FINISHING MY FOSMID
XAAA113
Andrew Nett

The assembly of fosmid XAAA113 initially consisted of three contigs with two gaps (Fig 1). This assembly was based on the sequencing reactions of subclones set up in class. Reduction of the assembly to one contig easily occurred upon incorporation of all other available reads.

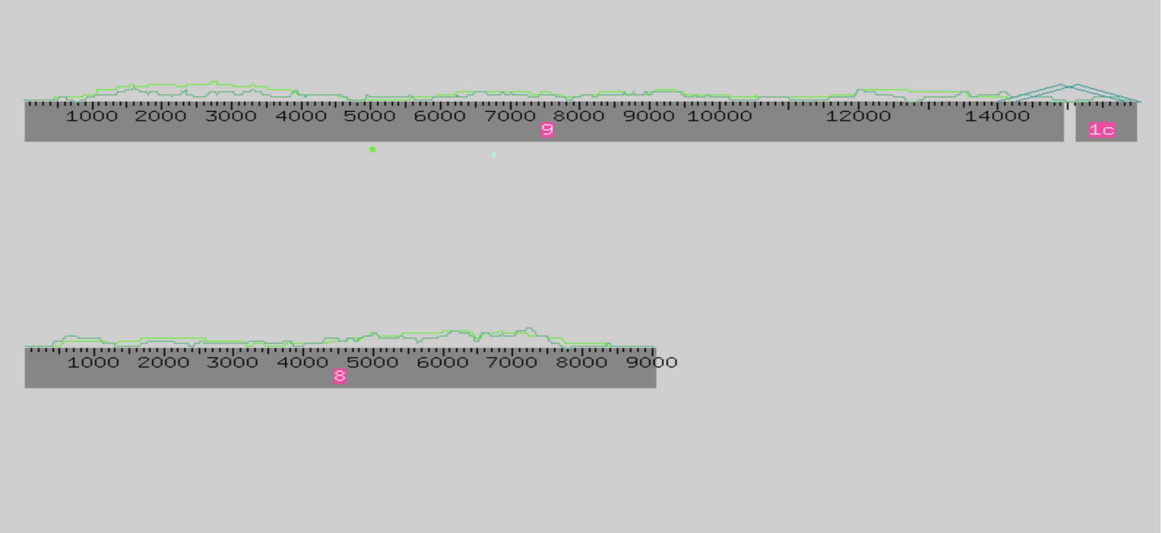


Fig 1. Initial assembly view based on phredphrap of just my reads.

After running phredphrap using all available reads of my fosmid clone, my assembly became one large contiguous sequence (barring contigs smaller than 2 kb). This contig was 37,265 base pairs long – close to the expected 40 kb fosmid size (Fig 2).

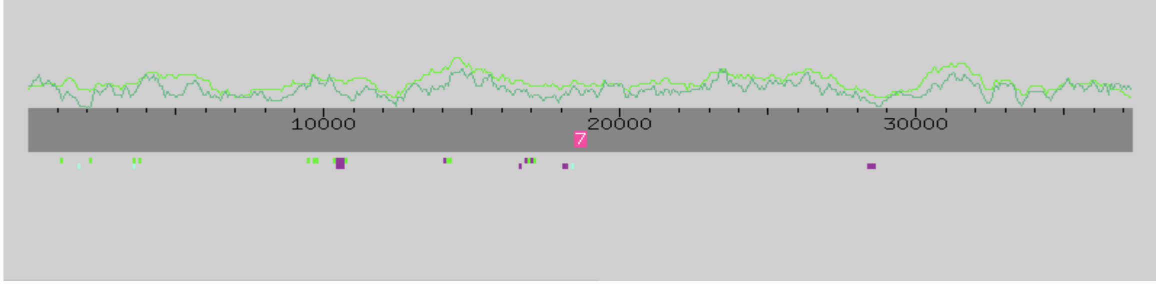
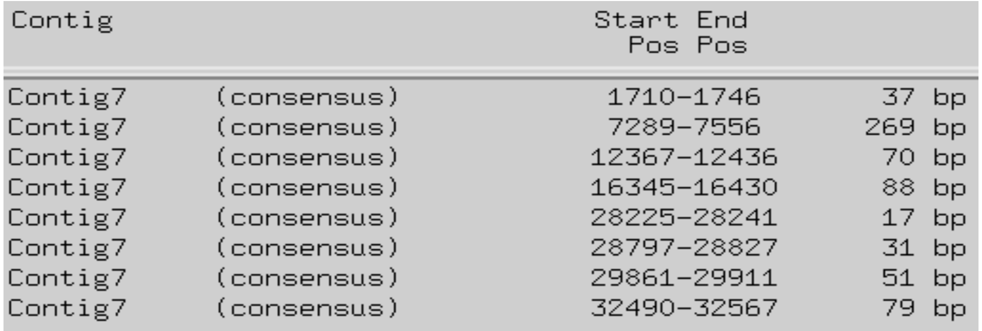


Fig 2. Assembly view based on all available reads.

ROUND I FINISHING

As no sequence gaps were present, I began the finishing process by navigating through regions of low base quality, single stranded chemistry, sequence discrepancies, single subclones, and unaligned high quality sequence. Though already existing as one contig, the assembly did contain certain problem areas: regions of single stranded chemistry, regions of low base quality, and high quality discrepancies in sequence.

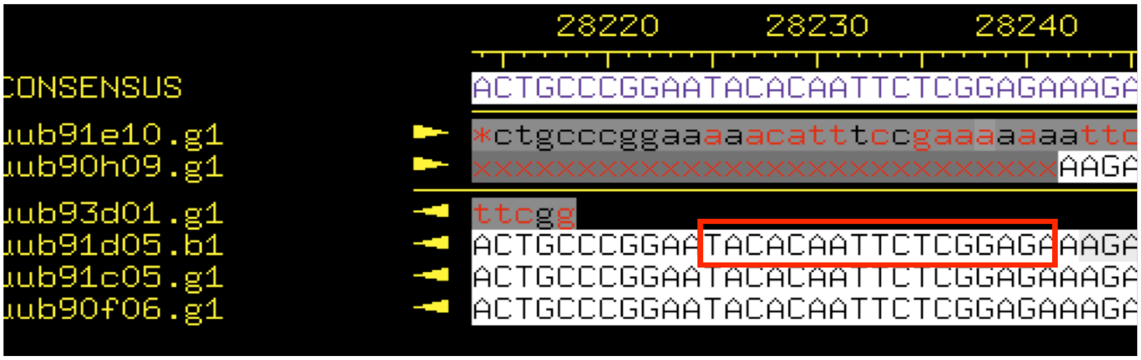
Single Stranded Chemistry



Contig		Start End Pos Pos	
Contig7	(consensus)	1710-1746	37 bp
Contig7	(consensus)	7289-7556	269 bp
Contig7	(consensus)	12367-12436	70 bp
Contig7	(consensus)	16345-16430	88 bp
Contig7	(consensus)	28225-28241	17 bp
Contig7	(consensus)	28797-28827	31 bp
Contig7	(consensus)	29861-29911	51 bp
Contig7	(consensus)	32490-32567	79 bp

Table 1. Regions of single stranded chemistry.

