MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

MASSEY GENOME SERVICE

Sanger Sequencing and Genotyping using Life Technologies 3730 capillary instrumentation
Operational since 2003

QUALTRACE SOFTWARE REPORT
February 2019

BULLETIN INCLUDES

Quality Reporting of Data using Nucleics QualTrace™ Software
Possible causes of and solutions to poor sequencing data
Identification of sequencing data problems not reported in QualTrace™ Report

Enquiries regarding online request submissions contact:
Xiaoxiao Lin – Laboratory Manager

Te Kunenga ki Pūrehuroa

Massey Genome Service
Private Bag 11222, Palmerston North 4442, New Zealand **Lab** +64 6 951 8735 **Office** +64 6 951 9080
http://genome.massey.ac.nz

MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

MASSEY GENOME SERVICE (MGS)

Sanger Sequencing and Genotyping using Life Technologies 3730 capillary instrumentation

QUATRACE SOFTWARE BULLETIN

INTRODUCTION

The Nucleics QualTrace™ Software is a quality control monitoring tool for core sequencing and genotyping centres. MGS uses this software for the monitoring of sequence quality from each run and reporting of problems relating to sequence quality to its customers. The QualTrace™ software program enables the MGS to rapidly detect sequencing problems that limit sequence read length, and allows for the real-time analysis of trace files and detection of nine different sequencing problem types. This software creates an ideal means for troubleshooting DNA sequencing problems and optimizing production protocols.

The QualTrace™ software examines the raw sequencing data contained within each sequencing trace file to automatically identify nine problems with either the preparation of the DNA sequencing reactions, or with the Life Technologies 3730 sequencing instrument itself. The sequencing issues that can be identified by the QualTrace software include:

- Sequence traces that contain no signal data due to either a failed reaction or a blockage of the Life Technologies 3730 instrument capillary.
- Mixed trace signals resulting from multiple DNA templates in the sequencing reaction.
- Noisy or very low signal data traces that provide only short reads.
- Very weak signal strength at the end of the DNA sequencing trace resulting in short reads.
- Early mixed trace signal due to template contamination by PCR products.
- Delays in the starts of the DNA sequencing trace signal which indicates a significant overloading of the Life Technologies 3730 instrument capillary with template DNA and/or other contaminants.
- Problems with the Life Technologies 3730 instrument spectral calibration resulting in major channel cross-talk signal.
- The presence of significant amounts of unincorporated fluorescent labelled ddNTP remaining in the loaded DNA sequencing reactions generating dye blobs.
- Traces which show a rapid decline in peak signal.
- Indels: Traces with insertions and deletions.
- The point in the traces where the trace signal reaches the noise threshold.
- The levels of peak blur (poor resolution).

The QualTrace™ software also reports the following:

- Total number of high quality bases.
- Peak signal start point.
- Instrument information such as instrument name, run module, capillary length, run date, run name, plate well, polymer lot number, sample identification.

QUALTRACE™ REPORT

MGS runs all sequencing sample data through the QualTrace™ software. Customer specific information from each QualTrace™ analysis (i.e. online request submission) is uploaded to the MGS server as a report for you to download, along with your sequence chromatogram files, sequence read test files, and Analysis Report for each request. A QualTrace™ Report is generated for each sequencing request you submit to the MGS. The information below provides an interpretation of the QualTrace™ Report results, causes for the results and possible solutions.

QualTrace™ Report Information

The QualTrace™ Report contains the following information for each sequence in the request:

- **Trace File:** The name of the of sequence file, which includes your username, request ID, template and primer name for identification.
- **QT Class:** A short description of the Qual Trace™ analysis result.
- **Noise:** The average signal noise (FU) across all four raw data channels.
- **Spectral:** The cross talk signal in the channel showing the greatest level of signal crosstalk (i.e. signal in one channel due to signal in another). Values above 0.1 indicate that a spectral calibration should be performed.
- **Dye Signal:** The non-removal of excess unincorporated fluorescent labelled ddNTP terminators will cause sequencing “dye blob’ artifacts. Dye signal is a measure of the ratio of the largest potential “dye blob” peak to the average sequencing peak signal. Values greater than 1.0 are indicative of the presence of dye blobs. Values greater than 5.0 indicate major dye blob peaks are present.
- **Peak Start:** The scan number of the first sequence peak as recorded in the .ab1 file field B1*Pt.*
- **Dye:** Dye peak signal intensity relative to average peak signal.
- **Early Signal:** Signal-to-noise ratio of the first 40% of the raw trace.
- **Mid Signal:** Signal-to-noise ratio of the second 40% of the raw trace.
- **Late Signal:** Signal-to-noise ratio of the last 20% of the raw trace.
- **Noise Start:** The approximate base number where the trace signal to noise ratio falls below 12.
- **QCUT:** The quality score threshold for a good quality base. The default is Q20.
- **Total:** The total number of bases in the trace file.
- **Good Bases:** The total number of high quality bases in the trace file with quality scores higher than the QCUT.
- **Early Mix:** The ratio of the primary peaks which have an underlying secondary peak of more than 20% of the height of the primary peak, to the total number of primary peaks in the region from base 50 to base 250. Values above 0.1 are indicative of mixed signal in the early region of the trace.
- **Late Mix:** The ratio of the primary peaks which have an underlying secondary peak of more than 20% of the height of the primary peak, to the total number of primary peaks in the region from base 250 to base 550. Values above 0.1 are indicative of mixed signal in the later region of the trace.
- **Peak Width:** The width of the peak at base 700 measured at half the peak height. The larger this value the lower the peak resolution for a given run condition.
- **Indel:** Indicates whether the trace contains an insertion of deletion, or not. A value of 0 indicates no indel, and any number above 0 indicates the approximate base position of the suspected indel.

Interpretation of QualTrace™ Report Results

In the QualTrace® Report, in the section called **“QT Class”**, when you click on the result there is a short description of the result for each sequence read. The QualTrace® software places the sequencing reads into the following categories for reporting:

- Excellent
- Good
- Signal Limited
- Noisy Signals
- No Peaks
- PCR Product
- Early Mixed Peaks
- Late Mixed Peaks

- Late Peak Start
- Spectral Calibration
- Poor Resolution

Below is a description of what each result means.

Excellent

The trace file has a peak signal-to-noise ratio above 12 after base position 950. Peak signal strength is not the limiting factor for trace read length. Align able Q20+ bases are typically high to very high. Refer to Figure 1 and 2.

