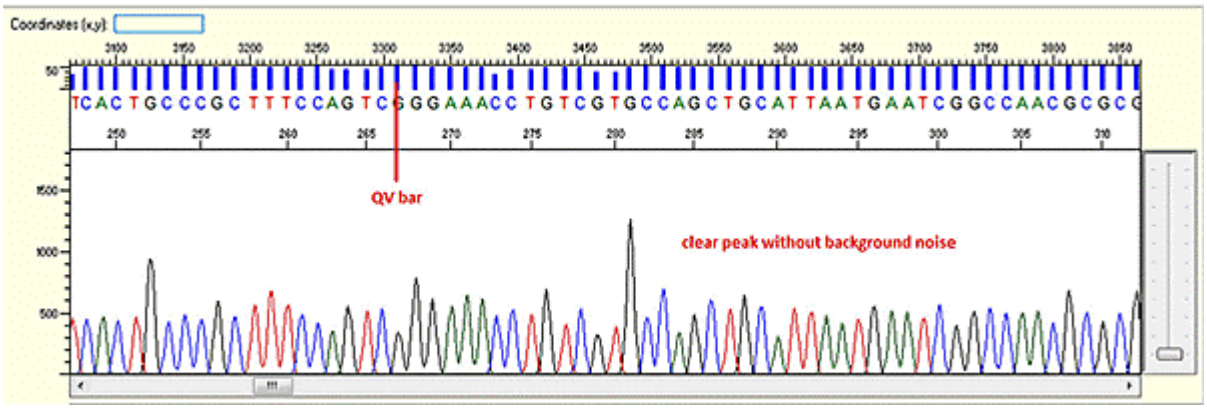
TROUBLESHOOTING GUIDE

1. Good Sequence Trace

It is important to identify what calls as a good sequencing trace. Our sequencing results show you the confidence of each base called by using KB Basecaller software. The confidence level is showed as quality value (QV) for each base.

Trace view for good sequence using Sequence Scanner software:



A good sequencing trace data will contain: -

1.Continuous long stretch or un-interrupted good quality value (QV) of the basecall. The QV will be displayed as rectangle bars at the upper part of the electropherogram.There are 3 categories of QV bars: -

1.Blue: high quality with >= QV 20

2.Yellow: medium quality with QV 15 to 19

3.Red: low quality with QV < 15

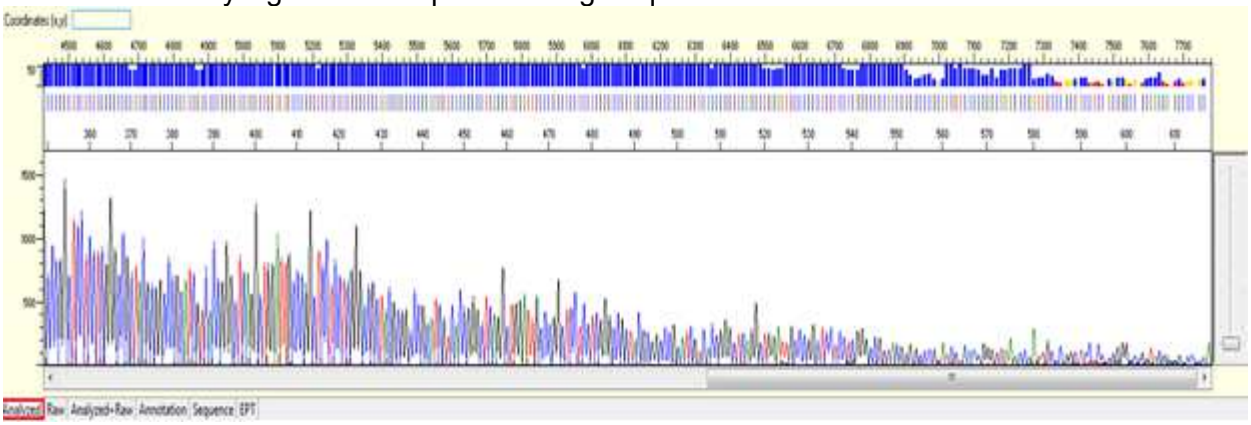
2.Well defined peaks with almost no or very minimum of background signal.