Although there were regions consisting only of data from single stranded chemistry reactions (Table 1), these regions all contained at least one read with a quality value of 30 or higher at each base. As such, I tagged these regions as phred30, and ordered no new reactions to cover them. For example, figure 3 shows sequence data for the single stranded chemistry region from bases 28225-28241. In this region, the quality values of the uub91d05.b1 subclone sequence range from 48 to 57 – well above the phred30 value, which serves as a threshold for concern.



```
                           28220     28230     28240
CONSENSUS                  ACTGCCCGGAATACACAATTCTCGGAGAAAGA
uub91e10.g1                *ctgcccggaaaaacatttccgaaaaaaattc
uub90h09.g1                XXXXXXXXXXXXXXXXXXXXXXXXXXXXAAGA
uub93d01.g1                ttcgg
uub91d05.b1                ACTGCCCGGAATACACAATTCTCGGAGAAAGA
uub91c05.g1                ACTGCCCGGAATACACAATTCTCGGAGAAAGA
uub90f06.g1                ACTGCCCGGAATACACAATTCTCGGAGAAAGA
```

Fig 3. Single stranded chemistry region. Quality > 30.

Regions of single stranded chemistry fortunately did not require additional coverage since they contained reads of sufficient quality. Not all of the trace data, however, warranted confidence in the contig sequence.

Low Base Quality

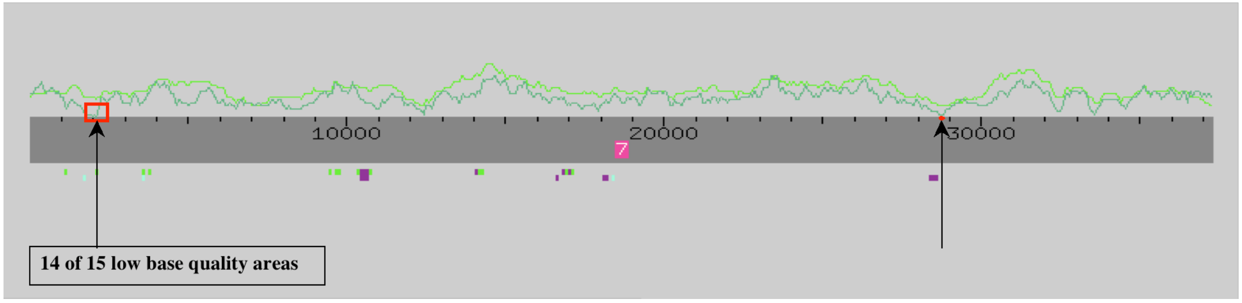
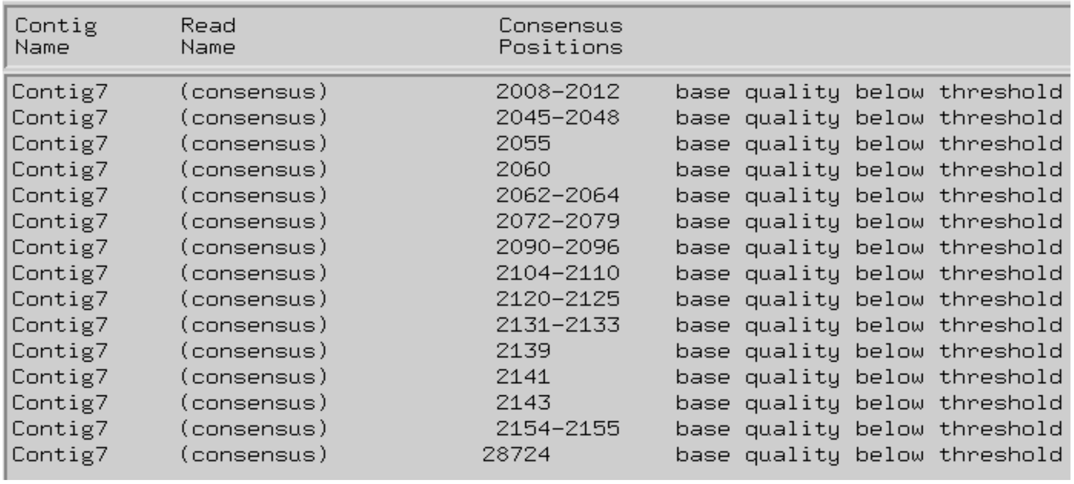


Fig 4. Low base quality regions.



Contig Name	Read Name	Consensus Positions	
Contig7	(consensus)	2008-2012	base quality below threshold
Contig7	(consensus)	2045-2048	base quality below threshold
Contig7	(consensus)	2055	base quality below threshold
Contig7	(consensus)	2060	base quality below threshold
Contig7	(consensus)	2062-2064	base quality below threshold
Contig7	(consensus)	2072-2079	base quality below threshold
Contig7	(consensus)	2090-2096	base quality below threshold
Contig7	(consensus)	2104-2110	base quality below threshold
Contig7	(consensus)	2120-2125	base quality below threshold
Contig7	(consensus)	2131-2133	base quality below threshold
Contig7	(consensus)	2139	base quality below threshold
Contig7	(consensus)	2141	base quality below threshold
Contig7	(consensus)	2143	base quality below threshold
Contig7	(consensus)	2154-2155	base quality below threshold
Contig7	(consensus)	28724	base quality below threshold

Table 2. Low base quality regions.

The initial assembly contained several low base quality regions (Table 2). I could possibly have discerned a consensus base sequence in some of these areas by looking at sequence traces. Upon examination, for example, one can convince themselves that low quality region 2062-2064 almost certainly has a GTG sequence. Further reads are not required for sequence determination (Figure 5). However, since my assembly only had one contig to begin with, I received instruction to order additional reactions to cover problem areas that perhaps would not normally necessitate additional coverage.

Conveniently, all but one of the low quality regions were clumped together within a 150 bp stretch between bases 2008 and 2155 (Table 2, Figure 4, previous page). Hence, I could cover all low quality areas by ordering only two reactions (Table 3, next page) – one to add coverage to the area from bases 2008-2155, and one to improve consensus sequence quality at base 28724. The reactions ordered used standard big dye chemistry and oligo sequences determined through the Consed program.

