaffold length difference	2551	
scaffold %A	29.87	
scaffold %C	20.10	
scaffold %G	20.12	
scaffold %T	29.91	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in unscaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	14822	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	14822	
Total size of contigs	124310199	
Longest contig	128347	
Shortest contig	404	
Number of contigs > 500 nt	14820	100.0%
Number of contigs > 1K nt	14799	99.8%
Number of contigs > 10K nt	3457	23.3%
Number of contigs > 100K nt	11	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	8387	
Median contig size	3810	
N50 contig length	20165	
L50 contig count	1762	
NG50 contig length	22716	
LG50 contig count	1487	
N50 contig - NG50 contig length difference	2551	
contig %A	29.87	
contig %C	20.10	
contig %G	20.12	
contig %T	29.91	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

M1_scaffolds.fa.gz

Number of scaffolds	657	
Total size of scaffolds	154090743	
Total scaffold length as percentage of known genome size	137.0%	
Longest scaffold	4022579	
Shortest scaffold	892	
Number of scaffolds > 500 nt	657	100.0%
Number of scaffolds > 1K nt	505	76.9%
Number of scaffolds > 10K nt	275	41.9%
Number of scaffolds > 100K nt	214	32.6%
Number of scaffolds > 1M nt	40	6.1%
Mean scaffold size	234537	
Median scaffold size	1418	
N50 scaffold length	901975	
L50 scaffold count	48	
NG50 scaffold length	1417207	
LG50 scaffold count	29	
N50 scaffold - NG50 scaffold length difference	515232	
scaffold %A	29.59	
scaffold %C	20.15	
scaffold %G	20.16	
scaffold %T	29.49	
scaffold %N	0.61	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	99.3%	
Percentage of assembly in unscaffolded contigs	0.7%	
Average number of contigs per scaffold	6.8	
Average length of break (>25 Ns) between contigs in scaffold	1435	
Number of contigs	4477	
Number of contigs in scaffolds	4191	
Number of contigs not in scaffolds	286	
Total size of contigs	153149674	
Longest contig	1239957	
Shortest contig	157	
Number of contigs > 500 nt	4279	95.6%
Number of contigs > 1K nt	4024	89.9%
Number of contigs > 10K nt	3009	67.2%
Number of contigs > 100K nt	280	6.3%
Number of contigs > 1M nt	2	0.0%
Mean contig size	34208	
Median contig size	21167	
N50 contig length	65510	
L50 contig count	696	
NG50 contig length	84611	
LG50 contig count	423	
N50 contig - NG50 contig length difference	19101	
contig %A	29.77	
contig %C	20.28	
contig %G	20.29	
contig %T	29.67	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

M2_scaffolds.fa.gz

Number of scaffolds	5000	
Total size of scaffolds	117797484	
Total scaffold length as percentage of known genome size	104.7%	
Longest scaffold	959012	
Shortest scaffold	892	
Number of scaffolds > 500 nt	5000	100.0%
Number of scaffolds > 1K nt	4834	96.7%
Number of scaffolds > 10K nt	3050	61.0%
Number of scaffolds > 100K nt	106	2.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	23559	
Median scaffold size	14900	
N50 scaffold length	41766	
L50 scaffold count	808	
NG50 scaffold length	43844	
LG50 scaffold count	746	
N50 scaffold - NG50 scaffold length difference	2078	
scaffold %A	29.74	
scaffold %C	20.21	
scaffold %G	20.27	
scaffold %T	29.74	
scaffold %N	0.04	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	22.0%	
Percentage of assembly in unscaffolded contigs	78.0%	
Average number of contigs per scaffold	1.2	
Average length of break (>25 Ns) between contigs in scaffold	9	
Number of contigs	5780	
Number of contigs in scaffolds	1395	
Number of contigs not in scaffolds	4385	
Total size of contigs	117752457	
Longest contig	959012	
Shortest contig	157	
Number of contigs > 500 nt	5563	96.2%
Number of contigs > 1K nt	5299	91.7%
Number of contigs > 10K nt	3301	57.1%
Number of contigs > 100K nt	68	1.2%
Number of contigs > 1M nt	0	0.0%
Mean contig size	20372	
Median contig size	13168	
N50 contig length	36443	
L50 contig count	941	
NG50 contig length	38171	
LG50 contig count	871	
N50 contig - NG50 contig length difference	1728	
contig %A	29.75	
contig %C	20.22	
contig %G	20.27	
contig %T	29.76	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

M3_scaffolds.fa.gz

Number of scaffolds	40344	
Total size of scaffolds	169568979	
Total scaffold length as percentage of known genome size	150.7%	
Longest scaffold	3720691	
Shortest scaffold	101	
Number of scaffolds > 500 nt	5011	12.4%
Number of scaffolds > 1K nt	1237	3.1%
Number of scaffolds > 10K nt	650	1.6%
Number of scaffolds > 100K nt	281	0.7%
Number of scaffolds > 1M nt	31	0.1%
Mean scaffold size	4203	
Median scaffold size	306	
N50 scaffold length	630371	
L50 scaffold count	81	
NG50 scaffold length	836600	
LG50 scaffold count	42	
N50 scaffold - NG50 scaffold length difference	206229	
scaffold %A	28.94	
scaffold %C	19.77	
scaffold %G	19.71	
scaffold %T	28.96	
scaffold %N	2.61	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	91.1%	
Percentage of assembly in unscaffolded contigs	8.9%	
Average number of contigs per scaffold	1.1	
Average length of break (>25 Ns) between contigs in scaffold	109	
Number of contigs	45200	
Number of contigs in scaffolds	6070	
Number of contigs not in scaffolds	39130	
Total size of contigs	165152290	
Longest contig	1240157	
Shortest contig	101	
Number of contigs > 500 nt	8436	18.7%
Number of contigs > 1K nt	4122	9.1%
Number of contigs > 10K nt	3002	6.6%
Number of contigs > 100K nt	268	0.6%
Number of contigs > 1M nt	3	0.0%
Mean contig size	3654	
Median contig size	324	
N50 contig length	58916	
L50 contig count	810	
NG50 contig length	83445	
LG50 contig count	433	
N50 contig - NG50 contig length difference	24529	
contig %A	29.72	
contig %C	20.30	
contig %G	20.24	
contig %T	29.74	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