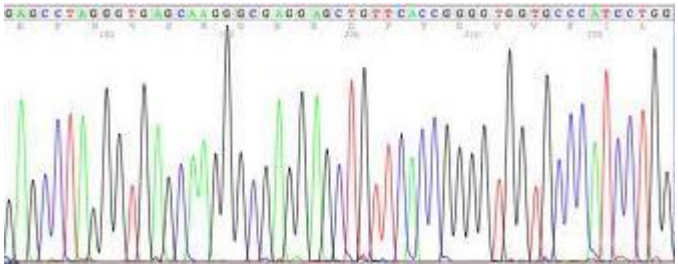

Fig.1

Figure 1: Excellent/Good Electropherogram image

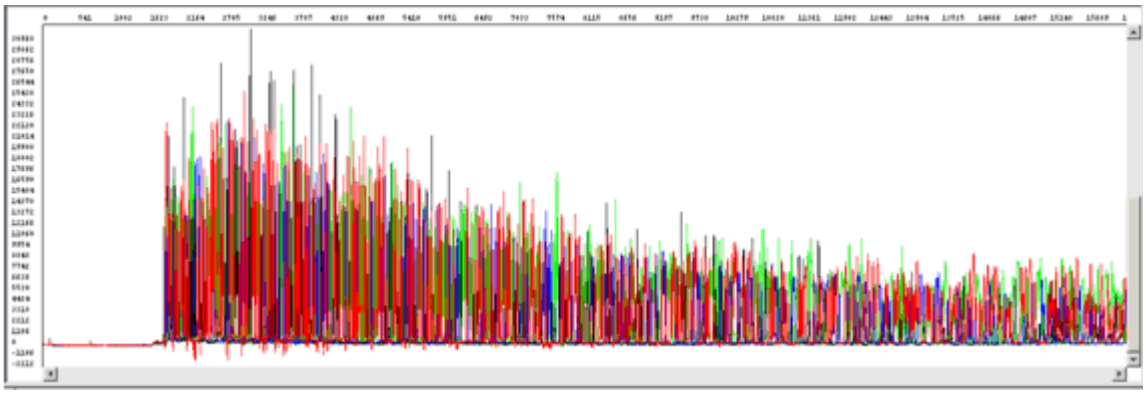

Fig.2

Figure 2: Excellent/Good Raw data image

Good

The trace file has a peak signal-to-noise ratio above 12 after base position 700, but not base position 950. Peak signal strength is a limiting factor for trace read length. Align able Q20+ bases are typically high. Refer to figure 1 and 2.

For excellent and good electropherogram traces, the raw data baseline should be uniformly flat across the entire length of the chromatogram. Secondly, the raw data should start at 20 minutes**.** The peak height at the start of the data should be around 125,000 FU. The peak height should gradually fall as time increases. You should see evenly-spaced peaks, each with only one colour. Peak heights may vary 3-fold, which is normal. "Noise" (baseline) peaks are very low but with good template and primer they will be quite minimal and would not disrupt the analyze data.

Signal Limited

The trace file has a low peak signal-to-noise ratio above 12 after base position 300, but not base position 700, resulting in short read lengths. Peak signal strength is a major limiting factor for trace read length. Align able Q20+ bases are typically low to moderate, with the sequence providing <700 bases of Q20+ bases. Trace signal ends abruptly. Refer to Figure 3 and 4. This result has many causes including:

- To little DNA in the sequencing reactions.
- Excessive sequencing chemistry dilution for the size of the product being sequenced.
- Contamination with salts, detergents, proteins, RNA, or other organic contaminants such as phenol. These contaminant can be preferentially injected into the capillary instead of the sequencing products.
- Excessive amounts of template DNA in the sequencing reactions which inhibits injection of the sequencing products into the capillary. Excess template can result in capillaries getting blocked.
- Loss of the sequencing products during the cleanup of the sequencing reactions. Loss of

sequencing products can be particularly problematic when using ethanol precipitation cleanup protocols.
- A region of the template which is difficult to sequence has been reached such as a secondary hairpin, or homopolymer region.

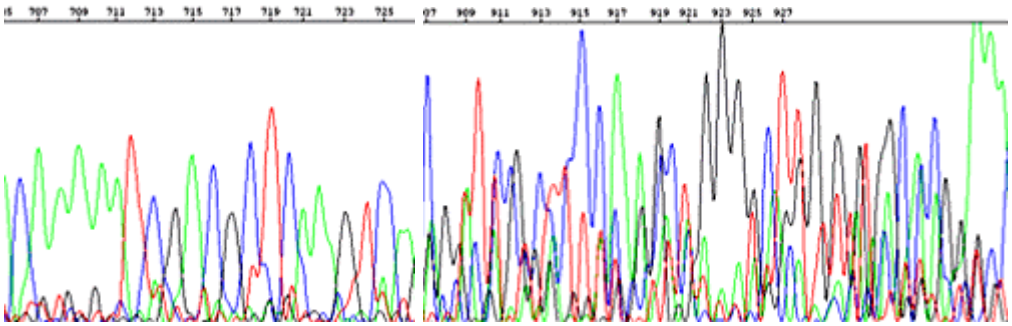

Fig. 3

Figure 3: Limited Signal Electropherogram image. Signal intensity decreases to the point where bases can no longer be called.

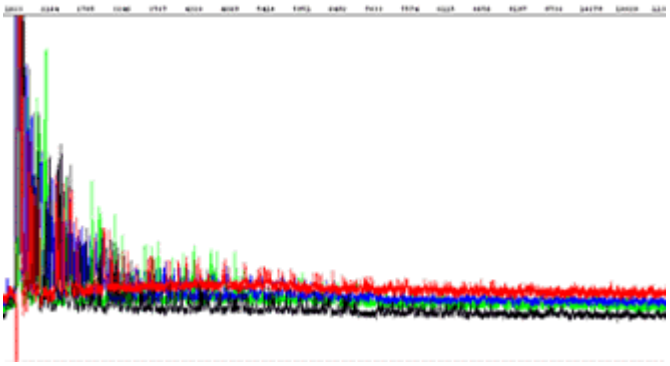

Fig. 4

Figure 4: Limited Signal Raw Data image.