As a measure of precaution, I compared the reactions I ordered to a list that the Autofinish program would have ordered if given the same sequence data. Autofinish's hypothetical list also only consisted of two reactions. One of these reactions covered the same low quality region as my Oligo I reaction. Autofinish, however, did not call for a reaction to add coverage at base 28724 as I did with my Oligo II reaction. Additionally, the second Autofinish reaction would have added coverage to a low quality region at bases 33592-33594. I had not initially called a reaction

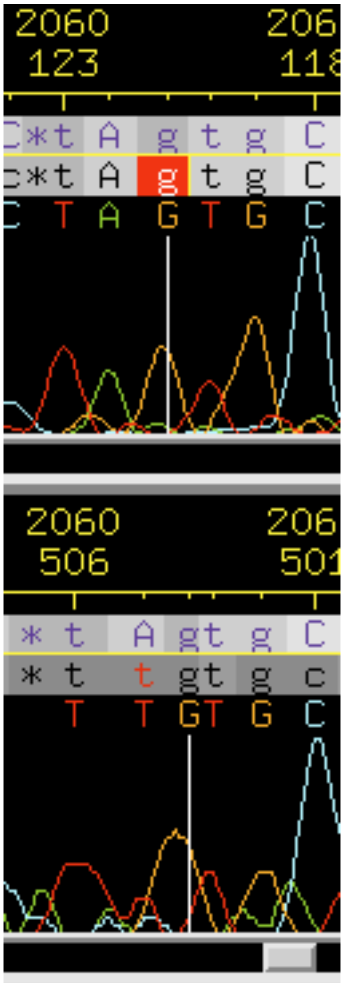
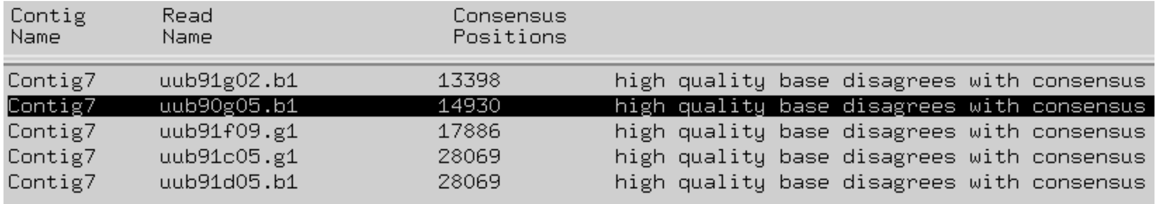


Fig 5. Trace data of low quality sequence.

for this area, nor did Consed navigation of low quality consensus sequences even register this region. I did order the reaction after seeing the Autofinish list.

Table 3. Ordered reactions and comparison to Autofinish.

	Oligo	Sequence	Template	Low Quality Region Covered
I ordered	I	ggcactcgtaaaagtaaacaag	uub90f04.g1	2008-2155
I ordered	II	cgctccattcataattaaaagtc	uub90c02.b1	28724
autofinish ordered	III	ttgttacatgaatttgaactacttg	uub93d06.b1	33592-33594
autofinish ordered	redundant	accgcgccttttcca	uub91h08	2008-2155



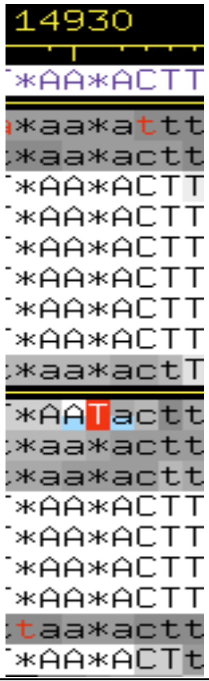
```
Contig    Read                 Consensus
Name      Name                 Positions

Contig7   uub91g02.b1          13398      high quality base disagrees with consensus
Contig7   uub90g05.b1          14930      high quality base disagrees with consensus
Contig7   uub91f09.g1          17886      high quality base disagrees with consensus
Contig7   uub91c05.g1          28069      high quality base disagrees with consensus
Contig7   uub91d05.b1          28069      high quality base disagrees with consensus
```

Table 4. High quality discrepancies in initial assembly.

High Quality Discrepancies

The final problem with my initial assembly was the presence of five high quality discrepancies between individual subclone sequences and the consensus sequence. By checking the trace data and its spacing at base 13398, I determined that the discrepancy was only a pad placed in one of the reads due to a trace compression.

The discrepancies existing at positions 14930 and 17886 also were not cause for concern, as they were most likely the result of capillary errors. Position 14930's discrepancy became much easier to ignore when contrasted to the ten other high quality reads available, all of which contain a pad in place of the incongruent read's T (Figure 6). Examination of available traces at this position supported ignoring the discrepant T as a capillary error (Figure 7). In two of the reads shown, there is no hint of a red T trace at the position in question. Furthermore, the spacing of the AATA



```
14930
*AA*ACTT
*aa*attt
*aa*actt
*AA*ACTT
*AA*ACTT
*AA*ACTT
*AA*ACTT
*AA*ACTT
*AA*ACTT
*aa*actT
*AATactt
*aa*actt
*aa*actt
*AA*ACTT
*AA*ACTT
*AA*ACTT
*AA*ACTT
taa*actt
*AA*ACTt
```

Fig 6. Discrepancy at base 14930.

sequence from bases 14929 to 14932 is consistent with the spacing of three bases, suggesting that one of the bases (the discrepant T) does not belong.

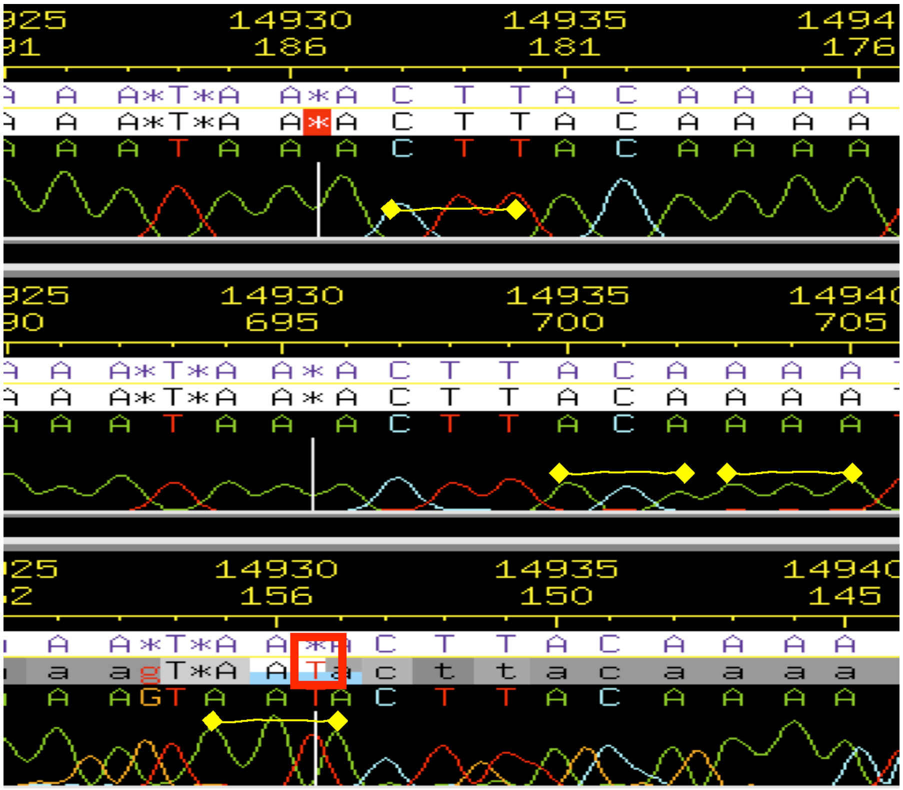
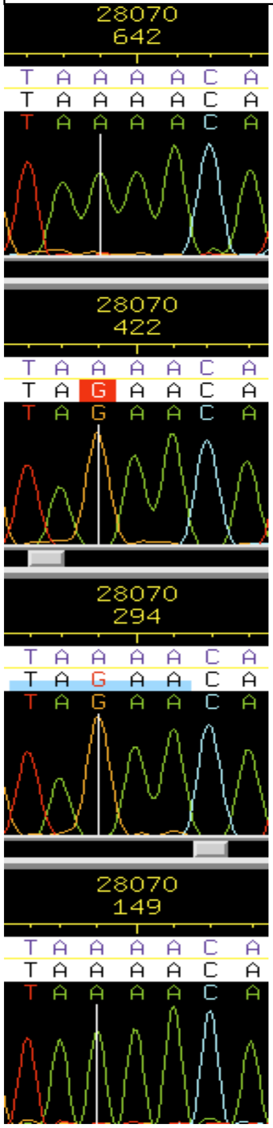