M4_scaffolds.fa.gz

Number of scaffolds	158	
Total size of scaffolds	118319732	
Total scaffold length as percentage of known genome size	105.2%	
Longest scaffold	3518296	
Shortest scaffold	4447	
Number of scaffolds > 500 nt	158	100.0%
Number of scaffolds > 1K nt	158	100.0%
Number of scaffolds > 10K nt	154	97.5%
Number of scaffolds > 100K nt	143	90.5%
Number of scaffolds > 1M nt	38	24.1%
Mean scaffold size	748859	
Median scaffold size	515477	
N50 scaffold length	1225267	
L50 scaffold count	30	
NG50 scaffold length	1350048	
LG50 scaffold count	28	
N50 scaffold - NG50 scaffold length difference	124781	
scaffold %A	29.16	
scaffold %C	19.89	
scaffold %G	19.93	
scaffold %T	29.15	
scaffold %N	1.87	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	99.8%	
Percentage of assembly in unscaffolded contigs	0.2%	
Average number of contigs per scaffold	16.9	
Average length of break (>25 Ns) between contigs in scaffold	14014	
Number of contigs	2672	
Number of contigs in scaffolds	2662	
Number of contigs not in scaffolds	10	
Total size of contigs	116106704	
Longest contig	1239957	
Shortest contig	3663	
Number of contigs > 500 nt	2672	100.0%
Number of contigs > 1K nt	2672	100.0%
Number of contigs > 10K nt	2251	84.2%
Number of contigs > 100K nt	223	8.3%
Number of contigs > 1M nt	2	0.1%
Mean contig size	43453	
Median contig size	30803	
N50 contig length	67017	
L50 contig count	516	
NG50 contig length	68799	
LG50 contig count	489	
N50 contig - NG50 contig length difference	1782	
contig %A	29.72	
contig %C	20.27	
contig %G	20.31	
contig %T	29.70	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

M5_scaffolds.fa.gz

Number of scaffolds	2500	
Total size of scaffolds	153736947	
Total scaffold length as percentage of known genome size	136.7%	
Longest scaffold	2390989	
Shortest scaffold	832	
Number of scaffolds > 500 nt	2500	100.0%
Number of scaffolds > 1K nt	2390	95.6%
Number of scaffolds > 10K nt	2120	84.8%
Number of scaffolds > 100K nt	481	19.2%
Number of scaffolds > 1M nt	3	0.1%
Mean scaffold size	61495	
Median scaffold size	38619	
N50 scaffold length	110135	
L50 scaffold count	410	
NG50 scaffold length	147982	
LG50 scaffold count	246	
N50 scaffold - NG50 scaffold length difference	37847	
scaffold %A	23.61	
scaffold %C	16.09	
scaffold %G	16.05	
scaffold %T	23.62	
scaffold %N	20.63	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	99.9%	
Percentage of assembly in unscaffolded contigs	0.1%	
Average number of contigs per scaffold	5.6	
Average length of break (>25 Ns) between contigs in scaffold	12689	
Number of contigs	13998	
Number of contigs in scaffolds	13846	
Number of contigs not in scaffolds	152	
Total size of contigs	122016803	
Longest contig	1111891	
Shortest contig	118	
Number of contigs > 500 nt	7262	51.9%
Number of contigs > 1K nt	5075	36.3%
Number of contigs > 10K nt	3294	23.5%
Number of contigs > 100K nt	69	0.5%
Number of contigs > 1M nt	1	0.0%
Mean contig size	8717	
Median contig size	524	
N50 contig length	36938	
L50 contig count	978	
NG50 contig length	40037	
LG50 contig count	855	
N50 contig - NG50 contig length difference	3099	
contig %A	29.75	
contig %C	20.28	
contig %G	20.22	
contig %T	29.75	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

N1_scaffolds.fa.gz

Number of scaffolds	31423	
Total size of scaffolds	139774320	
Total scaffold length as percentage of known genome size	124.2%	
Longest scaffold	26513	
Shortest scaffold	1081	
Number of scaffolds > 500 nt	31423	100.0%
Number of scaffolds > 1K nt	31423	100.0%
Number of scaffolds > 10K nt	1865	5.9%
Number of scaffolds > 100K nt	0	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	4448	
Median scaffold size	3571	
N50 scaffold length	5741	
L50 scaffold count	8103	
NG50 scaffold length	6704	
LG50 scaffold count	5907	
N50 scaffold - NG50 scaffold length difference	963	
scaffold %A	28.24	
scaffold %C	18.94	
scaffold %G	18.91	
scaffold %T	28.44	
scaffold %N	5.46	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	85.0%	
Percentage of assembly in unscaffolded contigs	15.0%	
Average number of contigs per scaffold	2.8	
Average length of break (>25 Ns) between contigs in scaffold	243	
Number of contigs	86428	
Number of contigs in scaffolds	77317	
Number of contigs not in scaffolds	9111	
Total size of contigs	132226959	
Longest contig	11501	
Shortest contig	1	
Number of contigs > 500 nt	63409	73.4%
Number of contigs > 1K nt	51141	59.2%
Number of contigs > 10K nt	20	0.0%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	1530	
Median contig size	1270	
N50 contig length	2387	
L50 contig count	18606	
NG50 contig length	2676	
LG50 contig count	14700	
N50 contig - NG50 contig length difference	289	
contig %A	29.86	
contig %C	20.02	
contig %G	19.99	
contig %T	30.07	
contig %N	0.07	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

N2_scaffolds.fa.gz

Number of scaffolds	33787	
Total size of scaffolds	113389083	
Total scaffold length as percentage of known genome size	100.8%	
Longest scaffold	18206	
Shortest scaffold	1082	
Number of scaffolds > 500 nt	33787	100.0%
Number of scaffolds > 1K nt	33787	100.0%
Number of scaffolds > 10K nt	359	1.1%
Number of scaffolds > 100K nt	0	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	3356	
Median scaffold size	2811	
N50 scaffold length	4035	
L50 scaffold count	9591	
NG50 scaffold length	4061	
LG50 scaffold count	9481	
N50 scaffold - NG50 scaffold length difference	26	
scaffold %A	28.28	
scaffold %C	18.98	
scaffold %G	18.96	
scaffold %T	28.52	
scaffold %N	5.25	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	73.3%	
Percentage of assembly in unscaffolded contigs	26.7%	
Average number of contigs per scaffold	2.1	
Average length of break (>25 Ns) between contigs in scaffold	176	
Number of contigs	69948	
Number of contigs in scaffolds	56215	
Number of contigs not in scaffolds	13733	
Total size of contigs	107460922	
Longest contig	10764	
Shortest contig	1	
Number of contigs > 500 nt	53749	76.8%
Number of contigs > 1K nt	45410	64.9%
Number of contigs > 10K nt	3	0.0%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	1536	
Median contig size	1365	
N50 contig length	2204	
L50 contig count	16690	
NG50 contig length	2129	
LG50 contig count	17853	
N50 contig - NG50 contig length difference	75	
contig %A	29.84	
contig %C	20.03	
contig %G	20.01	
contig %T	30.10	
contig %N	0.02	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

