Lisa Sudmeier (Revision 2)

Finishing *Drosophila mojavensis* Fosmid Clone 485-B16

Abstract

With the sequences of so many organisms' genomes finished and available for use, comparative genomics has become a rapidly growing scientific discipline. In the spring of 2007, we plan to use comparative genomics to study the 4th chromosome (the 'dot' chromosome) of *Drosophila* by comparing the sequence of the *Drosophila mojavensis* dot chromosome to the sequences of the *Drosophila melanogaster* and *Drosophila virilis* dot chromosomes. Although this chromosome appears to be mainly heterochromatic in *D. melanogaster* and euchromatic in *D. virilis*, Slawson et. al showed that it contains genes. We hope to learn about the *D. mojavensis* dot chromosome from sequence comparisons, but first the sequence of the *D. mojavensis* dot chromosome must be finished. This paper describes the first step in the comparative genomics work to be accomplished in the spring of 2007, the finishing of a *D. mojavensis* dot chromosome fosmid to mouse standard.

Beginning

When I first received the Consed assembly of fosmid 485-B16, four separate contigs appeared in Assembly View, and two more were present in the contig list of the Consed Main Window. However, the two that were not in Assembly View, contigs 1 and 2, contained only two and three reads each, while the smallest number of reads for any contig present in Assembly View was 44. Therefore, contigs 1 and 2 did not have enough coverage to be part of the assembly. Furthermore, usually contigs that are important for final assembly of a fosmid are over 2kb, but contigs 1 and 2 are only 1516 and 934 base pairs, respectively. Because of this, contigs 1 and 2 were ignored in the finishing process. Therefore, to successfully finish this fosmid, contigs 3-6 would need to be arranged in the correct order and orientation with the fosmid ends identified. The gaps between them would then need to be closed so the contigs could be joined and regions of low quality sequence would need to be resolved.

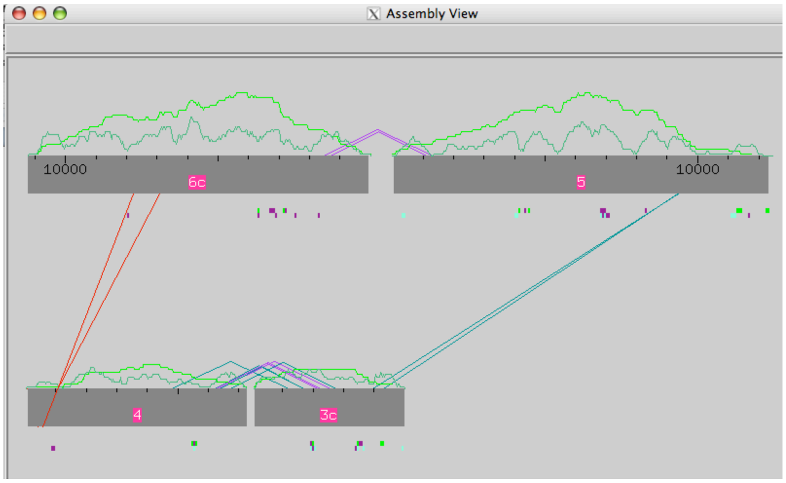
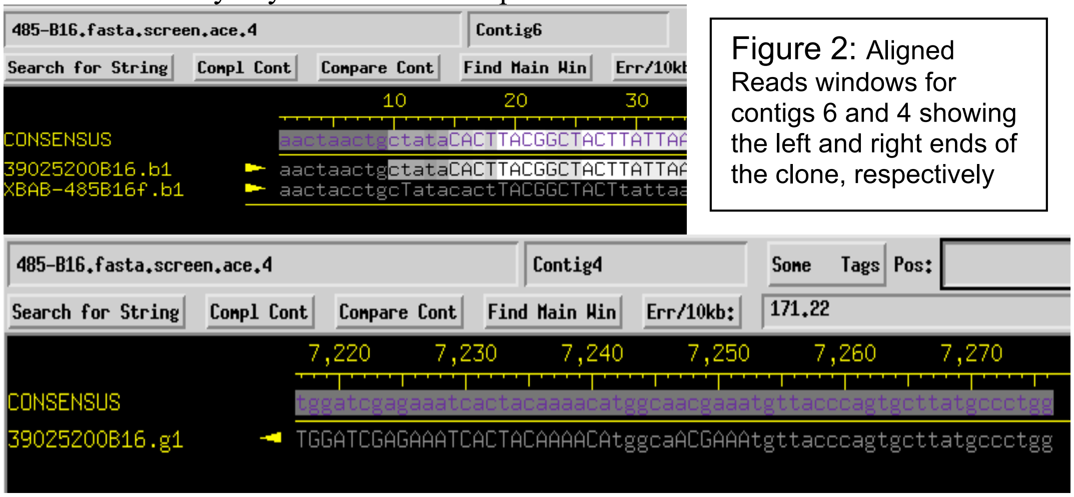