Fig 7. Discrepancy at position 14930.

Fig 8. Discrepancy at position 28069.



The other two discrepancies occurred at the same base location, position 28069. Here, two of five reads (all with a quality value ≥30) indicated a G to be present, while the other three designated the base as an A. Since the discrepant G existed in more than one read, I did not ignore it as merely a capillary error. Examination of the traces did not resolve the inconsistency, but instead solidified the G base calls as real. The traces were clear and had consistent spacing (Figure 8).

Presence of this sequence discrepancy suggested that the reads were perhaps misplaced by Consed. If the sequence in question was part of a duplication, this incongruency possibly resulted from a duplication that Consed assembled together, incorrectly overlapping some of the nearly identical reads as the same sequence. To check this possibility, I searched all available reads for the 11-base sequence centered around base 28069. This search was done for both an A and a G as the center base. If the sequence discrepancy was due to a misplaced read, one would have expected one of these 11-base sequences to match reads elsewhere. No matches were found, however, which suggested that a duplicated sequence did not exist. Additionally, examination of the relevant forward/reverse read pairs showed that the sequence discrepancy did not occur in a repeat region, and further, that the read pairs containing this discrepant sequence (and all others) were consistent (Fig 9). In other words, the read pairs were located at a distance to each other that was not larger than the expected read size. Thus, the sequence discrepancy was probably not due to a duplication overlap. To be certain, however, I compared real and *in silico* digests of my fosmid clone. If Consed had completely overlapped a duplicated sequence, conflicting real and *in silico* digest results might indicate such a misassembly. Restriction digest results were, in fact, compatible (Figure 10). Although the EcoRV digest illustrated an inconsistency between the real and *in silico* digest at a size of 1210, one did not actually exist.

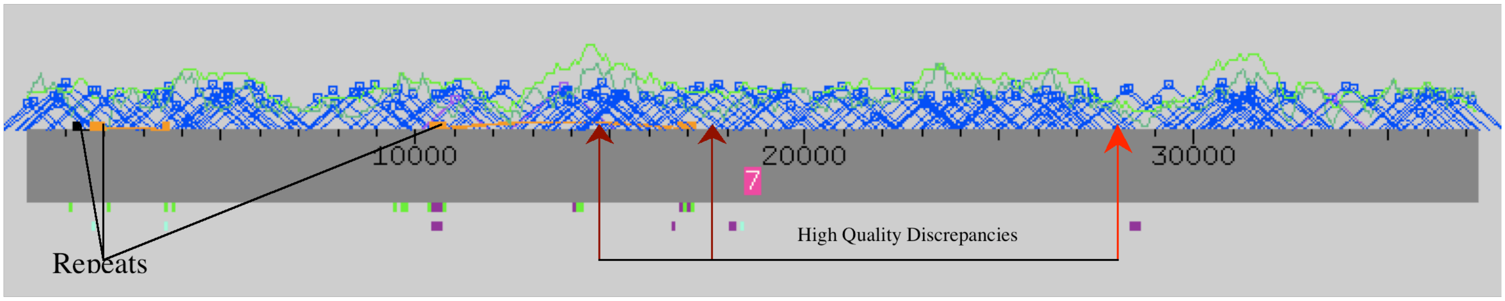


Fig 9. Contig assembly view. Read pairs consistent. No repeats close to base 28069.

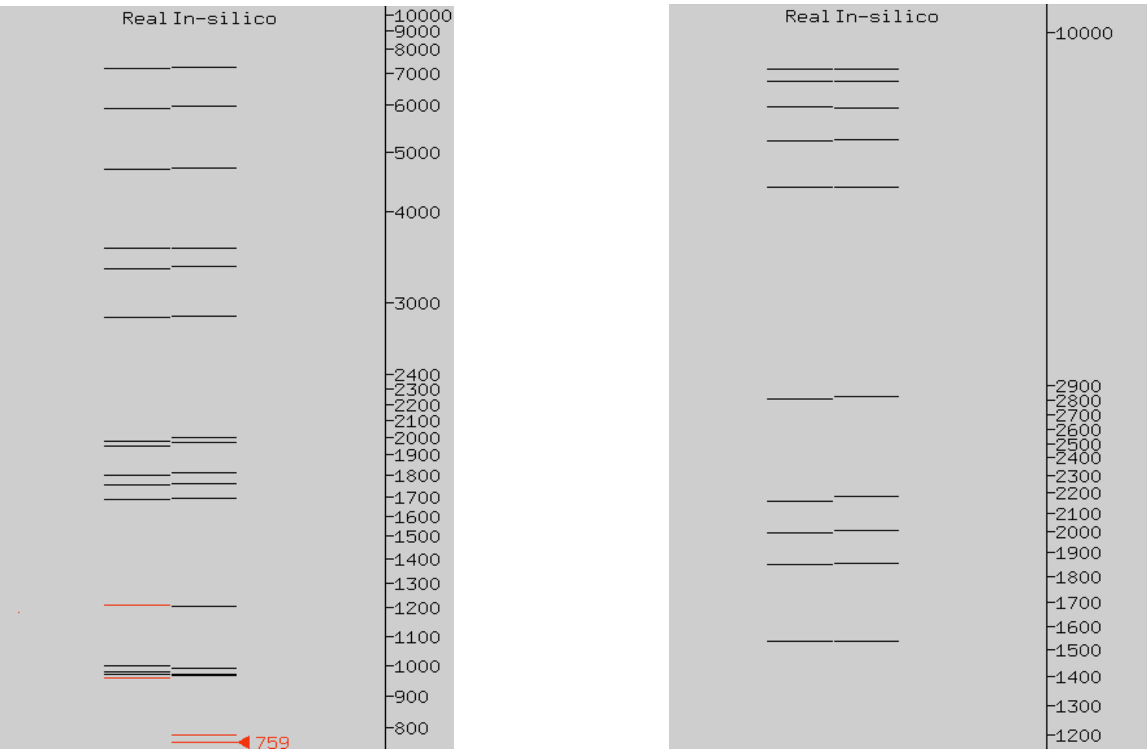


Fig 10.a. EcoRV real and *in-silico* digest **b.** HindIII real and *in-silico* digest.