Possible solutions for limited signal generating short trace reads:

- The concentration of the DNA template must be checked by running the product out on a 1-2% against a quantification standard of a similar size. Do not rely solely on spectrophotometer or fluorometer readings as they can often be inaccurate.
- Check to see if a secondary hairpin or homopolymer region has been reached. It may be possible to amplify the product region concerned by PCR using 7-deaza-deoxy guanosine triphoshate (7-deaza-dGTP), then sequence the resulting PCR product. This allows for sequencing through high G+C regions.
- Check the concentration of the oligonucleotide primer.
- Check the purity of the template.
- Consider sequencing from an amplified plasmid insert or change your method of template cleanup to a commercial plasmid preparation kit.

Noisy Signals

The trace file has a very low peak signal-to-noise ratio below 12 before base position 300. Very low signal-to-noise severely limits sequence read length. "Noisy" data can be identified by the presence of multiple peaks and numerous "N"s within your sequence. The “Sequencing Analysis” software assigns an "N’ as a base identification when there are two or more peaks present at one position. The electropherogram has noisy sequence peaks with low quality scores. Align able Q20+ bases are low. Refer to figure 5.

This result has many causes including:

- Incorrect template concentration.
- Excessive sequencing chemistry dilution for the size of the product being sequenced.
- Contamination with salts, detergents, proteins, RNA, or other organic contaminants such as phenol. These contaminants can be preferentially injected into the capillary instead of the sequencing products.
- Excessive amounts of template DNA in the sequencing reactions which inhibits injection of the sequencing products into the capillary. Excess template can result in capillaries getting blocked.

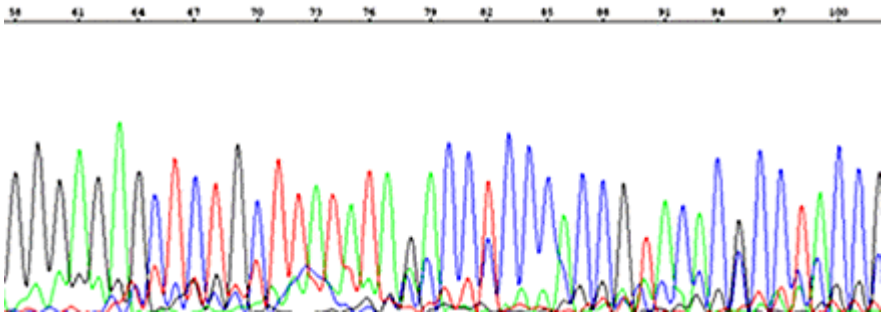

Fig. 5

Figure 5: Noisy Signals Electropherogram

Refer to “Solutions to failed and noisy sequencing reactions” below.

No Peaks

The trace file contains no or very weak peak signal in the raw data channels, which indicates a failed reaction. Refer to the section “Solutions to Failed sequencing Reactions” below, for possible solutions to this problem. The “Sequencing Analysis” software calls 5Ns. Align able Q20+ bases are very low to zero. Refer to figure 6. This result can be caused by:

- Wrong sequencing primers. Sequencing primers do not match the template being sequenced and hence the primers do not bind, and no sequence products are generated.
- Degraded sequencing primers which do not bind to the template.
- Losses of sequence products during the sequencing reaction clean up.
- Contamination with salts, detergents, proteins, RNA, or other organic contaminants such as phenol. These contaminant can be preferentially injected into the capillary instead of the sequencing products.
- Excessive amounts of template DNA in the sequencing reactions which inhibits injection of the sequencing products into the capillary. Excess template can result in capillaries getting blocked.

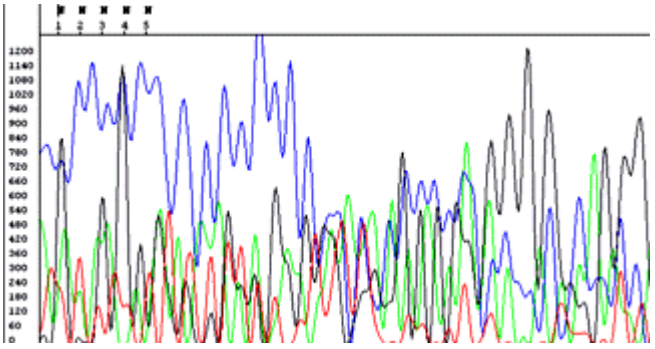

Fig. 6

Figure 6: Failed Sequencing Reaction. The “Sequencing Analysis” Software calls 5 Ns.