Figure 1: Assembly View of initial assembly

To begin the process of finishing this fosmid, I first found the ends of the fosmid by searching for reads containing ‘390252.’ The reactions designed to cover the ends of the clone, into the vector, were labeled 39025200B16.b1 and g1. The results of this search suggested that the ends of contigs 4 and 6 were the clone ends. However, the sequencing reactions had not extended far enough to sequence any of the vector, so I was unable to identify any actual vector sequence.



```
                                   10        20        30
CONSENSUS                  aactaactgctataCACTTACGGCTACTTATTAA
39025200B16.b1          ►  aactaactgctataCACTTACGGCTACTTATTAA
XBAB-485B16f.b1         ►  aactacctgcTatacactTACGGCTACTtattaa
```

```
                      7,220     7,230     7,240     7,250     7,260     7,270
CONSENSUS             tggatcgagaaatcactacaaaacatggcaacgaaatgttacccagtgcttatgccctgg
39025200B16.g1     ◄  TGGATCGAGAAATCACTACAAAACAtggcaACGAAAtgttacccagtgcttatgccctgg
```

Figure 2: Aligned Reads windows for contigs 6 and 4 showing the left and right ends of the clone, respectively

The original assembly also displayed contigs 6 and 3 in complemented orientation (shown by the ‘c’ after their number in Assembly View), meaning that they were shown in the 3’-5’ direction while the other two contigs were in the forward 5’-3’ orientation. I wanted all of the contigs to be in the same orientation (5’-3’) for assembly, so I reoriented contigs 3 and 6 using the ‘contig rearrangement’ button. The presence of forward-reverse paired reads between contigs 6 and 5 and between contigs 3 and 4, in addition to the knowledge of where the vector ends were located, led me to conclude that the contigs should be arranged in the following order: 6, 5, 3, 4. To obtain this order, I flipped the scaffold for contigs 3 and 4, resulting in the Assembly View shown below.

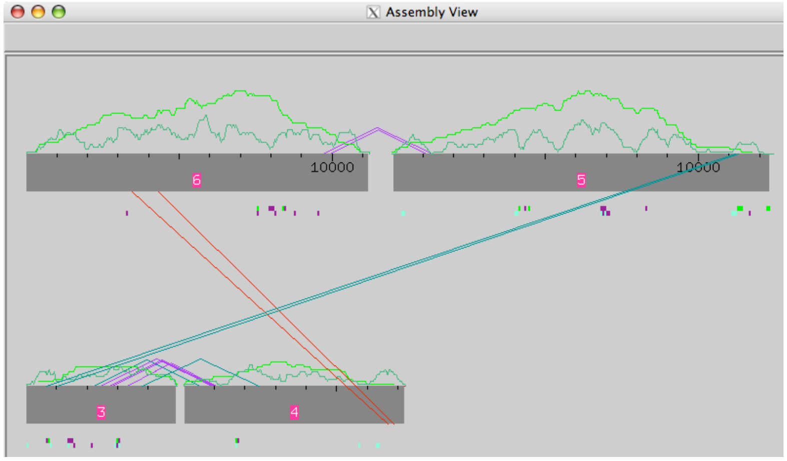


Figure 3: Arrangement of contigs after reorientation

Closing Gaps

The next goal I set out to accomplish to finish my fosmid to mouse standard was to close the gaps between the contigs. My first attempt used the 'search for string' command to try to find sequence matches at the ends of the contigs flanking the gaps. This, however, was unsuccessful, and I realized that I needed more data to cover the gaps. I planned to design oligonucleotides for reactions off the ends of each contig in the direction of its neighboring contig. For the first round of reactions, I was able to order reactions off both ends of contig 5 as well as off the ends of 6 and 3 that border contig 5. Because I knew these were difficult regions to sequence (since they didn't sequence well in any of the reactions that had already been done), I chose to use 4:1 chemistry in hopes that the combined benefits of both dGTP and Big Dye chemistries would be able to resolve these regions. The gap between contigs 3 and 4, however, was problematic. The presence of many repeats at the end of each contig flanking the gap made it very difficult to find a primer on contig 3 that did not match anywhere on contig 4. I decided that it would be best to design a PCR reaction to generate a template of the region spanning the gap, and then sequence this fragment. By amplifying this region, I could be sure that the primer would not sequence a different region flanked by the repeat on contig 3 or 4 because the amplified region would only contain one copy of the repeat, which would then be used to prime the sequencing reaction. Because I thought this region could only be resolved with a PCR-generated template, I did not design primers for sequencing reactions to close this gap in the first round.

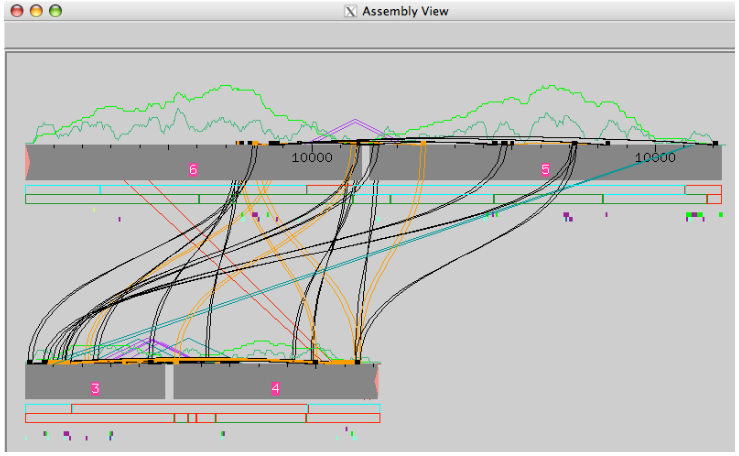
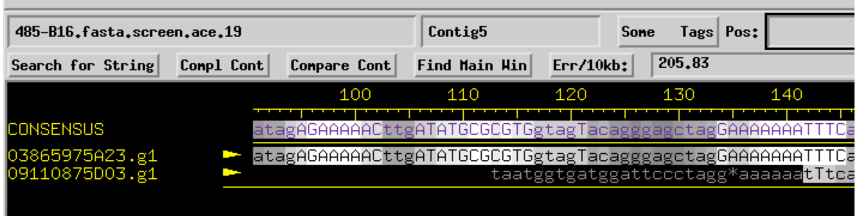