The real 1210 fragment band is red (indicating a supposed inconsistency) because the actual gel band was computer-designated as a doublet, whereas the *in silico* was not. By looking at the actual gel image, however, I determined that the band in question was, in fact, a singlet, and thus the real and *in silico* digests were consistent. (Bands smaller than 1kb were ignored.)

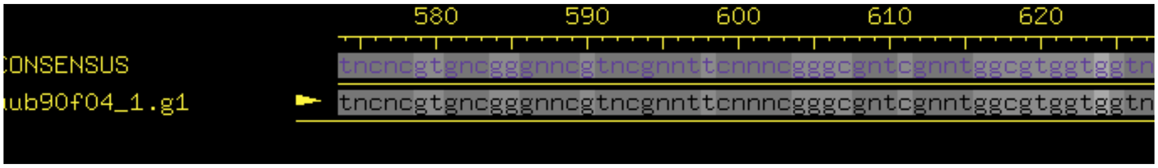
With this restriction digest comparison, I concluded that the high quality discrepancy at base 28069 was not the result of an overlapping duplicate sequence. Instead, the discrepancy most likely resulted from a mutation to either G or A that arose during clone amplification. This mutation then propagated permanently as cell replication continued, resulting in discrepant sequence between individual subclones.

ROUND II FINISHING

Low Base Quality

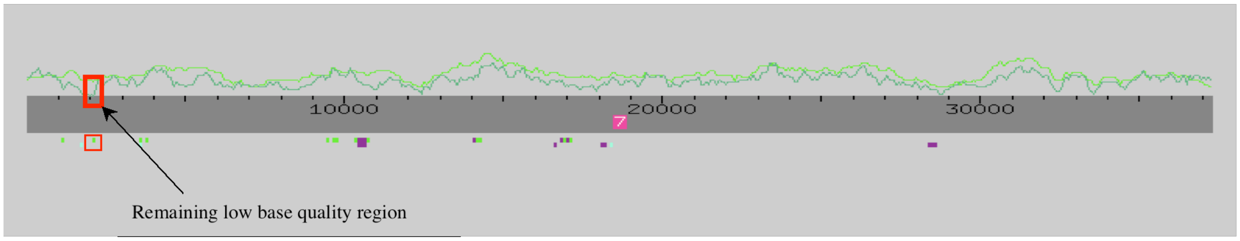
New reads ordered in the first round of finishing failed to completely resolve the low base quality areas between positions 2008 and 2155. The read ordered to add coverage to this area was not of high enough quality to incorporate into the Consed assembly. Figure 11 shows a representative stretch of this read.

The second new read, however, sufficiently improved sequence quality at position 28724, having a quality value above 30 at this position. The third reaction (using Oligo III) was never completed, although I did determine the consensus sequence in question as correct by examining the trace data. At this point, Consed navigation still did not register this third low base quality region highlighted by Autofinish.



```
                    580       590       600       610       620
CONSENSUS        tncncgtgncgggnncgtncgnnttcnnncgggcgntcgnntggcgtggtggtn
ub90f04_1.g1  ►  tncncgtgncgggnncgtncgnnttcnnncgggcgntcgnntggcgtggtggtn
```

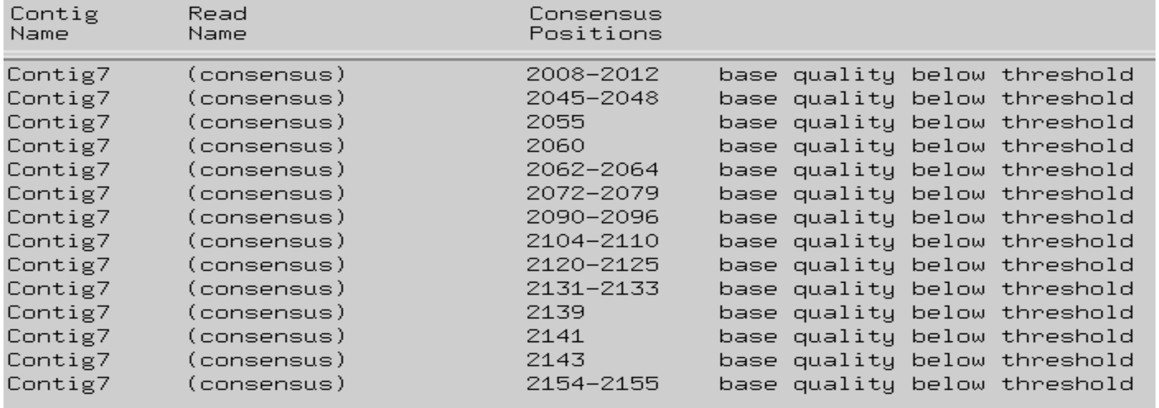
Fig 11. Failed reaction intended to cover low quality bases from position 2008 to 2155 (Oligo I from table 3).



Remaining low base quality region

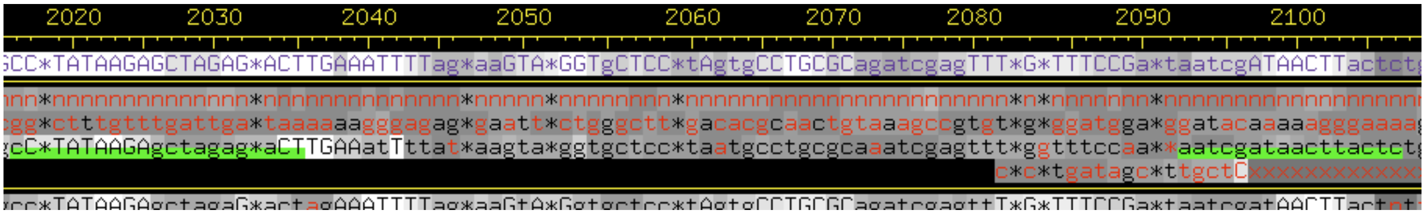
Fig 12. Assembly view after new reads added. Only reaction with Oligo II (table 3) improved sequence quality (at base 28724).