Refer to “Solutions to failed and noisy sequencing reactions” below.

PCR Product

Trace file appears to be derived from a PCR product less than 700 bases in length. The read length is limited by the length of the PCR product. Align able Q20+ bases are variable. All clear read PC products of less than 700bp in length will fall into this category.

Early Mixed Peaks

More than 50% of the primary peaks from base 50 to base 250 contain a secondary peak more than 20% of the height of the primary peak, resulting in a mixed signal. The electropherogram has two or more peaks at the same location. The peaks in the later regions of the trace are often good with only the beginning of the trace mixed. Align able Q20+ bases are typically moderate to high. The secondary peaks may be the same height as the primary peaks down to about 20%. Lower levels of mixed signal (below 20%) are normally base called well. Refer to figure 7.This result can be caused by:

- Two primer binding sites in the template
- Primer miss-annealing at secondary sites within the template

- Secondary primer PCR amplification in the sequencing reaction.
- A double pick of two colonies which are close together on the culture plate
- Poor quality PCR template containing multiple DNA fragments was used.

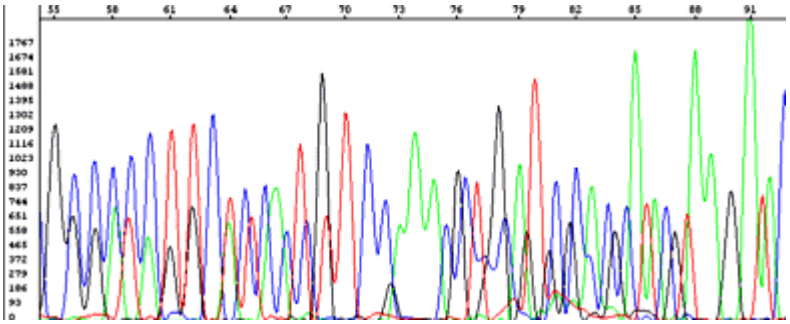

Fig.7

Figure 7: Electropherogram showing early mixed peak signals.

Possible solutions for early mixed peaks:

- **Check for the presence of multiple templates**. If sequencing from a plasmid prep making sure you select from a pure isolated colony. If sequencing a PCR product, check for the presence of a single product by running a 1-2% agarose gel. Even a relatively low amount of a contaminating PCR product can cause mixed template problems.
- **Check for the presence of multiple priming sites**. This can often occur when a fragment containing the primer priming site is sub-cloned into a vector that also contains the priming site. This problem is common when using the M13 forward and reverse universal primers.
- **PCR cleanup protocol.** Much sure you use a PCR clean-up protocol that remove leftover PCR primers. Even low levels of the PCR primers can cause mixed signal problems, especially if they have a high annealing temperature.
- **Melting temperature of sequencing primers.** Check the expected melting temperature of the sequencing primer. If the melting temperature is more than 5˚C above the annealing temperature used in the sequencing reaction then raise the annealing temperature. Because of the inclusion dITP in the BigDye™ sequencing mix the annealing/extension temperature can't be raised above 60˚C. If your primer is still mis-anneals at 60˚C then synthesize a new primer (this is easily done by removing bases from the 5' end until the Tm is below 60˚C).

Late Mixed Peaks

More than 50% of the primary peaks from base 250 to base 550, or the end of the basecall region on PCR product traces, contain a secondary peak of more than 20% of the height of the primary peak, resulting in a mixed signal. Align able Q20+ bases are low or moderate. Refer to figure 8.This result can be caused by:

- A plasmid prep that is contaminated by more than one product, such as two vectors with different inserts, or vector with insert and vector without, will generally show an early section of clean sequence data (common vector multiple cloning site sequence) followed by double peaks.
- A plasmid may contain more than one vector molecule or may encounter spontaneous deletions or insertions during growth. The point at which the double peaks begin corresponds to the start of the insert cloning site.