For sequencing trace of plasmid DNA or long PCR product (>1000bp), we will not be able to see the end of the trace (or we called as stop point) as the template size is larger than what the electrophoresis can separate. There should be a continuous read length (CRL) with good QV until 850 bases, then the resolution of the electrophoresis will gradually loss and there will be more low quality QVs.

2.Early signal loss

Pattern: Ski slope configuration of the trace.
The sequence starts with high quality peaks but become messy in the downstream. The raw data will show a good signal intensity at the beginning of the sequence and decreases when the sequence continues.

Trace view for early signal loss sequence using Sequence Scanner software:



Causes:

1.Improper amount of DNA template.
Too much starting DNA template will cause the used up of the chemistry at the beginning stage and thus little was left over for longer extension. For example, when the actual DNA template size given is too much shorter (e.g., 750bp) than what indicate in the order form (e.g., 1500bp), we may overload DNA template during the cycle sequencing.

Insufficient of starting DNA template, will cause fewer extending fragments generated in the sequencing reaction and thus the signal decrease when the sequencing continues. For example, when the actual DNA template size given is too much longer (e.g., 1500bp) than what indicate in the order form (e.g., 750bp), we may put too little template DNA during the cycle sequencing.

2.Present of contaminants in the DNA template / sequencing primer.
Present of contaminants (like RNA, phenol, ethanol, EDTA and salts) in the sample will give rise to premature termination because it affects the activity of the polymerase which leads to overabundance of short fragments generated.

3.Difficult region within the DNA template.
DNA polymerase had difficulty to process through some sequence context, for example homopolymer sequences, repetitive sequences, and region with high GC or AT content.

4.Degrading DNA template.

Degrading template DNA affect the amount of full template DNA which present in the cycle sequencing reaction. Extended fragments for degrading template will be short and the data will be noise with higher baseline especially at the start.

Solutions:

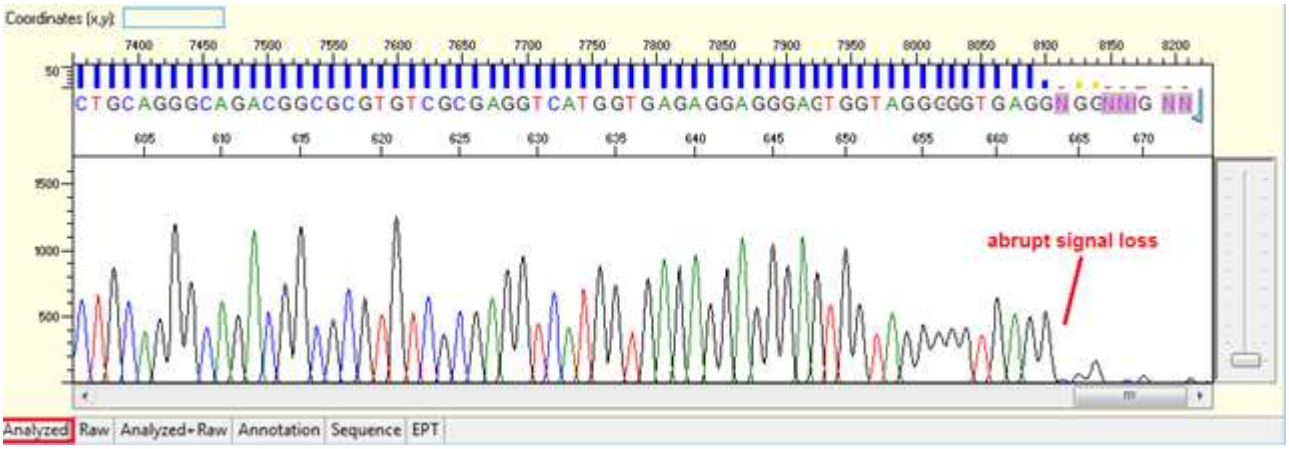
1.We 100% quantify each DNA template before cycle sequencing. We will thus be able to calculate the optimum volume required for each DNA template. Thus, providing us the correct or approximate size (less than 100 bases different from the actual size) of DNA template will prevent too much/ too little DNA template used in the cycle sequencing reaction. Please refer the guideline provided in our Order Preparation and Submission.

2.Any commercial column purification kit will be able to remove contaminants effectively. A good advice for researcher is that purifying bulk samples, we highly recommend processing the purification in several batches carefully. For purification kit using high salt as binding buffer, a skilful laboratory practitioner will prevent salts from contaminating the final elution. Alternatively, we also offer Sample Preparation and Quantification Services before sequencing.