Since the Oligo I reaction from the first round ended up failing, several low quality regions still existed between bases 2008 and 2155 (Table 5). The prevalence of low quality reads and failed reactions in this area may have resulted from the presence of a hairpin loop structure on the template DNA. Such a structure was suggested by two sequence strings within the low quality region from position 2008 to 2155, which had high quality matching sequences elsewhere in the assembly. These sequences are highlighted in green in Figure 12. The sequence on the left, at position 2018, matched a complementary sequence beginning at base 1119 while the right hand string, at base 2092, had a complementary sequence beginning at base 1047. A potential hairpin loop structure formed through alignment of these matching sequences could disrupt template-walking reactions in this area. Thus, a hairpin loop could perhaps be responsible for the low quality results.



Contig Name	Read Name	Consensus Positions	
Contig7	(consensus)	2008-2012	base quality below threshold
Contig7	(consensus)	2045-2048	base quality below threshold
Contig7	(consensus)	2055	base quality below threshold
Contig7	(consensus)	2060	base quality below threshold
Contig7	(consensus)	2062-2064	base quality below threshold
Contig7	(consensus)	2072-2079	base quality below threshold
Contig7	(consensus)	2090-2096	base quality below threshold
Contig7	(consensus)	2104-2110	base quality below threshold
Contig7	(consensus)	2120-2125	base quality below threshold
Contig7	(consensus)	2131-2133	base quality below threshold
Contig7	(consensus)	2139	base quality below threshold
Contig7	(consensus)	2141	base quality below threshold
Contig7	(consensus)	2143	base quality below threshold
Contig7	(consensus)	2154-2155	base quality below threshold

Table 5. Remaining regions of low consensus quality after 1st round reactions added to assembly.



```
  2020      2030       2040       2050        2060      2070      2080        2090       2100
GCC*TATAAGAGCTAGAG*ACTTGAAATTTTag*aaGTA*GGTgCTCC*tAgtgCCTGCGCagatcgagTTT*G*TTTCCGa*taatcgATAACTTactctg
nnn*nnnnnnnnnnnnnn*nnnnnnnnnnnnnn*nnnnn*nnnnnnnn*nnnnnnnnnnnnnnnnnnnnnnnn*n*nnnnnnnn*nnnnnnnnnnnnnnnnnnnn
cgg*ctttgtttgattga*taaaaaagggagag*gaatt*ctgggctt*gacacgcaactgtaaagccgtgt*g*ggatgga*ggatacaaaaagggaaaag
gcC*TATAAGAgctagag*aCTTGAAatTttat*aagta*ggtgctcc*taatgcctgcgcaaatcgagttt*ggtttccaa**aatcgataacttactctg
                                                                       c*c*tgatagc*ttgctCxxxxxxxxxxxxx
gcc*TATAAGAgctagaG*actagAAATTTTag*aaGtA*Ggtgctcc*tAgtgCCTGCGCagatcgagttT*G*TTTCCGa*taatcgatAACTTactnt
```

Fig 12. Green highlighted sequences have high quality matches elsewhere.

To improve quality in this region, I again ordered further reactions to increase coverage. These reactions used Oligo I (table 3) on three different templates with all types of sequencing chemistry – big dye, dGTP, 4 in 1 (Table 6).

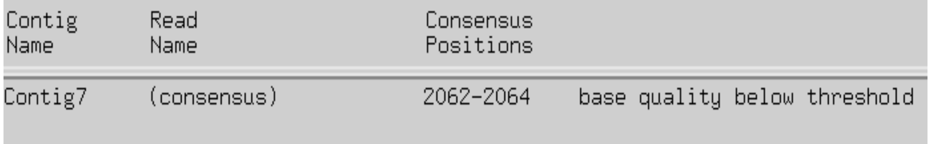
Sequence		Template	
Oligo I ggcactcgtaaaagtaaacaag	uub90f04.g1	uub91f12.g1	uub90h06.g1

Table 6. Round two reaction order list.

PRESENT STATE OF ASSEMBLY

Remaining problems

After adding reads from the reactions ordered in round two of finishing, almost all low base quality regions became resolved. Currently, one low quality region remains (Table 7, Figure 14, 15).



Contig Name	Read Name	Consensus Positions	
Contig7	(consensus)	2062-2064	base quality below threshold

Table 7. Low base quality regions.

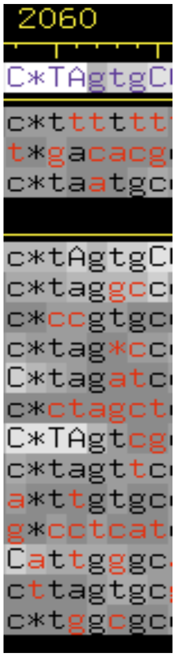


Fig 14. Low base quality region 2062-2064.

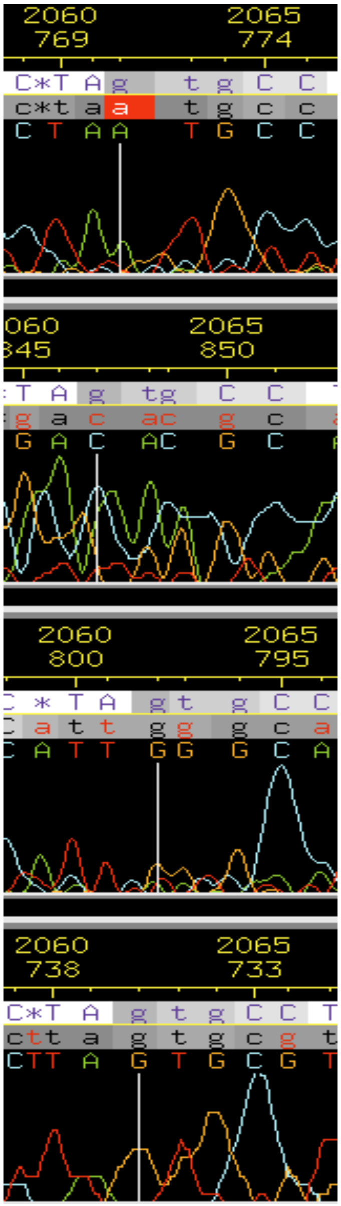


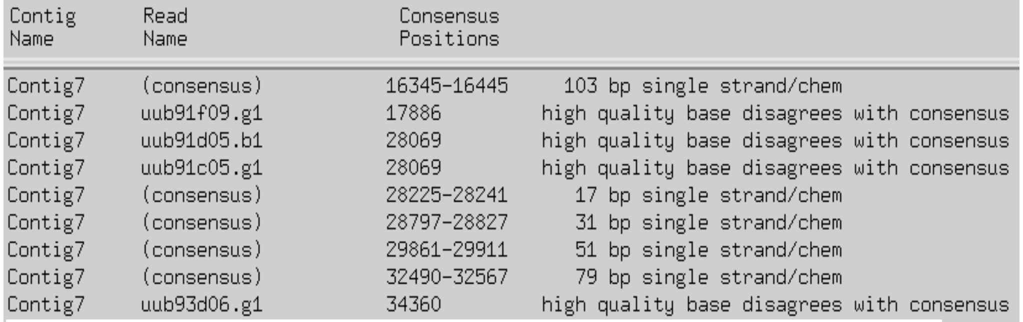
Fig 15. Trace data of bases 2062-2064.