Figure 4: Assembly View with crossmatch to show the prevalence of repeats, especially those shared between contigs 3 and 4

When the data was available from the first round of sequencing reactions, I ran Phred/Phrap on my assembly, which allowed the reads from the reactions to be automatically incorporated into the assembly. The reactions to span the 6-5 gap provided me with a significant amount of data, but not enough to close the gap. Therefore, I designed another set of primers on the new ends of both contigs 6 and 5 for the second (and final) round of reactions. The reactions off the ends of contigs 5 and 3 designed to close the 5-3 gap were unsuccessful and did not provide me with any additional data. After analyzing the 5-3 gap ends with the help of a professional finisher, it was determined that the presence of repeats and a region of high ‘A’ concentration could be making this region hard to resolve. Therefore, I again decided that a PCR reaction to generate a specific template would be the best strategy to close this gap.



```
                    100       110       120       130       140
CONSENSUS       atagAGAAAAACttgATATGCGCGTGgtagTacagggagctagGAAAAAAATTTCa
03865975A23.g1  atagAGAAAAACttgATATGCGCGTGgtagTacagggagctagGAAAAAAATTTCa
09110875D03.g1                  taatggtgatggattccctagg*aaaaaatTtca
```

Figure 5: ‘A’ repeats at the end of contig 5

Due to a technical difficulty, what would have been my second of three reaction orders was not successfully submitted through the pipeline. Therefore, I called all of my reactions in two orders. Furthermore, a miscommunication and the slowness of the PCR pipeline prevented me from ordering the PCR reactions to close the 5-3 and 3-4 gaps. Consequently, for my second and final round of sequencing reactions, I designed multiple primers on the ends of each contig flanking a gap for sequencing reactions, even though it was unlikely that they would work. I ordered these reactions with both BigDye and dGTP chemistries to lessen the chance that a reaction would fail just because of a chemistry problem.

After data was obtained from this second round of reactions I again ran Phred/Phrap on my assembly. Data from these reactions succeeded in closing the gap between contigs 6 and 5, but no additional sequence was generated off the ends of the contigs into the 5-3 and 3-4 gaps. The primers used for these reactions, which were chosen with relaxed primer picking preferences, were so far from the ends of the contigs that it was not surprising that they did not provide any more sequence off the contig ends. Therefore, without data from PCR products, two of the three original gaps remain. I did, however, design PCR primers for the professional finishers to use when completing this project.

I spent a great deal of time analyzing the restriction digest data after ordering the first round of reactions. Looking back, it would have made more sense to look at this data earlier so that I would have known the approximate size of the gaps when I started trying to close them. However, in my project, the presence of repeats suggested the need for a PCR-generated template in these regions, so inverting the order of action for this project did not affect the final outcome. The restriction digest data after the first round of reactions showed the gap between contigs 6 and 5 to be very small, as is evident by the similarity in size between the in-silico and real digest fragments in the figure below. Therefore, I hoped that another round of reactions would cover the region. However, the situation wasn't as hopeful for the other two gaps. For the gap between contigs 3 and 4, the difference between the real and in-silico fragment sizes suggested that the gap was around 500-700 base pairs. Since it was difficult to find good primers in the region, and most of them were 700bp or more from the gap, it appeared unlikely that direct sequencing reactions would be able to cover the gap. A PCR template is clearly the best choice for resolving this region. The largest of the three gaps, however, was that between contigs 5 and 3. Looking at the restriction digests, it was difficult to estimate, but when the fragments corresponding to the end of contig 5 and beginning of contig 3 were compared to each other, the gap appeared to be around a kb in size. As with the 3-4 gap, good primers close to the ends of the contigs would make covering the region simpler (a good sequencing reaction can read up to 700 base pairs). However, lack of these ideal primers again leads me to conclude that a PCR template would be the best option for finishing this gap.

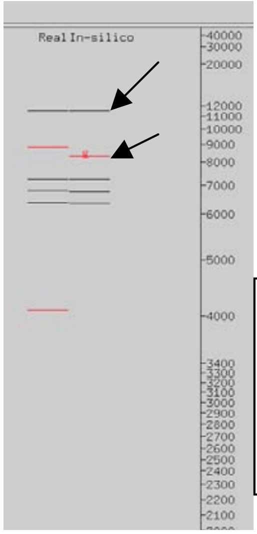
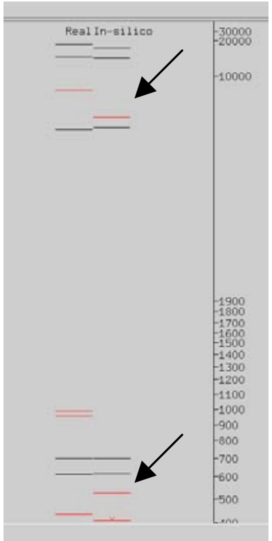


Figure 6: Restriction Digest Data

← **EcoRV**
Top arrow 6-5 gap.
Bottom arrow 3-4 gap (~500bp).