3. Abrupt signal loss

Pattern: The signal is good before, and it suddenly terminates or drops without continued basecalling.

Trace view for abrupt signal loss sequence using Sequence Scanner software:



Causes:

1.Difficult region found in the sequence, which stops the polymerase reaction.

DNA polymerase had difficulty to process through some sequence context, here is the homopolymer, repetitive sequence and high GC or AT content.

2.Formation of DNA secondary structure.

The complementary regions fold up on themselves forming hairpin structures that the polymerase cannot pass through.

Solutions:

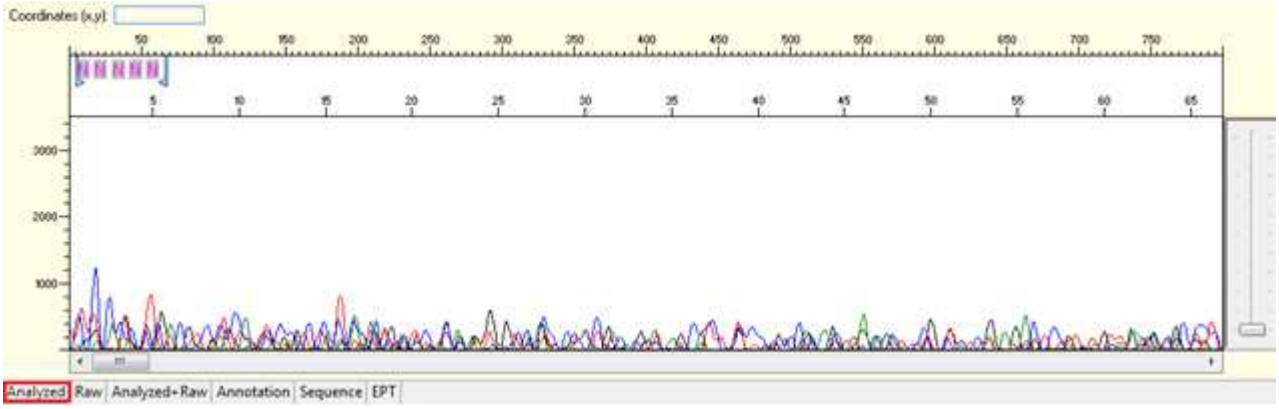
1.Sequence from the opposite strand.

2.Difficult region and DNA secondary structure from PCR product template tend to fail to be sequenced; cloning will be an option to get stable and thus longer read.

4. Noisy data with weak signal or no usable signal

Pattern for noisy data with weak signal: Undefined peaks, no continuous read length (CRL), and majority having low quality QV bars.

Trace view for noisy data with weak signal sequence using Sequence Scanner software:



Causes:

1.Too little DNA template or sequencing primer in the reaction.

2.Inhibitory contaminant present in template.

Present of contaminant (e.g salt, ethanol, phenol, RNA, etc.) will inhibit the binding of sequencing primer or polymerase to the DNA template to start the reaction.

3.Degraded DNA template or sequencing primer.

4.No priming site/ weak priming efficiency.

Solutions:

1.Quantify your DNA template through gel electrophoresis is highly recommended. For short DNA templates, as long as there is a visible DNA band when 1μL of your purified DNA loaded in agarose gel, it will be sufficient to be sequenced. However, large DNA templates will require more DNA amount to get the optimum reads.

2.Checking the purity of the dsDNA at ODA260/280 ratio with spectrophotometer e.g Nanodrop is recommended. A pure dsDNA should have ODA260/280 ratio between 1.8 to 2.0. In our opinion, if you are using a commercial kit, and they are not expired, the quality of the purified sample should be working well for sequencing typically. The contamination that leads to weak/ no usable signal usually is the serious degradation of DNA templates.

3.Contamination with DNase tend to cause different degree of degradation. Elution with TE buffer can help to protect your DNA from degradation, but EDTA in the TE buffer is one of the inhibitors for sequencing. To avoid inhibitory effect of EDTA in the sequencing reaction, you must make sure the DNA amount that present in the sample is higher/ fall under our sample requirement.