N3_scaffolds.fa.gz

Number of scaffolds	24298	
Total size of scaffolds	146503047	
Total scaffold length as percentage of known genome size	130.2%	
Longest scaffold	40835	
Shortest scaffold	1082	
Number of scaffolds > 500 nt	24298	100.0%
Number of scaffolds > 1K nt	24298	100.0%
Number of scaffolds > 10K nt	3631	14.9%
Number of scaffolds > 100K nt	0	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	6029	
Median scaffold size	5016	
N50 scaffold length	7850	
L50 scaffold count	6300	
NG50 scaffold length	9358	
LG50 scaffold count	4317	
N50 scaffold - NG50 scaffold length difference	1508	
scaffold %A	27.61	
scaffold %C	18.51	
scaffold %G	18.48	
scaffold %T	27.78	
scaffold %N	7.62	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	95.5%	
Percentage of assembly in unscaffolded contigs	4.5%	
Average number of contigs per scaffold	4.3	
Average length of break (>25 Ns) between contigs in scaffold	459	
Number of contigs	103555	
Number of contigs in scaffolds	100634	
Number of contigs not in scaffolds	2921	
Total size of contigs	135479743	
Longest contig	14662	
Shortest contig	1	
Number of contigs > 500 nt	69424	67.0%
Number of contigs > 1K nt	51787	50.0%
Number of contigs > 10K nt	12	0.0%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	1308	
Median contig size	1001	
N50 contig length	2196	
L50 contig count	20445	
NG50 contig length	2526	
LG50 contig count	15562	
N50 contig - NG50 contig length difference	330	
contig %A	29.86	
contig %C	20.01	
contig %G	19.99	
contig %T	30.04	
contig %N	0.11	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

O1_scaffolds.fa.gz

Number of scaffolds	14994	
Total size of scaffolds	90333710	
Total scaffold length as percentage of known genome size	80.3%	
Longest scaffold	74215	
Shortest scaffold	2000	
Number of scaffolds > 500 nt	14994	100.0%
Number of scaffolds > 1K nt	14994	100.0%
Number of scaffolds > 10K nt	2267	15.1%
Number of scaffolds > 100K nt	0	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	6025	
Median scaffold size	4360	
N50 scaffold length	7847	
L50 scaffold count	3543	
NG50 scaffold length	6000	
LG50 scaffold count	5153	
N50 scaffold - NG50 scaffold length difference	1847	
scaffold %A	29.94	
scaffold %C	20.06	
scaffold %G	20.07	
scaffold %T	29.93	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in un scaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	14994	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	14994	
Total size of contigs	90333710	
Longest contig	74215	
Shortest contig	2000	
Number of contigs > 500 nt	14994	100.0%
Number of contigs > 1K nt	14994	100.0%
Number of contigs > 10K nt	2267	15.1%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	6025	
Median contig size	4360	
N50 contig length	7847	
L50 contig count	3543	
NG50 contig length	6000	
LG50 contig count	5153	
N50 contig - NG50 contig length difference	1847	
contig %A	29.94	
contig %C	20.06	
contig %G	20.07	
contig %T	29.93	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

P1_scaffolds.fa.gz

Number of scaffolds	4143	
Total size of scaffolds	114258553	
Total scaffold length as percentage of known genome size	101.6%	
Longest scaffold	4397505	
Shortest scaffold	100	
Number of scaffolds > 500 nt	905	21.8%
Number of scaffolds > 1K nt	316	7.6%
Number of scaffolds > 10K nt	125	3.0%
Number of scaffolds > 100K nt	116	2.8%
Number of scaffolds > 1M nt	42	1.0%
Mean scaffold size	27579	
Median scaffold size	175	
N50 scaffold length	1707112	
L50 scaffold count	24	
NG50 scaffold length	1716225	
LG50 scaffold count	23	
N50 scaffold - NG50 scaffold length difference	9113	
scaffold %A	29.83	
scaffold %C	19.94	
scaffold %G	19.95	
scaffold %T	29.86	
scaffold %N	0.42	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	93.9%	
Percentage of assembly in unscaffolded contigs	6.1%	
Average number of contigs per scaffold	1.1	
Average length of break (>25 Ns) between contigs in scaffold	115	
Number of contigs	4566	
Number of contigs in scaffolds	521	
Number of contigs not in scaffolds	4045	
Total size of contigs	113781081	
Longest contig	1814562	
Shortest contig	100	
Number of contigs > 500 nt	1321	28.9%
Number of contigs > 1K nt	730	16.0%
Number of contigs > 10K nt	510	11.2%
Number of contigs > 100K nt	343	7.5%
Number of contigs > 1M nt	4	0.1%
Mean contig size	24919	
Median contig size	205	
N50 contig length	343889	
L50 contig count	103	
NG50 contig length	346960	
LG50 contig count	101	
N50 contig - NG50 contig length difference	3071	
contig %A	29.96	
contig %C	20.02	
contig %G	20.03	
contig %T	29.98	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

Q1_scaffolds.fa.gz

Number of scaffolds	990	
Total size of scaffolds	120855878	
Total scaffold length as percentage of known genome size	107.4%	
Longest scaffold	17101185	
Shortest scaffold	1004	
Number of scaffolds > 500 nt	990	100.0%
Number of scaffolds > 1K nt	990	100.0%
Number of scaffolds > 10K nt	123	12.4%
Number of scaffolds > 100K nt	36	3.6%
Number of scaffolds > 1M nt	20	2.0%
Mean scaffold size	122077	
Median scaffold size	4528	
N50 scaffold length	8285827	
L50 scaffold count	6	
NG50 scaffold length	8396795	
LG50 scaffold count	5	
N50 scaffold - NG50 scaffold length difference	110968	
scaffold %A	29.34	
scaffold %C	19.81	
scaffold %G	19.81	
scaffold %T	29.33	
scaffold %N	1.70	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	93.5%	
Percentage of assembly in unscaffolded contigs	6.5%	
Average number of contigs per scaffold	2.0	
Average length of break (>25 Ns) between contigs in scaffold	2077	
Number of contigs	1946	
Number of contigs in scaffolds	1000	
Number of contigs not in scaffolds	946	
Total size of contigs	119174112	
Longest contig	858807	
Shortest contig	1	
Number of contigs > 500 nt	1869	96.0%
Number of contigs > 1K nt	1861	95.6%
Number of contigs > 10K nt	904	46.5%
Number of contigs > 100K nt	397	20.4%
Number of contigs > 1M nt	0	0.0%
Mean contig size	61241	
Median contig size	8554	
N50 contig length	208256	
L50 contig count	181	
NG50 contig length	219906	
LG50 contig count	165	
N50 contig - NG50 contig length difference	11650	
contig %A	29.76	
contig %C	20.09	
contig %G	20.09	
contig %T	29.75	
contig %N	0.31	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