SacI →
Top arrow: fragment off left end of contig 3
Bottom arrow: fragment off right end of contig 5



Low Quality and other Regions of Interest

I also spent time trying to resolve low quality, single stranded, and single subclone regions. After the first round of reactions, I went through the new reads I'd obtained and decided to reorder some using the same primers with different chemistries, as in the case shown below. The primer for this reaction clearly annealed well, as shown by the strong peaks for the first few bases, but the sequencing reaction failed to generate much data. For the first round of reactions, I ordered Big Dye chemistry to cover all low quality regions. When reusing primers, I chose dGTP chemistry in hopes that it would be able to overcome the problems that prevented Big Dye from sequencing effectively. In the second round, I ordered many reactions with both chemistries anticipating that if one didn't work the other might. For many low quality regions I had to relax the primer picking preferences in order to obtain primers reasonably close to the region of interest. By relaxing these preferences, I allowed a greater range of melting temperatures and a higher percent match elsewhere. I performed 'search for string' with the first 10 bases of each primer to ensure that it didn't have an exact match somewhere else in the assembly.

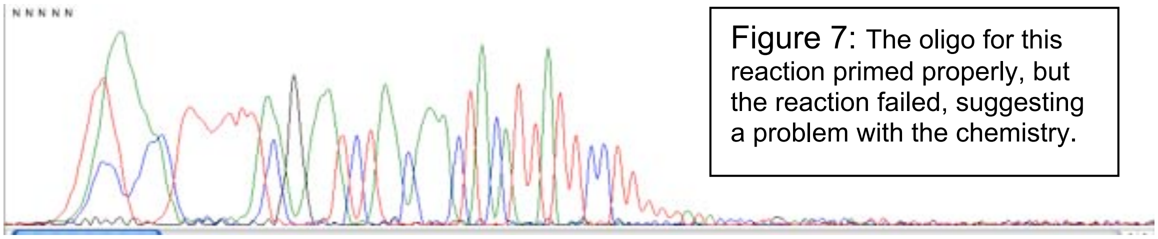
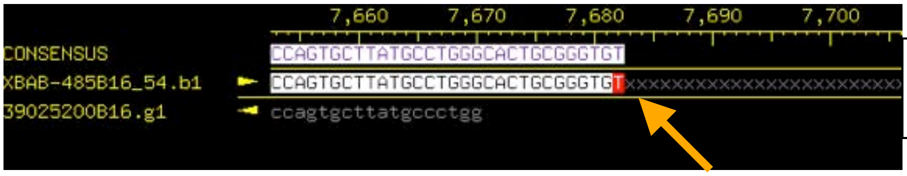


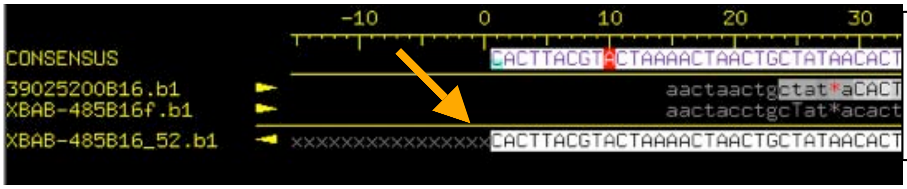
Figure 7: The oligo for this reaction primed properly, but the reaction failed, suggesting a problem with the chemistry.

I also designed primers for sequencing reactions to cover the ends of contigs 6 and 4 that went into the vector. After the second round of reactions, I had data showing clear vector ends, which I tagged to mark the ends of my clone. Consed recognized the vector sequence and called the bases x's. I did 'search for string' to look for any other bases in the assembly that Consed may have called 'x,' but they were only present in the vector sequence.



```
                          7,660     7,670     7,680     7,690     7,700
CONSENSUS                 CCAGTGCTTATGCCTGGGCACTGCGGGTGT
XBAB-485B16_54.b1     ►   CCAGTGCTTATGCCTGGGCACTGCGGGTGTxxxxxxxxxxxxxxxxxxxxxxx
39025200B16.g1        ◄   ccagtgcttatgccctgg
```

Figure 8: Vector at end of contig 4



```
                            -10         0        10        20        30
CONSENSUS                              CACTTACGTACTAAAACTAACTGCTATAACACT
39025200B16.b1        ►                                   aactaactgctat*aCACT
XBAB-485B16f.b1       ►                                   aactacctgcTat*acact
XBAB-485B16_52.b1     ◄     xxxxxxxxxxxxxxxCACTTACGTACTAAAACTAACTGCTATAACACT
```

Figure 9: Vector at end of contig 6, which became part of 5 when they joined

Mononucleotide regions are areas of concern as well because it's much easier for Phred/Phrap to miscall them, and flanking regions may be affected. I searched for regions containing 15 or more C's or A's in a row, and found two runs of A's. Traces from both regions, however, had high quality peaks both through and flanking the mononucleotide region. I would be concerned about such regions if the traces all looked like the second of the traces shown below because it is not of very high quality.