Upon examining trace data of this remaining low quality region, I do not think I can confidently conclude a consensus sequence. For example, it is hard to conclude a G exists at base 2062 when considering the top two traces of figure 15.

Final checklist

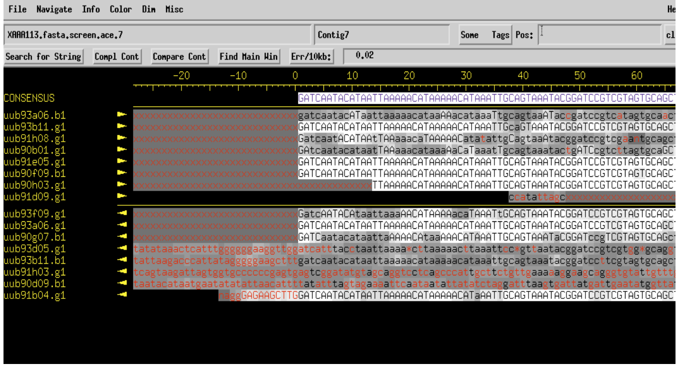
Thus, one low base quality region remains in my assembly. Beyond this region, navigation through Consed does not highlight any additional problem areas (Table 8). All regions of single stranded chemistry contain one or more reads with a quality value of 30 or higher at each base position. Additionally, I have checked and accounted for all high quality sequence discrepancies. (Though not listed elsewhere, the high quality discrepancy at base 34360 is due to a capillary error.)

The cloning sites of the assembly begin and end with the expected GATC sequence (Figure 16, next page). The full contig is 37,265 base pairs long, It contains one inverted repeat, and two tandem



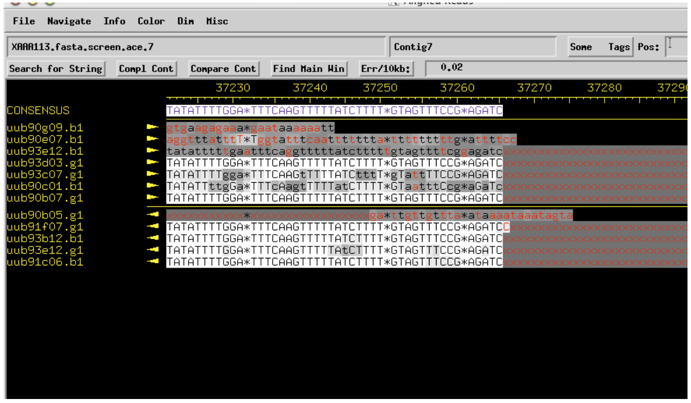
Contig Name	Read Name	Consensus Positions	
Contig7	(consensus)	16345-16445	103 bp single strand/chem
Contig7	uub91f09.g1	17886	high quality base disagrees with consensus
Contig7	uub91d05.b1	28069	high quality base disagrees with consensus
Contig7	uub91c05.g1	28069	high quality base disagrees with consensus
Contig7	(consensus)	28225-28241	17 bp single strand/chem
Contig7	(consensus)	28797-28827	31 bp single strand/chem
Contig7	(consensus)	29861-29911	51 bp single strand/chem
Contig7	(consensus)	32490-32567	79 bp single strand/chem
Contig7	uub93d06.g1	34360	high quality base disagrees with consensus

Table 8. Low quality/single stranded chemistry regions, and high quality discrepancies of final assembly.



```
                                -20       -10         0         10        20        30        40        50        60
CONSENSUS                                          GATCAATACATAATTAAAAACATAAAAACATAAATTGCAGTAAATACGGATCCGTCGTAGTGCAGC
uub93a06.b1 ►  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxgatcaatacATaattaaaaacataaAAacataaaTtgcagtaaATaccgatccgtcatagtgcaac
uub93b11.g1 ►  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxGATCAATACATAATTAAAAACATAAAAACATAAATTGcaGTAAATACGGATCCGTCGTAGTGCAGC
uub91h08.g1 ►  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxGatcaatACATAAtTAAaaacaTAAAAACatatattgCagtaaAtacggatccgtcgaantgcagc
uub90b01.g1 ►  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxGAtcaatacataatTAAaaacataaaAaCaTaaatTgcagtaaatactgATCcgtcttagtgcaGC
uub91e05.g1 ►  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxGATCAATACATAATTAAAAACATAAAAACATAAATTGCAGTAAATACGGATCCGTCGTAGTGCAGC
uub90f09.b1 ►  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxGATCAATACATAATTAAAAACATAAAAACATAAATTGCAGTAAATACGGATCCGTCGTAGTGCAGC
uub90h03.g1 ►  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxTTAAAAACATAAAAACATAAATTGCAGTAAATACGGATCCGTCGTAGTGCAGC
uub91d09.g1 ►                                                                      ccatattagcxxxxxxxxxxxxxxxxxxx
uub93f09.g1 ◄  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxGatcAATACAtaattaaaAACATAAAAacaTAAATtGCAGTAAATACGGATCCGTCGTAGTGCAGC
uub93a06.g1 ◄  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxGATCAATACATAATTAAAAACATAAAAACATAAATTGCAGTAAATACGGATCCGTCGTAGTGCAGC
uub90g07.b1 ◄  xxxxxxxxxxxxxxxxxxxxxxxxxxxxxGATCaatacataattaAAAACAtaaAAACATAAATTGCAGTAAATaCGGATccgTCGTAGTGCAGc
uub93d05.g1 ◄  tatataaactcatttggggggaaggttggatcatttacctaattaaaa*cttaaaaacttaaattcc*gttaatacggatccgtcgtgg*gcagg
uub93b11.b1 ◄  tattaagacccattataggggaagctttgatcaatacataattaaaaacataaaaacataaattgcagtaaatacggatccttcgtagtgcagc
uub91h03.g1 ◄  tcagtaagattagtggtgccccccgagtgagtcggatatgtagcaggtcctcagcccattgcttctgttgaaaaaggaagcagggtgtattgttt
uub90d09.b1 ◄  taatacataatgaatatatattaacattttatatttagtagaaaattcaataatattatatctaggatttaagtgattatgattgaatatggtta
uub91b04.g1 ◄                naggGAGAAGCTTGGATCAATACATAATTAAAAACATAAAAACATaAATTGCAGTAAATACGGATCCGTCGTAGTGCAGC
```