V1_scaffolds.fa.gz

Number of scaffolds	3230202	
Total size of scaffolds	227922042	
Total scaffold length as percentage of known genome size	202.6%	
Longest scaffold	352	
Shortest scaffold	61	
Number of scaffolds > 500 nt	0	0.0%
Number of scaffolds > 1K nt	0	0.0%
Number of scaffolds > 10K nt	0	0.0%
Number of scaffolds > 100K nt	0	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	71	
Median scaffold size	61	
N50 scaffold length	63	
L50 scaffold count	1365014	
NG50 scaffold length	82	
LG50 scaffold count	557254	
N50 scaffold - NG50 scaffold length difference	19	
scaffold %A	29.58	
scaffold %C	20.44	
scaffold %G	20.39	
scaffold %T	29.60	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in un scaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	3230202	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	3230202	
Total size of contigs	227922042	
Longest contig	352	
Shortest contig	61	
Number of contigs > 500 nt	0	0.0%
Number of contigs > 1K nt	0	0.0%
Number of contigs > 10K nt	0	0.0%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	71	
Median contig size	61	
N50 contig length	63	
L50 contig count	1365014	
NG50 contig length	82	
LG50 contig count	557254	
N50 contig - NG50 contig length difference	19	
contig %A	29.58	
contig %C	20.44	
contig %G	20.39	
contig %T	29.60	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

V2_scaffolds.fa.gz

Number of scaffolds	3239006	
Total size of scaffolds	228736023	
Total scaffold length as percentage of known genome size	203.3%	
Longest scaffold	367	
Shortest scaffold	61	
Number of scaffolds > 500 nt	0	0.0%
Number of scaffolds > 1K nt	0	0.0%
Number of scaffolds > 10K nt	0	0.0%
Number of scaffolds > 100K nt	0	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	71	
Median scaffold size	61	
N50 scaffold length	64	
L50 scaffold count	1367408	
NG50 scaffold length	82	
LG50 scaffold count	556281	
N50 scaffold - NG50 scaffold length difference	18	
scaffold %A	29.56	
scaffold %C	20.47	
scaffold %G	20.36	
scaffold %T	29.62	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in un scaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	3239006	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	3239006	
Total size of contigs	228736023	
Longest contig	367	
Shortest contig	61	
Number of contigs > 500 nt	0	0.0%
Number of contigs > 1K nt	0	0.0%
Number of contigs > 10K nt	0	0.0%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	71	
Median contig size	61	
N50 contig length	64	
L50 contig count	1367408	
NG50 contig length	82	
LG50 contig count	556281	
N50 contig - NG50 contig length difference	18	
contig %A	29.56	
contig %C	20.47	
contig %G	20.36	
contig %T	29.62	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

V3_scaffolds.fa.gz

Number of scaffolds	3243118	
Total size of scaffolds	228987512	
Total scaffold length as percentage of known genome size	203.5%	
Longest scaffold	503	
Shortest scaffold	61	
Number of scaffolds > 500 nt	1	0.0%
Number of scaffolds > 1K nt	0	0.0%
Number of scaffolds > 10K nt	0	0.0%
Number of scaffolds > 100K nt	0	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	71	
Median scaffold size	61	
N50 scaffold length	63	
L50 scaffold count	1369364	
NG50 scaffold length	82	
LG50 scaffold count	556288	
N50 scaffold - NG50 scaffold length difference	19	
scaffold %A	29.56	
scaffold %C	20.46	
scaffold %G	20.36	
scaffold %T	29.62	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in un scaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	3243118	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	3243118	
Total size of contigs	228987512	
Longest contig	503	
Shortest contig	61	
Number of contigs > 500 nt	1	0.0%
Number of contigs > 1K nt	0	0.0%
Number of contigs > 10K nt	0	0.0%
Number of contigs > 100K nt	0	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	71	
Median contig size	61	
N50 contig length	63	
L50 contig count	1369364	
NG50 contig length	82	
LG50 contig count	556288	
N50 contig - NG50 contig length difference	19	
contig %A	29.56	
contig %C	20.46	
contig %G	20.36	
contig %T	29.62	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

V4_scaffolds.fa.gz

Number of scaffolds	32617	
Total size of scaffolds	127185684	
Total scaffold length as percentage of known genome size	113.1%	
Longest scaffold	4743985	
Shortest scaffold	61	
Number of scaffolds > 500 nt	1324	4.1%
Number of scaffolds > 1K nt	689	2.1%
Number of scaffolds > 10K nt	278	0.9%
Number of scaffolds > 100K nt	207	0.6%
Number of scaffolds > 1M nt	32	0.1%
Mean scaffold size	3899	
Median scaffold size	69	
N50 scaffold length	798306	
L50 scaffold count	46	
NG50 scaffold length	879312	
LG50 scaffold count	38	
N50 scaffold - NG50 scaffold length difference	81006	
scaffold %A	27.28	
scaffold %C	18.53	
scaffold %G	18.54	
scaffold %T	27.22	
scaffold %N	8.42	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	97.1%	
Percentage of assembly in unscaffolded contigs	2.9%	
Average number of contigs per scaffold	1.6	
Average length of break (>25 Ns) between contigs in scaffold	328	
Number of contigs	51760	
Number of contigs in scaffolds	19633	
Number of contigs not in scaffolds	32127	
Total size of contigs	116549762	
Longest contig	1368779	
Shortest contig	51	
Number of contigs > 500 nt	19175	37.0%
Number of contigs > 1K nt	17258	33.3%
Number of contigs > 10K nt	3080	6.0%
Number of contigs > 100K nt	12	0.0%
Number of contigs > 1M nt	1	0.0%
Mean contig size	2252	
Median contig size	92	
N50 contig length	8856	
L50 contig count	3838	
NG50 contig length	9192	
LG50 contig count	3613	
N50 contig - NG50 contig length difference	336	
contig %A	29.77	
contig %C	20.22	
contig %G	20.23	
contig %T	29.71	
contig %N	0.07	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

V5_scaffolds.fa.gz

Number of scaffolds	44364	
Total size of scaffolds	117239847	
Total scaffold length as percentage of known genome size	104.2%	
Longest scaffold	134380	
Shortest scaffold	61	
Number of scaffolds > 500 nt	12376	27.9%
Number of scaffolds > 1K nt	11531	26.0%
Number of scaffolds > 10K nt	4091	9.2%
Number of scaffolds > 100K nt	3	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	2643	
Median scaffold size	78	
N50 scaffold length	14739	
L50 scaffold count	2386	
NG50 scaffold length	15369	
LG50 scaffold count	2229	
N50 scaffold - NG50 scaffold length difference	630	
scaffold %A	29.62	
scaffold %C	20.12	
scaffold %G	20.15	
scaffold %T	29.60	
scaffold %N	0.50	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	64.6%	
Percentage of assembly in unscaffolded contigs	35.4%	
Average number of contigs per scaffold	1.2	
Average length of break (>25 Ns) between contigs in scaffold	13	
Number of contigs	53949	
Number of contigs in scaffolds	14996	
Number of contigs not in scaffolds	38953	
Total size of contigs	116706574	
Longest contig	121503	
Shortest contig	51	
Number of contigs > 500 nt	20744	38.5%
Number of contigs > 1K nt	18817	34.9%
Number of contigs > 10K nt	2993	5.5%
Number of contigs > 100K nt	2	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	2163	
Median contig size	98	
N50 contig length	8141	
L50 contig count	4284	
NG50 contig length	8473	
LG50 contig count	4031	
N50 contig - NG50 contig length difference	332	
contig %A	29.76	
contig %C	20.21	
contig %G	20.25	
contig %T	29.74	
contig %N	0.05	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