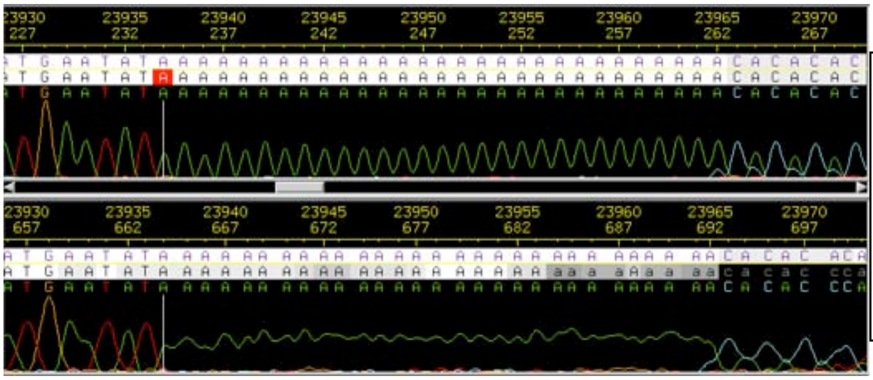
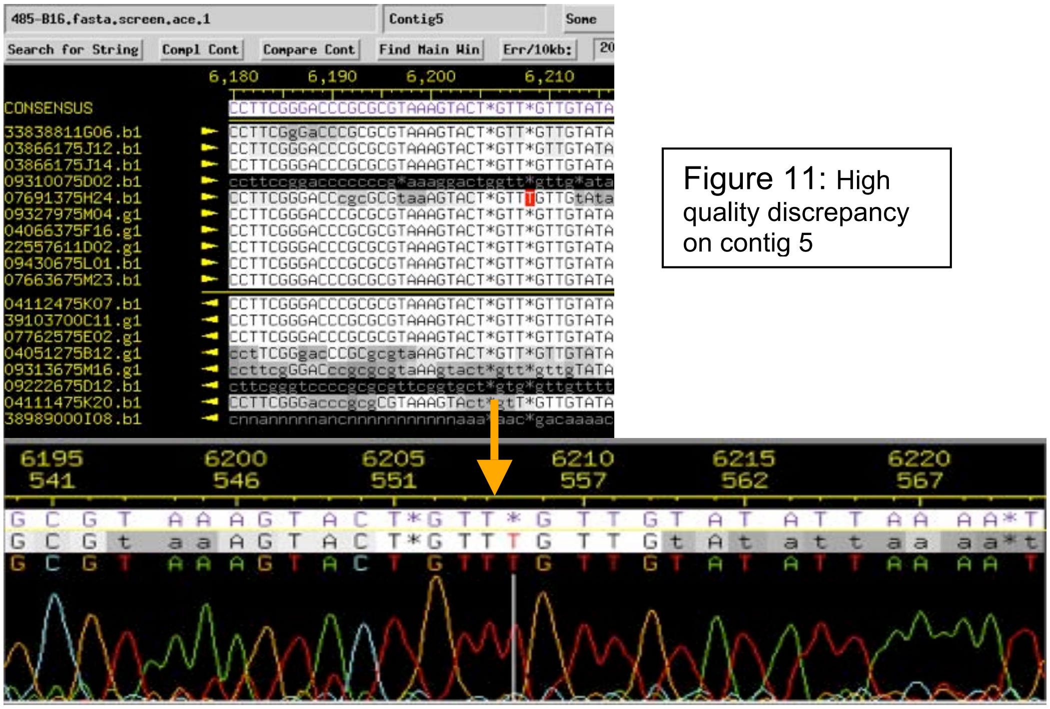


Figure 10: Mononucleotide region on contig 5; *above*, a trace with high confidence; *below*, a low quality trace.

Sometimes Phred/Phrap calls a base with very high confidence that doesn't match the consensus sequence. These 'high quality discrepancies' are important to resolve in an assembly when trying to achieve mouse standard. I found one high quality discrepancy before ordering any reactions, and another after obtaining all of my new reads. The first, shown below, was a triple 'T' peak. Phred/Phrap called 3 T's because of the shape of the peak, but from looking at the distance from the left side of the first peak to the right side of the third, I was able to determine that there was not enough space for more than two full base peaks. Therefore, I changed the third T to a * (pad) so that this read would only show 2 T's in a row, as does the consensus. The second high quality discrepancy was from one of my new reads in which a pad was called at an A position in the consensus sequence. When I looked at the trace, I could see the A peak partially hidden behind a C peak.



```
                   6,180     6,190     6,200     6,210
CONSENSUS          CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
33838811G06.b1  ►  CCTTCGgGaCCCGCGCGTAAAGTACT*GTT*GTTGTATA
03866175J12.b1  ►  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
03866175J14.b1  ►  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
09310075D02.b1  ►  ccttccggaccccccccg*aaaggactggtt*gttg*ata
07691375H24.b1  ►  CCTTCGGGACCcgcGCGtaaAGTACT*GTTTGTTGtAta
09327975M04.g1  ►  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
04066375F16.g1  ►  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
22557611D02.g1  ►  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
09430675L01.b1  ►  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
07663675M23.b1  ►  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
04112475K07.b1  ◄  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
39103700C11.g1  ◄  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
07762575E02.g1  ◄  CCTTCGGGACCCGCGCGTAAAGTACT*GTT*GTTGTATA
04051275B12.g1  ◄  cctTCGGgacCCGCgcgtaAAGTACT*GTT*GTTGTATA
09313675M16.g1  ◄  ccttcgGGACccgcgcgtaAAgtact*gtt*gttgTATA
09222675D12.b1  ◄  cttcgggtccccgcgcgttcggtgct*gtg*gttgtttt
04111475K20.b1  ◄  CCTTCGGGacccgcgCGTAAAGTAct*gtT*GTTGTATA
38989000I08.b1  ◄  cnnannnnnancnnnnnnnnnnnaaa*aac*gacaaaac
```