Fig 16. Verified cloning sites.



```
                      37230     37240     37250     37260     37270     37280     37290
CONSENSUS         TATATTTTGGA*TTTCAAGTTTTTATCTTTT*GTAGTTTCCG*AGATC
uub90g09.b1 ►     gtgaagagaaa*gaataaaaaatt
uub90e07.b1 ►     aggtttatttT*Tggtatttcaattttttta*tttttttttg*atttcc
uub93e12.b1 ►     tatattttgaatttcaggttttatcttttgtagttttcggagatcxxxxxxxxxxxxxxxxxxxx
uub93d03.g1 ►     TATATTTTGGA*TTTCAAGTTTTTATCTTTT*GTAGTTTCCG*AGATCxxxxxxxxxxxxxxxxxxxx
uub93c07.g1 ►     TATATTTTgga*TTTCAAGtTTTTATCtttT*gTattTTCCG*AGATCxxxxxxxxxxxxxxxxxxxx
uub90c01.b1 ►     TATATTttgGa*TTTcAagtTTTTatCTTTT*GTaatttCcg*aGaTcxxxxxxxxxxxxxxxxxxxx
uub90b07.g1 ►     TATATTTTGGA*TTTCAAGTTTTTATCTTTT*GTAGTTTCCG*AGATCxxxxxxxxxxxxxxxxxxxx
uub90b05.g1 ◄     xxxxxxxxxx*xxxxxxxxxxxxxxxxxga*ttgttgttta*ataaaataaatagta
uub91f07.g1 ◄     TATATTTTGGA*TTTCAAGTTTTTATCTTTT*GTAGTTTCCG*AGATCCxxxxxxxxxxxxxxxxxxx
uub93b12.b1 ◄     TATATTTTGGA*TTTCAAGTTTTTATCTTTT*GTAGTTTCCG*AGATCxxxxxxxxxxxxxxxxxxxx
uub93e12.g1 ◄     TATATTTTGGA*TTTCAAGTTTTTAtCTTTT*GTAGTTTCCG*AGATCxxxxxxxxxxxxxxxxxxxx
uub91c06.b1 ◄     TATATTTTGGA*TTTCAAGTTTTTATCTTTT*GTAGTTTCCG*AGATCxxxxxxxxxxxxxxxxxxxx
```

repeats – one of which is a 412 bp long repeat with 100% identity with copies at position 10375 and 16767 (Figure 18). As shown above, real and *in silico* digests of the contig by EcoRV and HindIII give matching results (Figure 10). These findings assure that the contig assembly is correct and accurate.

Additionally, sequences following long runs of T's or C's are free of potential mistakes in base calling. Searching for 15 bp stretches of a single base, I found a stretch of T's at position 4394-4408, and a stretch of A's from position 22361 to 22375. No long stretches of G's or C's are present. Toward the end of such long base stretches, a sort of fatiguing may occur that results in base calling mistakes. For example, the traces of T's following a stretch of consecutive T's may become diminished – possibly being lost or hidden under other base traces, leading to incorrect base calls. Trace data shows that this did not actually occur following long stretches in my contig (Figure 17).

As a final assurance of assembly quality, results of the program Findid establish that no contaminating E. Coli sequences exist within my contig.

As stated above, my initial assembly yields one 37kb contiguous sequence containing one inverted repeat, two tandem repeats, and one as of yet unresolved low base quality region (Fig 18, next page).

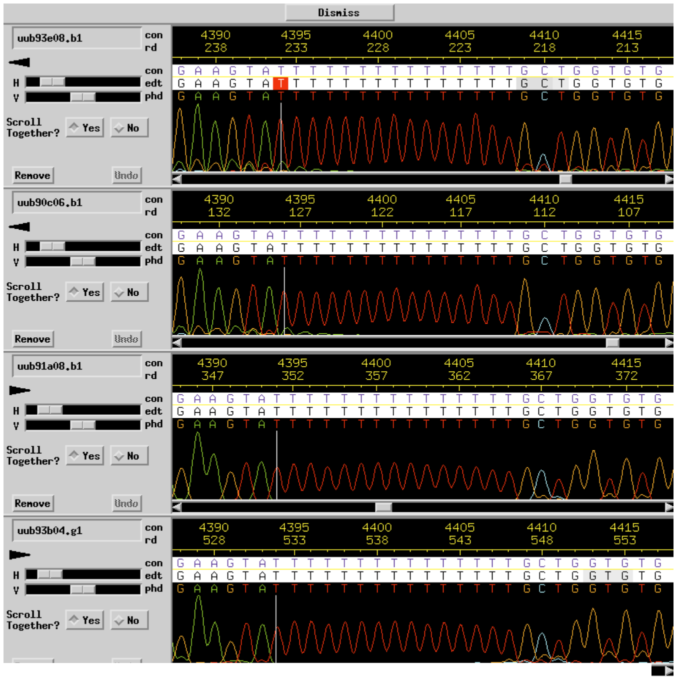

Dismiss
uub93e08.b1
con
rd
4390 238
4395 233
4400 228
4405 223
4410 218
4415 213
con G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
edt G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
phd G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
H
V
Scroll Together?
Yes
No
Remove
Undo
uub90c06.b1
con
rd
4390 132
4395 127
4400 122
4405 117
4410 112
4415 107
con G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
edt G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
phd G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
H
V
Scroll Together?
Yes
No
Remove
Undo
uub91a08.b1
con
rd
4390 347
4395 352
4400 357
4405 362
4410 367
4415 372
con G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
edt G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
phd G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
H
V
Scroll Together?
Yes
No
Remove
Undo
uub93b04.g1
con
rd
4390 528
4395 533
4400 538
4405 543
4410 548
4415 553
con G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
edt G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
phd G A A G T A T T T T T T T T T T T T T T T G C T G G T G T G
H
V
Scroll Together?
Yes
No

Fig 17. T-run. No T's are missed after long stretch in either direction.

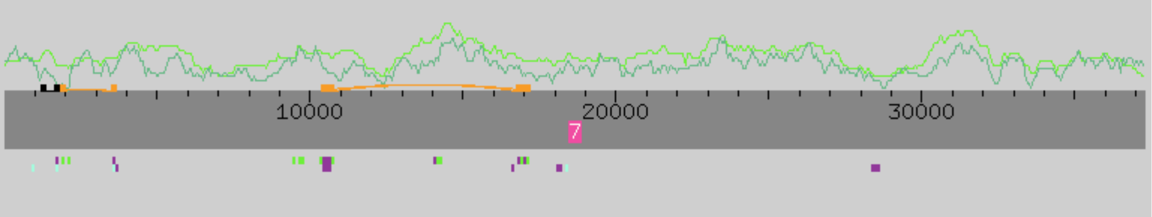


Fig 18. Final assembly view as of 3/4/04.