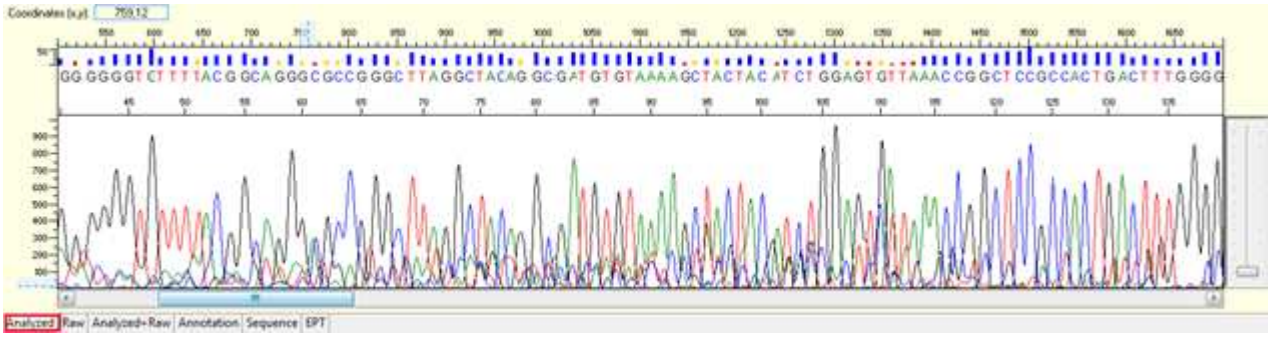
4.If you are sequencing a purified plasmid DNA, we highly recommend using the vector's sequencing primer for your first reaction to avoid disappointment.

For purified PCR products, most of the researchers will use one of their PCR primers as sequencing primer. If the design of PCR primers also fulfil the requirements as a sequencing primer, it will be fine.

5. Noisy data with strong signal or high level of back ground noises

Pattern: Multiple overlapping peaks with strong signal from beginning. The secondary peak from the same position may be the same, lower, or higher than the major peak.

Trace view for noisy data with strong signal or high level of background noises using Sequence Scanner software:



Causes:

1.Unspecific priming site.

Presence of more than 1 region within the DNA template, which complement partially to the sequencing primer. This cause the unspecific priming of the sequencing primer, and thus there will be more than 1 type of extension products produced from a single reaction tube. During the capillary electrophoresis, they only do the job of size separation. If the extension product is not unique at each size, the signal will be mixed.

2.Trace of excess PCR primers in the DNA template.

If the PCR products are not purified before sequencing, the excess PCR primers will act as sequencing primer in the sequencing reaction. Each primer will give different cluster of extension products, which lead to mixed signals.

3.Presence of more than one DNA template.

1.Occasionally, we will find some purified PCR products show more than 1 one DNA band in the agarose gel photo. Most of the commercial PCR Clean-up kits will not be able to remove the contaminated double stranded DNA or primer dimer. If the sequencing primer can bind on these unspecific DNA bands or primer dimer, it will generate different cluster of extension products and being analysed together to with the major extension cycle sequencing products, which leads to mixed signal. In some cases, it is pretty easy to identify the mixed signal under this category. For example: you have an unspecific DNA band at the size of 200bp, and your target PCR product is at the size of 750bp; you will find the mixed signal that will stop near the region of 180-base, following is the clean signal until the end of the sequencing trace.

2.There is a mutation in your DNA template, which causing frame shift deletion in the sequencing result. 1 frame shift = 1 amino acid = 3 basepairs. If your actual target PCR products is 750bp, a frame shift deletion sample will have both 750bp and 747bp of PCR products. You will not be able to distinguish such a small difference in size using any type of agarose gel electrophoresis. But the sequencing result will show different end point with 3 bases of differences. The mixed signal will start immediately after the mutation site, due to different signal from the extension products. This observation is however limited to the PCR product size. When the size of PCR products is greater than 850bp, it is very difficult to judge the different of the end point since the resolution of electrophoresis gradually loss beyond 850 bases.

4.Sequencing primer with low annealing temperature.

When the Ta is lower than 48°C, we will follow your recommended Ta on the order form. The same Ta may work for your PCR, but not necessarily works for sequencing. PCR is an exponential reaction that involves 2 PCR primers, but sequencing is a linear reaction that involves 1 sequencing primer.

Solutions:

1.Redesign primer specifically for sequencing. If you are sequencing a plasmid DNA, the vector's primer usually works very well.