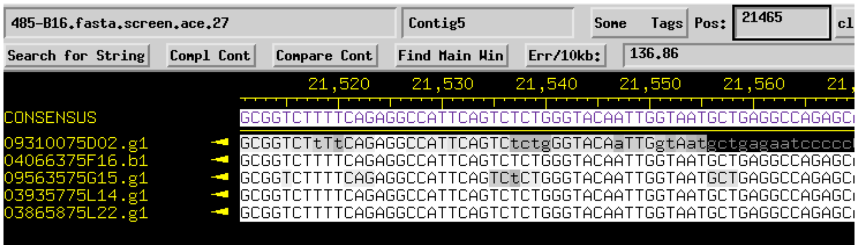
6195 541
6200 546
6205 551
6210 557
6215 562
6220 567
G C G T A A A G T A C T*G T T * G T T G T A T A T T A A A A*T
G C G t a a A G T A C T*G T T T G T T G t A t a t t a a a a*t
G C G T A A A G T A C T G T T T G T T G T A T A T T A A A A T

Figure 11: High quality discrepancy on contig 5

Finally, I searched for bases that hadn't been called and were marked with an 'N.' There was only one of these in the assembly, and by looking at the trace, I could clearly see a 'C' peak. This was the last locus to be resolved without needing additional data. Since many of the reactions called did not generate new data, there are still many low quality and single stranded/single subclone regions in the assembly that need additional sequencing to be resolved.

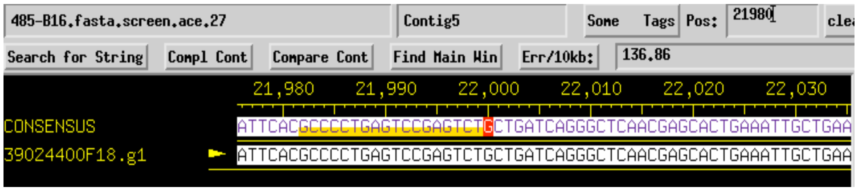
Autofinish

The high frequency of repeats in this assembly made designing primers very difficult. I looked at the primers Autofinish had designed after two rounds of reactions. It would have been better to do this after the first round of reactions, so that I could have used primers from Autofinish that I had not picked myself for the second round of reactions. About half of the primers Autofinish designed were identical to ones I had made. Most of the other primers Autofinish made were in the same region as primers I had made and in the same direction, so I assume they were designed to resolve the same region. There was one single stranded region for which Autofinish had designed a primer, which I never tried to resolve because there were so many high quality reads available (although they were all from the same strand). An additional location where Autofinish designed a primer and I didn't, but should have, was for a single stranded/single chemistry/single subclone region on contig 5. Overlooking this region that Autofinish tried to resolve is one example of why I should have looked at Autofinish earlier. Since so many of my reactions didn't work, it would have been wise to try Autofinish's primers for some of the regions I was having trouble resolving with different primers.



```
                  21,520    21,530    21,540    21,550    21,560    21,
CONSENSUS         GCGGTCTTTTCAGAGGCCATTCAGTCTCTGGGTACAATTGGTAATGCTGAGGCCAGAGC
09310075D02.g1  ◄ GCGGTCTtTtCAGAGGCCATTCAGTCtctgGGTACAaTTGgtAatgctgagaatccccc
04066375F16.b1  ◄ GCGGTCTTTTCAGAGGCCATTCAGTCTCTGGGTACAATTGGTAATGCTGAGGCCAGAGC
09563575G15.g1  ◄ GCGGTCTTTTCAGAGGCCATTCAGTCtCTGGGTACAATTGGTAATGCTGAGGCCAGAGC
03935775L14.g1  ◄ GCGGTCTTTTCAGAGGCCATTCAGTCTCTGGGTACAATTGGTAATGCTGAGGCCAGAGC
03865875L22.g1  ◄ GCGGTCTTTTCAGAGGCCATTCAGTCTCTGGGTACAATTGGTAATGCTGAGGCCAGAGC
```

Figure 12: Autofinish designed a primer to resolve this region, but I didn't try to resolve it because the quality was already >Phred30.



```
                  21,980    21,990    22,000    22,010    22,020    22,030
CONSENSUS         ATTCACGCCCCTGAGTCCGAGTCTGCTGATCAGGGCTCAACGAGCACTGAAATTGCTGAA
39024400F18.g1  ► ATTCACGCCCCTGAGTCCGAGTCTGCTGATCAGGGCTCAACGAGCACTGAAATTGCTGAA
```

Figure 13: I missed this region when designing reactions, but it clearly needs more reactions for higher confidence. I added a 'data needed tag'

Conclusion

Had I been able to obtain PCR templates and order three rounds of reactions, I believe I would have been able to successfully finish my fosmid. However, in addition to missing out on opportunities to order reactions, I did make a number of mistakes that I would try to avoid if I find myself finishing a fosmid again. First, I would look at

Autofinish earlier in the process of finishing. Second, I wouldn't wait so long to consult the results of the restriction digest. Finally, I would spend more time designing reactions to cover the low quality regions. The high number of failed reactions, which should be anticipated, prevented me from adequately covering these regions.