V6_scaffolds.fa.gz

Number of scaffolds	65820	
Total size of scaffolds	116793074	
Total scaffold length as percentage of known genome size	103.8%	
Longest scaffold	132122	
Shortest scaffold	61	
Number of scaffolds > 500 nt	28662	43.5%
Number of scaffolds > 1K nt	24833	37.7%
Number of scaffolds > 10K nt	1607	2.4%
Number of scaffolds > 100K nt	2	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	1774	
Median scaffold size	220	
N50 scaffold length	5468	
L50 scaffold count	6344	
NG50 scaffold length	5671	
LG50 scaffold count	5958	
N50 scaffold - NG50 scaffold length difference	203	
scaffold %A	29.76	
scaffold %C	20.23	
scaffold %G	20.23	
scaffold %T	29.75	
scaffold %N	0.02	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.1%	
Percentage of assembly in unscaffolded contigs	99.9%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	65834	
Number of contigs in scaffolds	28	
Number of contigs not in scaffolds	65806	
Total size of contigs	116792643	
Longest contig	132122	
Shortest contig	61	
Number of contigs > 500 nt	28671	43.6%
Number of contigs > 1K nt	24840	37.7%
Number of contigs > 10K nt	1603	2.4%
Number of contigs > 100K nt	2	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	1774	
Median contig size	220	
N50 contig length	5467	
L50 contig count	6349	
NG50 contig length	5665	
LG50 contig count	5963	
N50 contig - NG50 contig length difference	198	
contig %A	29.76	
contig %C	20.23	
contig %G	20.23	
contig %T	29.75	
contig %N	0.02	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W10_scaffolds.fa.gz

Number of scaffolds	24778	
Total size of scaffolds	115023920	
Total scaffold length as percentage of known genome size	102.2%	
Longest scaffold	364958	
Shortest scaffold	108	
Number of scaffolds > 500 nt	21385	86.3%
Number of scaffolds > 1K nt	18900	76.3%
Number of scaffolds > 10K nt	2880	11.6%
Number of scaffolds > 100K nt	20	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	4642	
Median scaffold size	2992	
N50 scaffold length	8190	
L50 scaffold count	4144	
NG50 scaffold length	8380	
LG50 scaffold count	3992	
N50 scaffold - NG50 scaffold length difference	190	
scaffold %A	29.78	
scaffold %C	20.23	
scaffold %G	20.25	
scaffold %T	29.74	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in unscaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	24778	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	24778	
Total size of contigs	115023920	
Longest contig	364958	
Shortest contig	108	
Number of contigs > 500 nt	21385	86.3%
Number of contigs > 1K nt	18900	76.3%
Number of contigs > 10K nt	2880	11.6%
Number of contigs > 100K nt	20	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	4642	
Median contig size	2992	
N50 contig length	8190	
L50 contig count	4144	
NG50 contig length	8380	
LG50 contig count	3992	
N50 contig - NG50 contig length difference	190	
contig %A	29.78	
contig %C	20.23	
contig %G	20.25	
contig %T	29.74	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W11_scaffolds.fa.gz

Number of scaffolds	17581	
Total size of scaffolds	115453668	
Total scaffold length as percentage of known genome size	102.6%	
Longest scaffold	665472	
Shortest scaffold	200	
Number of scaffolds > 500 nt	15531	88.3%
Number of scaffolds > 1K nt	14163	80.6%
Number of scaffolds > 10K nt	3995	22.7%
Number of scaffolds > 100K nt	17	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	6567	
Median scaffold size	4139	
N50 scaffold length	11836	
L50 scaffold count	2985	
NG50 scaffold length	12127	
LG50 scaffold count	2862	
N50 scaffold - NG50 scaffold length difference	291	
scaffold %A	29.75	
scaffold %C	20.22	
scaffold %G	20.24	
scaffold %T	29.75	
scaffold %N	0.04	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	2.7%	
Percentage of assembly in unscaffolded contigs	97.3%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	2	
Number of contigs	17979	
Number of contigs in scaffolds	783	
Number of contigs not in scaffolds	17196	
Total size of contigs	115441384	
Longest contig	665472	
Shortest contig	24	
Number of contigs > 500 nt	15613	86.8%
Number of contigs > 1K nt	14233	79.2%
Number of contigs > 10K nt	3991	22.2%
Number of contigs > 100K nt	17	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	6421	
Median contig size	4023	
N50 contig length	11777	
L50 contig count	3009	
NG50 contig length	12032	
LG50 contig count	2885	
N50 contig - NG50 contig length difference	255	
contig %A	29.75	
contig %C	20.23	
contig %G	20.24	
contig %T	29.75	
contig %N	0.03	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W1_scaffolds.fa.gz

Number of scaffolds	17187	
Total size of scaffolds	115264020	
Total scaffold length as percentage of known genome size	102.5%	
Longest scaffold	428217	
Shortest scaffold	200	
Number of scaffolds > 500 nt	14995	87.2%
Number of scaffolds > 1K nt	13745	80.0%
Number of scaffolds > 10K nt	4019	23.4%
Number of scaffolds > 100K nt	19	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	6706	
Median scaffold size	4444	
N50 scaffold length	11936	
L50 scaffold count	2996	
NG50 scaffold length	12187	
LG50 scaffold count	2882	
N50 scaffold - NG50 scaffold length difference	251	
scaffold %A	29.76	
scaffold %C	20.22	
scaffold %G	20.24	
scaffold %T	29.73	
scaffold %N	0.05	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	4.6%	
Percentage of assembly in unscaffolded contigs	95.4%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	3	
Number of contigs	17759	
Number of contigs in scaffolds	1119	
Number of contigs not in scaffolds	16640	
Total size of contigs	115245941	
Longest contig	428217	
Shortest contig	24	
Number of contigs > 500 nt	15165	85.4%
Number of contigs > 1K nt	13891	78.2%
Number of contigs > 10K nt	3999	22.5%
Number of contigs > 100K nt	19	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	6489	
Median contig size	4237	
N50 contig length	11765	
L50 contig count	3039	
NG50 contig length	12061	
LG50 contig count	2923	
N50 contig - NG50 contig length difference	296	
contig %A	29.76	
contig %C	20.23	
contig %G	20.24	
contig %T	29.74	
contig %N	0.03	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W3_scaffolds.fa.gz