2.Optimize the PCR protocol to have depleted primers during the end of PCR, so you can send your unpurified PCR products directly for sequencing. However, sometimes it will fail when the PCR is out of control. The best safeguard is to purify your PCR products before sequencing. Alternatively, please order purification service before sequencing.

3.Examine your purified DNA template on agarose gel before sending for sequencing. If there is more than one DNA band observed on agarose gel, purify them with gel extraction purification rather than PCR clean-up.

4.If you are bothering with the DNA template that showing mutation, the best solution is to clone the fragment and pick some colonies for sequencing. You should justify the cost of DNA cloning only if you are really interested with the mutated fragment. Otherwise, put it aside and sequence the next sample.

5.If you are using one of your PCR primers as your sequencing primer, you may simply compliance the rules of sequencing primer when select the best PCR primer that designed by the primer software.

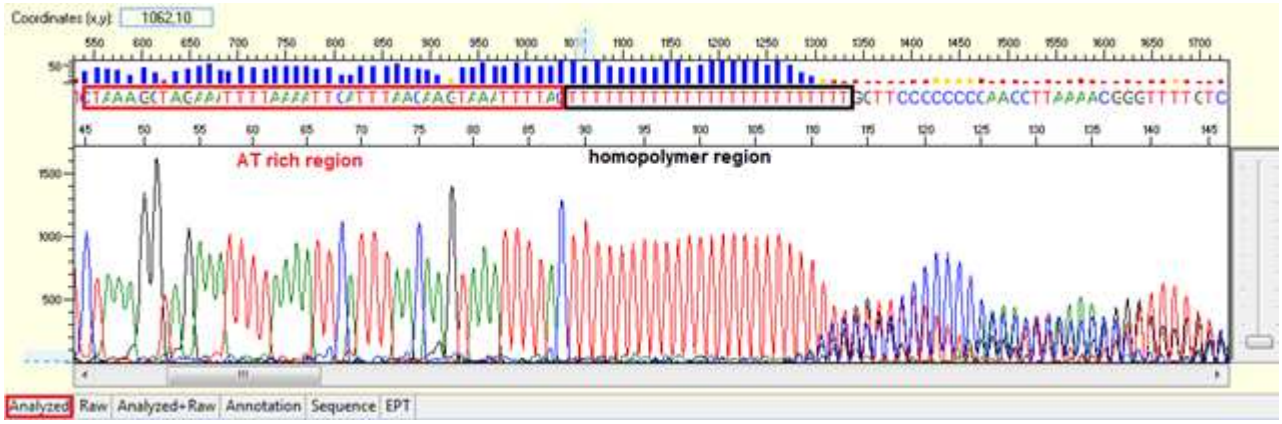
6. Difficult template (homopolymer, repetitive sequence)

Pattern:

1.Homopolymer: Noisy data after a long stretch of mono- (homopolymer) or di-nucleotide repeat.

2.Repetitive sequence: Noisy data after a region of the template DNA that is especially high in one or two of the nucleotides.

Trace view for homopolymer (difficult template) using Sequence Scanner software:



Causes:

1.Polymerase did not pair correctly with the template.
The polymerase dissociates and re-hybridized in a different location when met with a long stretch of mono- or di-nucleotide repeat generating various sized fragments and creating mixed signal after the region. This normally will generate a slippage pattern in the results. The condition tends to be more problematic in PCR products.

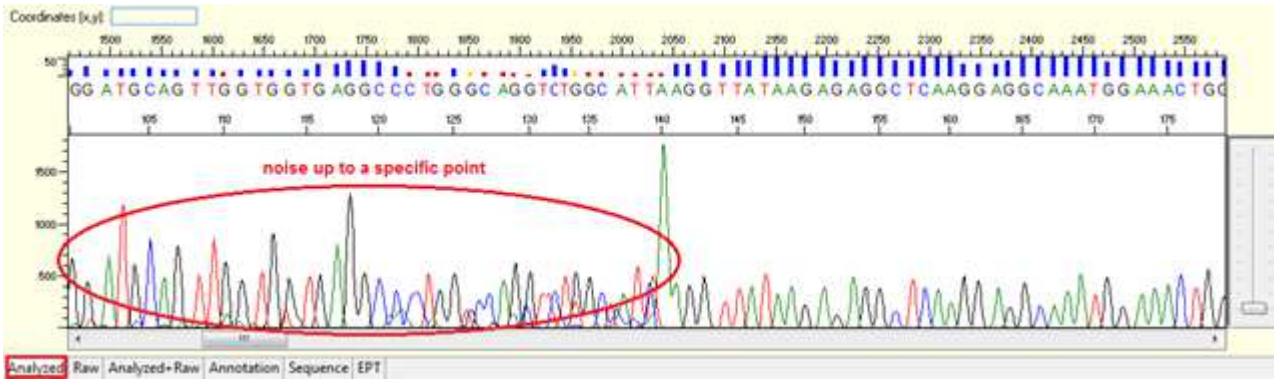
Solutions:

1.For PCR products, clone the fragment into a holding vector will help to stabilize the sequencing through the same homopolymer region. It is advisable to test on a few samples firstly before use on large number of samples.
2.Sequence from the opposite strand.
Although the sequencing will still fail during passing through the same homopolymer region, but the alignment will be successful. E.g. AAAAAAAAAA… will align with TTTTTTTTTT… from the opposite strand. You probably will lose the information of how many repeats of A. But if this is not what you are looking for, it is still worth for trying to save the cost from cloning.
3.Use anchored sequencing primer
The most common anchored primer is 21TV, where T is repeated for 21 units, followed by a degenerated primer of V (mixed of G, C and A). It looks like this: TTT TTT TTT TTT TTT TTT TTT V.The anchored primer will bind specifically on the homopolymer region and allows the polymerase to work normally right after them. It works very well when there is only ONE homopolymer region found within the DNA template. This however will lose some information (~ 20 to 30 bases) after the priming site, which is one of the disadvantages for BigDye Terminator v3.1 chemistry. If you need to recover this lost region, you need to have another sequencing primer that either sequence from the opposite strand until reaching the homopolymer region, or the original primer that binds upstream from the anchored primer.

7. Noise upto a specific point

Pattern: Data showing noisy peaks from the start until a specific point. The continue data showing good QVs.

Trace view for noise up to a specific point using Sequence Scanner software:



Causes:

1.Contamination of primer dimers.

The contamination of primer dimers will generate noise up to 150 bases from the beginning of the sequencing trace. Normally primer dimers can be seen as a cloudy band/ spot in the bottom of agarose gel quantification.

2.Contamination of smaller fragment/ **illegitimate** products.

Instead of seen as a cloudy band/ spot in the bottom of agarose gel quantification, smaller fragment/ illegitimate products will be seen as DNA band. In some cases, they are not easily detected using agarose gel when the loading volume is smaller than 5μL. They are not necessary generate high level of noise, especially when there is a high similarity of their sequence with the target fragment. But you can trace them easily by detecting the end point of extension products that produced by these smaller fragment/ **illegitimate** products in the beginning of the sequence. Their end point is just like the other PCR products that usually end by a high A peak by the Taq polymerase.

Solutions:

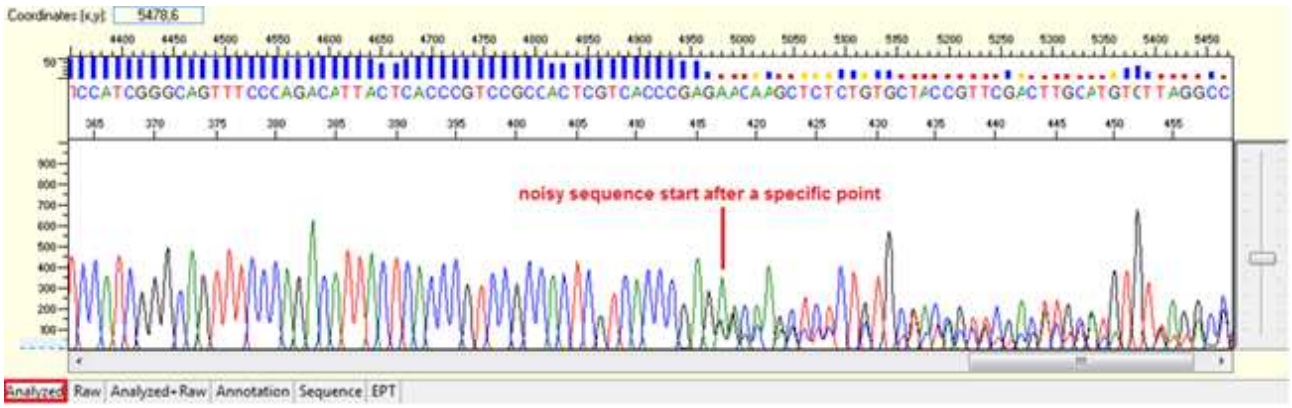
1.The quick way is to reduce the amount of PCR primers while performing the PCR. It works fine only if your PCR primers do not have any degenerate base(s).When you design your PCR primers, use the software to choose the best PCR primer sequence that does not perform self-dimers.