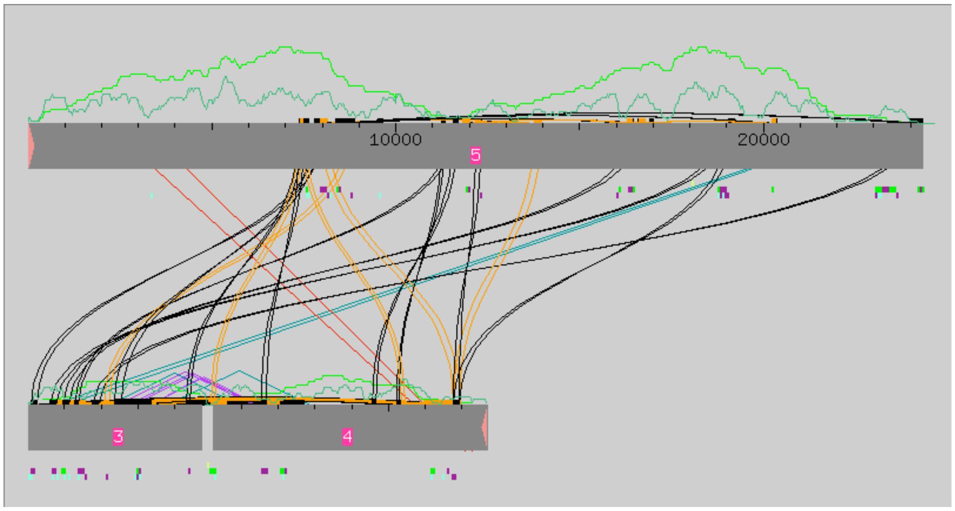
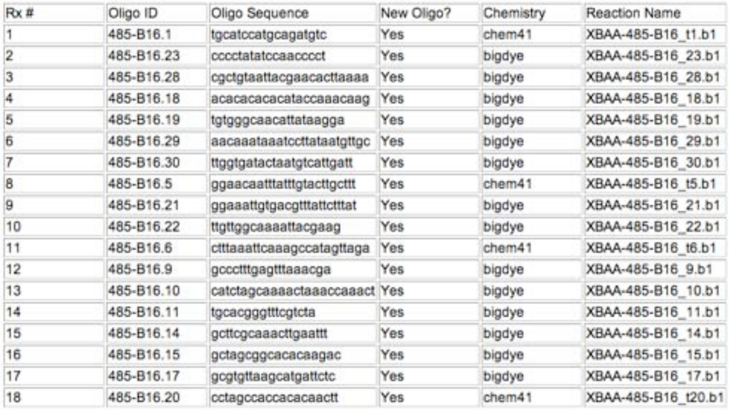


Figure 14: Final assembly in Assembly View

Appendix

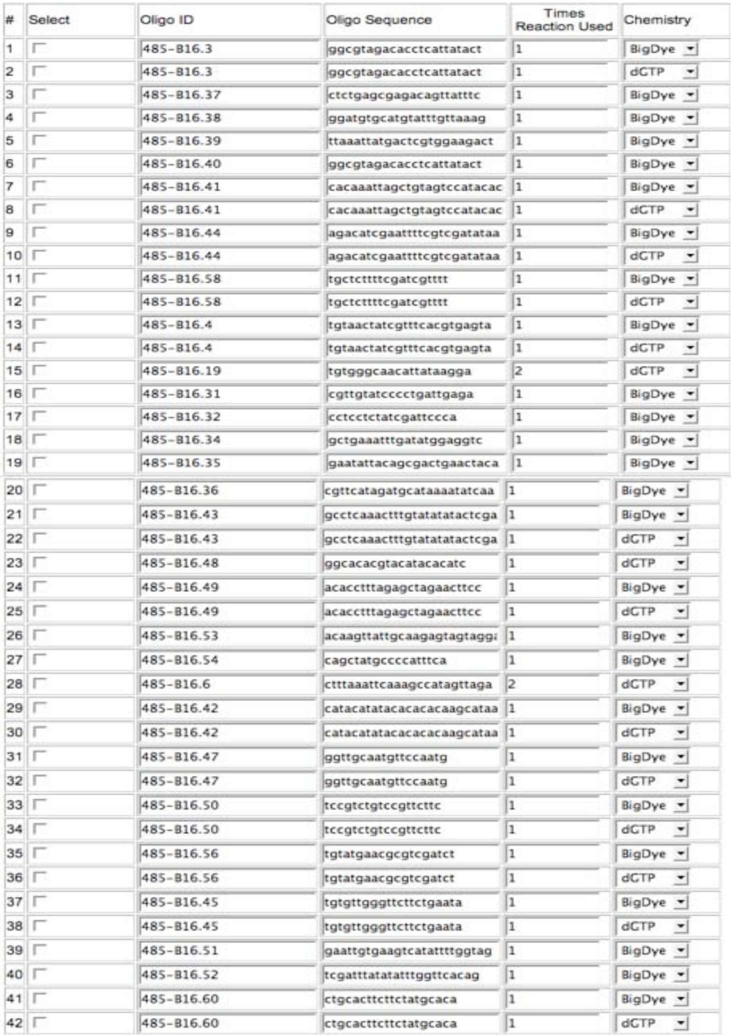
Reactions Ordered

Round 1:



Rx #	Oligo ID	Oligo Sequence	New Oligo?	Chemistry	Reaction Name
1	485-B16.1	tgcatccatgcagatgtc	Yes	chem41	XBAA-485-B16_t1.b1
2	485-B16.23	cccctatatccaacccct	Yes	bigdye	XBAA-485-B16_23.b1
3	485-B16.28	cgctgtaattacgaacacttaaaa	Yes	bigdye	XBAA-485-B16_28.b1
4	485-B16.18	acacacacacataccaaacaag	Yes	bigdye	XBAA-485-B16_18.b1
5	485-B16.19	tgtgggcaacattataagga	Yes	bigdye	XBAA-485-B16_19.b1
6	485-B16.29	aacaaataaatccttataatgttgc	Yes	bigdye	XBAA-485-B16_29.b1
7	485-B16.30	ttggtgatactaatgtcattgatt	Yes	bigdye	XBAA-485-B16_30.b1
8	485-B16.5	ggaacaatttatttgtacttgcttt	Yes	chem41	XBAA-485-B16_t5.b1
9	485-B16.21	ggaaattgtgacgtttattctttat	Yes	bigdye	XBAA-485-B16_21.b1
10	485-B16.22	ttgttggcaaaattacgaag	Yes	bigdye	XBAA-485-B16_22.b1
11	485-B16.6	ctttaaattcaaagccatagttaga	Yes	chem41	XBAA-485-B16_t6.b1
12	485-B16.9	gccctttgagtttaaacga	Yes	bigdye	XBAA-485-B16_9.b1
13	485-B16.10	catctagcaaaactaaaccaaact	Yes	bigdye	XBAA-485-B16_10.b1
14	485-B16.11	tgcacgggtttcgtcta	Yes	bigdye	XBAA-485-B16_11.b1
15	485-B16.14	gcttcgcaaacttgaattt	Yes	bigdye	XBAA-485-B16_14.b1
16	485-B16.15	gctagcggcacacaagac	Yes	bigdye	XBAA-485-B16_15.b1
17	485-B16.17	gcgtgttaagcatgattctc	Yes	bigdye	XBAA-485-B16_17.b1
18	485-B16.20	cctagccaccacacaactt	Yes	chem41	XBAA-485-B16_t20.b1

Round 2:



#	Select	Oligo ID	Oligo Sequence	Times Reaction Used	Chemistry
1	☐	485-B16.3	ggcgtagacacctcattatact	1	BigDye
2	☐	485-B16.3	ggcgtagacacctcattatact	1	dGTP
3	☐	485-B16.37	ctctgagcgagacagttatttc	1	BigDye
4	☐	485-B16.38	ggatgtgcatgtatttgttaaag	1	BigDye
5	☐	485-B16.39	ttaaattatgactcgtggaagact	1	BigDye
6	☐	485-B16.40	ggcgtagacacctcattatact	1	BigDye
7	☐	485-B16.41	cacaaattagctgtagtccatacac	1	BigDye
8	☐	485-B16.41	cacaaattagctgtagtccatacac	1	dGTP
9	☐	485-B16.44	agacatcgaattttcgtcgatataa	1	BigDye
10	☐	485-B16.44	agacatcgaattttcgtcgatataa	1	dGTP
11	☐	485-B16.58	tgctcttttcgatcgtttt	1	BigDye
12	☐	485-B16.58	tgctcttttcgatcgtttt	1	dGTP
13	☐	485-B16.4	tgtaactatcgtttcacgtgagta	1	BigDye
14	☐	485-B16.4	tgtaactatcgtttcacgtgagta	1	dGTP
15	☐	485-B16.19	tgtgggcaacattataagga	2	dGTP
16	☐	485-B16.31	cgttgtatcccctgattgaga	1	BigDye
17	☐	485-B16.32	cctcctctatcgattccca	1	BigDye
18	☐	485-B16.34	gctgaaatttgatatggaggtc	1	BigDye
19	☐	485-B16.35	gaatattacagcgactgaactaca	1	BigDye
20	☐	485-B16.36	cgttcatagatgcataaaatatcaa	1	BigDye
21	☐	485-B16.43	gcctcaaactttgtatatatactcga	1	BigDye
22	☐	485-B16.43	gcctcaaactttgtatatatactcga	1	dGTP
23	☐	485-B16.48	ggcacacgtacatacacatc	1	dGTP
24	☐	485-B16.49	acacctttagagctagaacttcc	1	BigDye
25	☐	485-B16.49	acacctttagagctagaacttcc	1	dGTP
26	☐	485-B16.53	acaagttattgcaagagtagtagga	1	BigDye
27	☐	485-B16.54	cagctatgccccatttca	1	BigDye
28	☐	485-B16.6	ctttaaattcaaagccatagttaga	2	dGTP
29	☐	485-B16.42	catacatatacacacacaagcataa	1	BigDye
30	☐	485-B16.42	catacatatacacacacaagcataa	1	dGTP
31	☐	485-B16.47	ggttgcaatgttccaatg	1	BigDye
32	☐	485-B16.47	ggttgcaatgttccaatg	1	dGTP
33	☐	485-B16.50	tccgtctgtccgttcttc	1	BigDye
34	☐	485-B16.50	tccgtctgtccgttcttc	1	dGTP
35	☐	485-B16.56	tgtatgaacgcgtcgatct	1	BigDye
36	☐	485-B16.56	tgtatgaacgcgtcgatct	1	dGTP
37	☐	485-B16.45	tgtgttgggttcttctgaata	1	BigDye
38	☐	485-B16.45	tgtgttgggttcttctgaata	1	dGTP
39	☐	485-B16.51	gaattgtgaagtcatattttggtag	1	BigDye
40	☐	485-B16.52	tcgatttatatatttggttcacag	1	BigDye
41	☐	485-B16.60	ctgcacttcttctatgcaca	1	BigDye
42	☐	485-B16.60	ctgcacttcttctatgcaca	1	dGTP