Number of scaffolds	24858	
Total size of scaffolds	115114064	
Total scaffold length as percentage of known genome size	102.3%	
Longest scaffold	364958	
Shortest scaffold	200	
Number of scaffolds > 500 nt	21431	86.2%
Number of scaffolds > 1K nt	18974	76.3%
Number of scaffolds > 10K nt	2865	11.5%
Number of scaffolds > 100K nt	20	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	4631	
Median scaffold size	2991	
N50 scaffold length	8180	
L50 scaffold count	4158	
NG50 scaffold length	8372	
LG50 scaffold count	4000	
N50 scaffold - NG50 scaffold length difference	192	
scaffold %A	29.75	
scaffold %C	20.22	
scaffold %G	20.24	
scaffold %T	29.74	
scaffold %N	0.04	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	2.2%	
Percentage of assembly in unscaffolded contigs	97.8%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	1	
Number of contigs	25328	
Number of contigs in scaffolds	936	
Number of contigs not in scaffolds	24392	
Total size of contigs	115098816	
Longest contig	364958	
Shortest contig	24	
Number of contigs > 500 nt	21464	84.7%
Number of contigs > 1K nt	18997	75.0%
Number of contigs > 10K nt	2864	11.3%
Number of contigs > 100K nt	20	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	4544	
Median contig size	2899	
N50 contig length	8169	
L50 contig count	4166	
NG50 contig length	8357	
LG50 contig count	4009	
N50 contig - NG50 contig length difference	188	
contig %A	29.76	
contig %C	20.22	
contig %G	20.25	
contig %T	29.75	
contig %N	0.03	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W5_scaffolds.fa.gz

Number of scaffolds	19928	
Total size of scaffolds	115373922	
Total scaffold length as percentage of known genome size	102.6%	
Longest scaffold	665472	
Shortest scaffold	200	
Number of scaffolds > 500 nt	17625	88.4%
Number of scaffolds > 1K nt	15967	80.1%
Number of scaffolds > 10K nt	3465	17.4%
Number of scaffolds > 100K nt	18	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	5790	
Median scaffold size	3704	
N50 scaffold length	10286	
L50 scaffold count	3310	
NG50 scaffold length	10560	
LG50 scaffold count	3172	
N50 scaffold - NG50 scaffold length difference	274	
scaffold %A	29.74	
scaffold %C	20.22	
scaffold %G	20.24	
scaffold %T	29.75	
scaffold %N	0.05	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	3.8%	
Percentage of assembly in unscaffolded contigs	96.2%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	2	
Number of contigs	20561	
Number of contigs in scaffolds	1248	
Number of contigs not in scaffolds	19313	
Total size of contigs	115354785	
Longest contig	665472	
Shortest contig	24	
Number of contigs > 500 nt	17728	86.2%
Number of contigs > 1K nt	16050	78.1%
Number of contigs > 10K nt	3441	16.7%
Number of contigs > 100K nt	18	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	5610	
Median contig size	3533	
N50 contig length	10181	
L50 contig count	3332	
NG50 contig length	10465	
LG50 contig count	3194	
N50 contig - NG50 contig length difference	284	
contig %A	29.75	
contig %C	20.22	
contig %G	20.24	
contig %T	29.75	
contig %N	0.03	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W6_scaffolds.fa.gz

Number of scaffolds	25991	
Total size of scaffolds	115297130	
Total scaffold length as percentage of known genome size	102.5%	
Longest scaffold	364958	
Shortest scaffold	200	
Number of scaffolds > 500 nt	22503	86.6%
Number of scaffolds > 1K nt	19817	76.2%
Number of scaffolds > 10K nt	2658	10.2%
Number of scaffolds > 100K nt	18	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	4436	
Median scaffold size	2845	
N50 scaffold length	7739	
L50 scaffold count	4366	
NG50 scaffold length	7936	
LG50 scaffold count	4188	
N50 scaffold - NG50 scaffold length difference	197	
scaffold %A	29.77	
scaffold %C	20.22	
scaffold %G	20.24	
scaffold %T	29.73	
scaffold %N	0.04	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	2.4%	
Percentage of assembly in unscaffolded contigs	97.6%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	1	
Number of contigs	26531	
Number of contigs in scaffolds	1068	
Number of contigs not in scaffolds	25463	
Total size of contigs	115280793	
Longest contig	364958	
Shortest contig	24	
Number of contigs > 500 nt	22554	85.0%
Number of contigs > 1K nt	19855	74.8%
Number of contigs > 10K nt	2648	10.0%
Number of contigs > 100K nt	18	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	4345	
Median contig size	2768	
N50 contig length	7715	
L50 contig count	4381	
NG50 contig length	7899	
LG50 contig count	4203	
N50 contig - NG50 contig length difference	184	
contig %A	29.77	
contig %C	20.22	
contig %G	20.25	
contig %T	29.73	
contig %N	0.03	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W7_scaffolds.fa.gz

Number of scaffolds	18342	
Total size of scaffolds	115217841	
Total scaffold length as percentage of known genome size	102.4%	
Longest scaffold	428202	
Shortest scaffold	200	
Number of scaffolds > 500 nt	16085	87.7%
Number of scaffolds > 1K nt	14740	80.4%
Number of scaffolds > 10K nt	3839	20.9%
Number of scaffolds > 100K nt	20	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	6282	
Median scaffold size	4193	
N50 scaffold length	11107	
L50 scaffold count	3228	
NG50 scaffold length	11284	
LG50 scaffold count	3107	
N50 scaffold - NG50 scaffold length difference	177	
scaffold %A	29.77	
scaffold %C	20.22	
scaffold %G	20.24	
scaffold %T	29.72	
scaffold %N	0.05	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	4.2%	
Percentage of assembly in unscaffolded contigs	95.8%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	3	
Number of contigs	18929	
Number of contigs in scaffolds	1158	
Number of contigs not in scaffolds	17771	
Total size of contigs	115199002	
Longest contig	428202	
Shortest contig	24	
Number of contigs > 500 nt	16226	85.7%
Number of contigs > 1K nt	14862	78.5%
Number of contigs > 10K nt	3815	20.2%
Number of contigs > 100K nt	20	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	6086	
Median contig size	3992	
N50 contig length	11013	
L50 contig count	3259	
NG50 contig length	11204	
LG50 contig count	3138	
N50 contig - NG50 contig length difference	191	
contig %A	29.77	
contig %C	20.22	
contig %G	20.24	
contig %T	29.73	
contig %N	0.03	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W8_scaffolds.fa.gz