2.If you are able to recover this region from your opposite sequencing result, what you need to do is probably just cut off/ trim off the noise region before performing the pairwise alignment. This is workable if your DNA template is short (<850bp).

3.The easy way to resolve this is to gel purify the DNA template before sequencing.

8. Noise after a specific point

Pattern: Data is good at the beginning turns to double or more peaks after a specific position.
Trace view for noise after a specific point using Sequence Scanner software:



Causes:

1.Mixed preparation of a plasmid DNA sample.
There is more than one colony picked and extracted for sequencing. The noise is found right after the multiple cloning site (MCS).

2.Frameshift mutation.
Insertion or deletion of nucleotide occur in PCR product. The noise will start right after the mutation site.

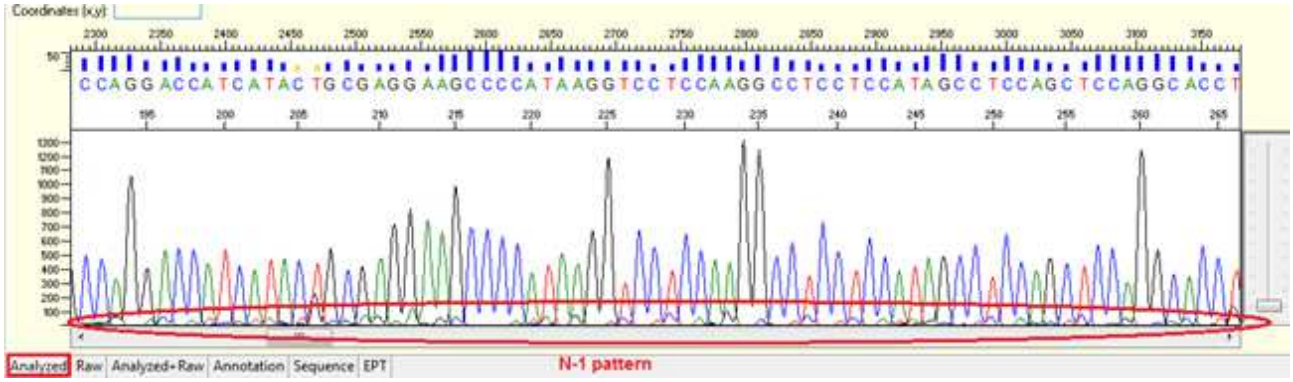
Solutions:

1.Re-extract the plasmid DNA from the single colony for re-sequencing.
2.The best solution to sequence the mutated samples is still to clone the fragments and pick few colonies for sequencing. You should justify the cost of cloning only if you are really interested with the mutated fragment. Otherwise, you should sequence another sample.

9. N-1 pattern

Pattern: Overlapped peaks throughout with the second peaks generally 1 base smaller and being the same base as that of the true base immediately to the right. It looks like a small “tail” before the peak of the correspondent nucleotide.

Trace view for N-1 result pattern using Sequence Scanner software:



Causes:

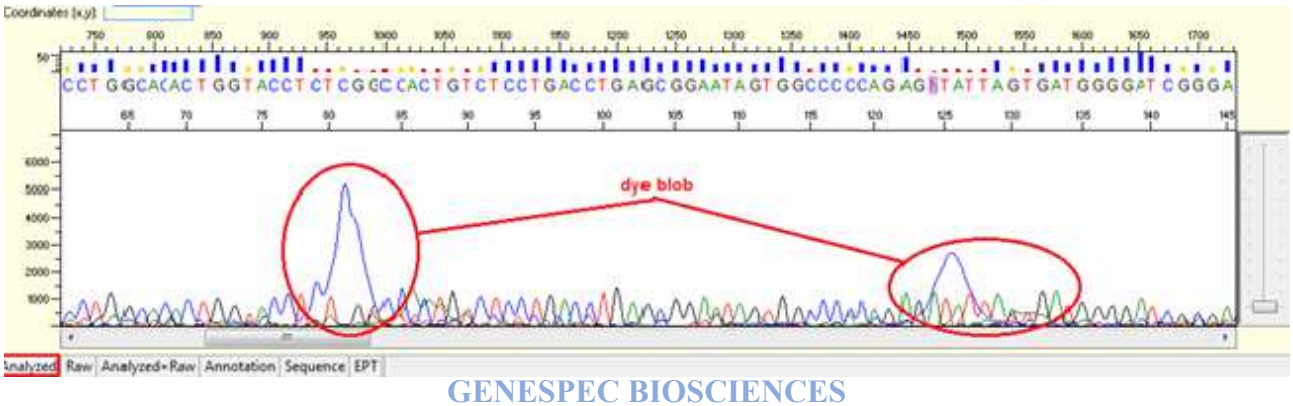
1.Poor synthesis quality of the sequencing primer because of low coupling efficiency.
If the contamination of N-1 primers is higher than 10%, it will be detected from the sequencing reads and affects the QVs.
The bad sequencing primer will affect each reaction without a miss.
2.Mis-priming
The same sequencing primer works very well for the other DNA templates in the same order, but only shows this pattern on some selected DNA templates.
3.Degrading of sequencing primer.
Degrading PCR primer may carry truncated oligo with -1, -2, -3… bases from the full length primer and affect the priming of the sequencing primer to the DNA template.

Solutions:

1.Ask replacement from your primer synthesis company. A new synthesis usually will resolve this problem immediately.
2.Re-design the sequencing primer by extending the primer length will help to increase the priming specificity.

10. Dye blobs

Pattern: The sequence contain dye blob may or may not affect the basecalling.
Trace view for dye blob result using Sequence Scanner software:



Causes:

1.Unincorporated dye terminator molecules that did not remove efficiently during the steps of post sequencing clean-up.
2.Degraded DNA template.

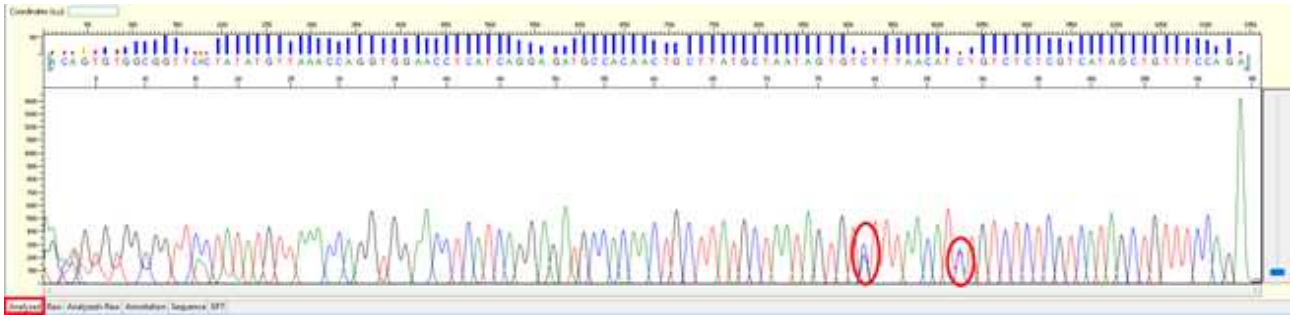
Solutions:

1.When we see dye blobs, we will automatically re-run the reaction. If you notice this in your results and you are very sure that the DNA template is freshly prepared, kindly contact us to request for the re-run.
2.Use freshly prepare DNA template for re-sequencing.

11. Single Nucleotide Polymorphism (SNP)

Pattern: A clear double peak present at a specific position within the continuous read length (CRL).

Trace view for result with SNP using Sequence Scanner software:



Cause:

1.The nature of genomic DNA that carrying SNP in the different allele of the gene. When PCR, it generates mix of PCR products which have one base different.

Solution:

1.SNP should not affect the analysis of a sequencing result. However, if one really wishes to get a clean sequence, DNA cloning will be an option before sequencing.