Number of scaffolds	18213	
Total size of scaffolds	115424301	
Total scaffold length as percentage of known genome size	102.6%	
Longest scaffold	339806	
Shortest scaffold	200	
Number of scaffolds > 500 nt	16167	88.8%
Number of scaffolds > 1K nt	14818	81.4%
Number of scaffolds > 10K nt	3826	21.0%
Number of scaffolds > 100K nt	19	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	6337	
Median scaffold size	4251	
N50 scaffold length	11079	
L50 scaffold count	3212	
NG50 scaffold length	11326	
LG50 scaffold count	3082	
N50 scaffold - NG50 scaffold length difference	247	
scaffold %A	29.76	
scaffold %C	20.22	
scaffold %G	20.24	
scaffold %T	29.74	
scaffold %N	0.04	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	3.5%	
Percentage of assembly in unscaffolded contigs	96.5%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	2	
Number of contigs	18725	
Number of contigs in scaffolds	1005	
Number of contigs not in scaffolds	17720	
Total size of contigs	115408235	
Longest contig	339806	
Shortest contig	24	
Number of contigs > 500 nt	16266	86.9%
Number of contigs > 1K nt	14908	79.6%
Number of contigs > 10K nt	3803	20.3%
Number of contigs > 100K nt	19	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	6163	
Median contig size	4077	
N50 contig length	11004	
L50 contig count	3235	
NG50 contig length	11256	
LG50 contig count	3104	
N50 contig - NG50 contig length difference	252	
contig %A	29.76	
contig %C	20.23	
contig %G	20.24	
contig %T	29.74	
contig %N	0.03	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

W9_scaffolds.fa.gz

Number of scaffolds	19862	
Total size of scaffolds	115154477	
Total scaffold length as percentage of known genome size	102.4%	
Longest scaffold	665681	
Shortest scaffold	110	
Number of scaffolds > 500 nt	17404	87.6%
Number of scaffolds > 1K nt	15765	79.4%
Number of scaffolds > 10K nt	3524	17.7%
Number of scaffolds > 100K nt	18	0.1%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	5798	
Median scaffold size	3680	
N50 scaffold length	10472	
L50 scaffold count	3277	
NG50 scaffold length	10727	
LG50 scaffold count	3151	
N50 scaffold - NG50 scaffold length difference	255	
scaffold %A	29.79	
scaffold %C	20.24	
scaffold %G	20.24	
scaffold %T	29.74	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in unscaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	19862	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	19862	
Total size of contigs	115154477	
Longest contig	665681	
Shortest contig	110	
Number of contigs > 500 nt	17404	87.6%
Number of contigs > 1K nt	15765	79.4%
Number of contigs > 10K nt	3524	17.7%
Number of contigs > 100K nt	18	0.1%
Number of contigs > 1M nt	0	0.0%
Mean contig size	5798	
Median contig size	3680	
N50 contig length	10472	
L50 contig count	3277	
NG50 contig length	10727	
LG50 contig count	3151	
N50 contig - NG50 contig length difference	255	
contig %A	29.79	
contig %C	20.24	
contig %G	20.24	
contig %T	29.74	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

X1_scaffolds.fa.gz

Number of scaffolds	151203	
Total size of scaffolds	159381550	
Total scaffold length as percentage of known genome size	141.7%	
Longest scaffold	524383	
Shortest scaffold	31	
Number of scaffolds > 500 nt	15215	10.1%
Number of scaffolds > 1K nt	14237	9.4%
Number of scaffolds > 10K nt	5744	3.8%
Number of scaffolds > 100K nt	18	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	1054	
Median scaffold size	42	
N50 scaffold length	15303	
L50 scaffold count	3187	
NG50 scaffold length	20226	
LG50 scaffold count	1849	
N50 scaffold - NG50 scaffold length difference	4923	
scaffold %A	29.85	
scaffold %C	20.14	
scaffold %G	20.15	
scaffold %T	29.82	
scaffold %N	0.03	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	5733	
Percentage of assembly in scaffolded contigs	3.5%	
Percentage of assembly in unscaffolded contigs	96.5%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	151852	
Number of contigs in scaffolds	1281	
Number of contigs not in scaffolds	150571	
Total size of contigs	159336958	
Longest contig	524383	
Shortest contig	31	
Number of contigs > 500 nt	15373	10.1%
Number of contigs > 1K nt	14388	9.5%
Number of contigs > 10K nt	5724	3.8%
Number of contigs > 100K nt	18	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	1049	
Median contig size	43	
N50 contig length	15165	
L50 contig count	3210	
NG50 contig length	20054	
LG50 contig count	1861	
N50 contig - NG50 contig length difference	4889	
contig %A	29.86	
contig %C	20.15	
contig %G	20.16	
contig %T	29.83	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	5733	

X2_scaffolds.fa.gz

Number of scaffolds	140598	
Total size of scaffolds	161428633	
Total scaffold length as percentage of known genome size	143.5%	
Longest scaffold	722867	
Shortest scaffold	31	
Number of scaffolds > 500 nt	14037	10.0%
Number of scaffolds > 1K nt	13192	9.4%
Number of scaffolds > 10K nt	5975	4.2%
Number of scaffolds > 100K nt	18	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	1148	
Median scaffold size	43	
N50 scaffold length	16267	
L50 scaffold count	3067	
NG50 scaffold length	21533	
LG50 scaffold count	1756	
N50 scaffold - NG50 scaffold length difference	5266	
scaffold %A	29.84	
scaffold %C	20.14	
scaffold %G	20.15	
scaffold %T	29.84	
scaffold %N	0.02	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	5708	
Percentage of assembly in scaffolded contigs	2.7%	
Percentage of assembly in unscaffolded contigs	97.3%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	141144	
Number of contigs in scaffolds	1079	
Number of contigs not in scaffolds	140065	
Total size of contigs	161394359	
Longest contig	722867	
Shortest contig	31	
Number of contigs > 500 nt	14106	10.0%
Number of contigs > 1K nt	13274	9.4%
Number of contigs > 10K nt	5959	4.2%
Number of contigs > 100K nt	18	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	1143	
Median contig size	43	
N50 contig length	16191	
L50 contig count	3078	
NG50 contig length	21476	
LG50 contig count	1761	
N50 contig - NG50 contig length difference	5285	
contig %A	29.85	
contig %C	20.14	
contig %G	20.16	
contig %T	29.85	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	5708	

X3_scaffolds.fa.gz

Number of scaffolds	328797	
Total size of scaffolds	126688003	
Total scaffold length as percentage of known genome size	112.6%	
Longest scaffold	156082	
Shortest scaffold	31	
Number of scaffolds > 500 nt	30708	9.3%
Number of scaffolds > 1K nt	25514	7.8%
Number of scaffolds > 10K nt	1471	0.4%
Number of scaffolds > 100K nt	2	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	385	
Median scaffold size	39	
N50 scaffold length	4611	
L50 scaffold count	7832	
NG50 scaffold length	5249	
LG50 scaffold count	6388	
N50 scaffold - NG50 scaffold length difference	638	
scaffold %A	29.79	
scaffold %C	20.20	
scaffold %G	20.20	
scaffold %T	29.81	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in unscaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	328797	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	328797	
Total size of contigs	126688003	
Longest contig	156082	
Shortest contig	31	
Number of contigs > 500 nt	30708	9.3%
Number of contigs > 1K nt	25514	7.8%
Number of contigs > 10K nt	1471	0.4%
Number of contigs > 100K nt	2	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	385	
Median contig size	39	
N50 contig length	4611	
L50 contig count	7832	
NG50 contig length	5249	
LG50 contig count	6388	
N50 contig - NG50 contig length difference	638	
contig %A	29.79	
contig %C	20.20	
contig %G	20.20	
contig %T	29.81	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

X4_scaffolds.fa.gz

Number of scaffolds	173372	
Total size of scaffolds	127031925	
Total scaffold length as percentage of known genome size	112.9%	
Longest scaffold	541257	
Shortest scaffold	31	
Number of scaffolds > 500 nt	11920	6.9%
Number of scaffolds > 1K nt	10955	6.3%
Number of scaffolds > 10K nt	4267	2.5%
Number of scaffolds > 100K nt	18	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	733	
Median scaffold size	39	
N50 scaffold length	15662	
L50 scaffold count	2370	
NG50 scaffold length	17580	
LG50 scaffold count	1933	
N50 scaffold - NG50 scaffold length difference	1918	
scaffold %A	29.79	
scaffold %C	20.19	
scaffold %G	20.17	
scaffold %T	29.76	
scaffold %N	0.08	
scaffold %non-ACGTN	0.01	
Number of scaffold non-ACGTN nt	17573	
Percentage of assembly in scaffolded contigs	18.8%	
Percentage of assembly in unscaffolded contigs	81.2%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	175163	
Number of contigs in scaffolds	3414	
Number of contigs not in scaffolds	171749	
Total size of contigs	126937664	
Longest contig	541257	
Shortest contig	31	
Number of contigs > 500 nt	13036	7.4%
Number of contigs > 1K nt	11949	6.8%
Number of contigs > 10K nt	4209	2.4%
Number of contigs > 100K nt	18	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	725	
Median contig size	39	
N50 contig length	14169	
L50 contig count	2605	
NG50 contig length	16018	
LG50 contig count	2124	
N50 contig - NG50 contig length difference	1849	
contig %A	29.82	
contig %C	20.21	
contig %G	20.18	
contig %T	29.78	
contig %N	0.00	
contig %non-ACGTN	0.01	
Number of contig non-ACGTN nt	17573	

X5_scaffolds.fa.gz

Number of scaffolds	172923	
Total size of scaffolds	126671540	
Total scaffold length as percentage of known genome size	112.6%	
Longest scaffold	541257	
Shortest scaffold	31	
Number of scaffolds > 500 nt	11936	6.9%
Number of scaffolds > 1K nt	10957	6.3%
Number of scaffolds > 10K nt	4263	2.5%
Number of scaffolds > 100K nt	18	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	733	
Median scaffold size	39	
N50 scaffold length	15505	
L50 scaffold count	2386	
NG50 scaffold length	17451	
LG50 scaffold count	1956	
N50 scaffold - NG50 scaffold length difference	1946	
scaffold %A	29.76	
scaffold %C	20.18	
scaffold %G	20.18	
scaffold %T	29.79	
scaffold %N	0.08	
scaffold %non-ACGTN	0.01	
Number of scaffold non-ACGTN nt	17326	
Percentage of assembly in scaffolded contigs	18.3%	
Percentage of assembly in unscaffolded contigs	81.7%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	174679	
Number of contigs in scaffolds	3352	
Number of contigs not in scaffolds	171327	
Total size of contigs	126580753	
Longest contig	541257	
Shortest contig	31	
Number of contigs > 500 nt	13009	7.4%
Number of contigs > 1K nt	11915	6.8%
Number of contigs > 10K nt	4210	2.4%
Number of contigs > 100K nt	18	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	725	
Median contig size	39	
N50 contig length	14168	
L50 contig count	2614	
NG50 contig length	15872	
LG50 contig count	2144	
N50 contig - NG50 contig length difference	1704	
contig %A	29.78	
contig %C	20.20	
contig %G	20.19	
contig %T	29.81	
contig %N	0.00	
contig %non-ACGTN	0.01	
Number of contig non-ACGTN nt	17326	

X6_scaffolds.fa.gz

Number of scaffolds	311185	
Total size of scaffolds	126049808	
Total scaffold length as percentage of known genome size	112.0%	
Longest scaffold	156082	
Shortest scaffold	31	
Number of scaffolds > 500 nt	30949	9.9%
Number of scaffolds > 1K nt	25660	8.2%
Number of scaffolds > 10K nt	1443	0.5%
Number of scaffolds > 100K nt	2	0.0%
Number of scaffolds > 1M nt	0	0.0%
Mean scaffold size	405	
Median scaffold size	41	
N50 scaffold length	4601	
L50 scaffold count	7832	
NG50 scaffold length	5192	
LG50 scaffold count	6446	
N50 scaffold - NG50 scaffold length difference	591	
scaffold %A	29.80	
scaffold %C	20.22	
scaffold %G	20.21	
scaffold %T	29.77	
scaffold %N	0.00	
scaffold %non-ACGTN	0.00	
Number of scaffold non-ACGTN nt	0	
Percentage of assembly in scaffolded contigs	0.0%	
Percentage of assembly in un scaffolded contigs	100.0%	
Average number of contigs per scaffold	1.0	
Average length of break (>25 Ns) between contigs in scaffold	0	
Number of contigs	311185	
Number of contigs in scaffolds	0	
Number of contigs not in scaffolds	311185	
Total size of contigs	126049808	
Longest contig	156082	
Shortest contig	31	
Number of contigs > 500 nt	30949	9.9%
Number of contigs > 1K nt	25660	8.2%
Number of contigs > 10K nt	1443	0.5%
Number of contigs > 100K nt	2	0.0%
Number of contigs > 1M nt	0	0.0%
Mean contig size	405	
Median contig size	41	
N50 contig length	4601	
L50 contig count	7832	
NG50 contig length	5192	
LG50 contig count	6446	
N50 contig - NG50 contig length difference	591	
contig %A	29.80	
contig %C	20.22	
contig %G	20.21	
contig %T	29.77	
contig %N	0.00	
contig %non-ACGTN	0.00	
Number of contig non-ACGTN nt	0	

Appendix 2