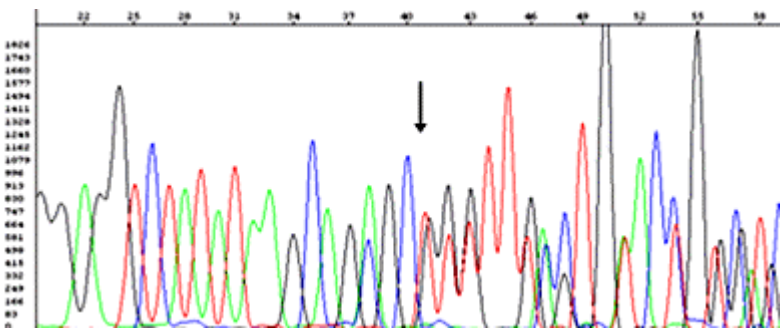

Fig. 8

Figure 8: Electropherogram showing late mixed peak signals, starting part way through the sequence read.

Possible solutions for late mixed peaks:

- It is important to pick a single colony from your growth plate. Re-streaking if necessary, to be sure that your colony is completely isolated and pure. You should follow this up with a restriction digest of your plasmid run out on an agarose gel to ensure vector and insert are present as expected.

Late Peak Start

The first signal peak collection has been time delayed by 33% compared to the expected collection time for this trace run condition. This error can occur when the capillary is overloaded with protein, template DNA, or salt. The templates and or primers are contaminated. No Align able bases. Refer to figure 9A and 9B.

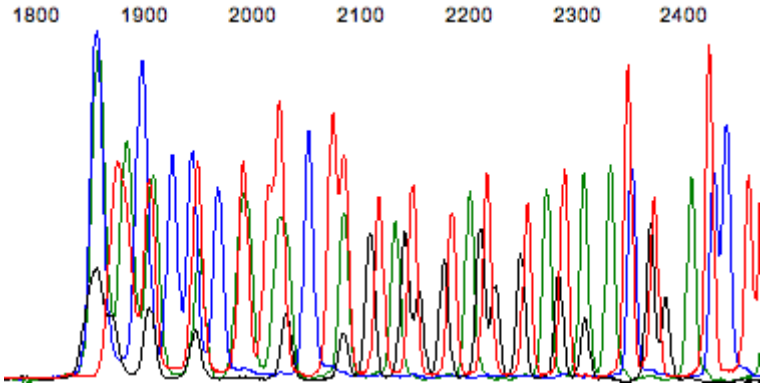

Fig. 9A

Figure 9A: Good start point at between 1800 and 1900 scans.

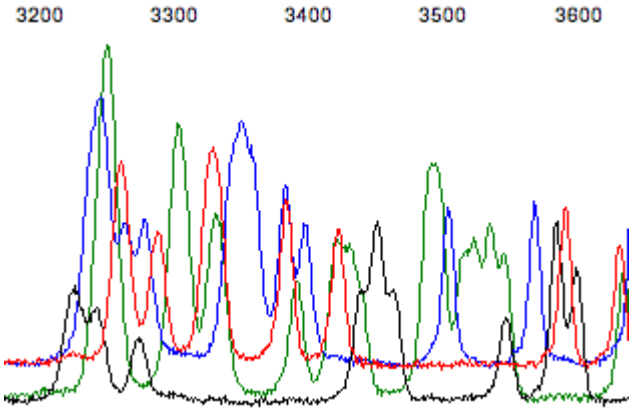

Fig. 9B

Figure 9B: Delayed start point at between approx. 3200 scans, showing poor peak resolution.

This result can be caused by:

- The capillary being overloaded.
- Dirty template DNA contaminated with proteins and/or salt.

Possible solutions to capillary overloading problems:

- Reduce the amount of DNA loaded into the sequencing reaction if you are setting up your own sequencing reactions, or supply less DNA template in the template/primer premix to the MGS.
- Supply cleaner DNA templates to the MGS.

Spectral Calibration

This indicates that the trace file contains significant fluorescent dye channel signal crosstalk (a minimum of 15% in at least one channel). Multiple occurrences of this suggest a machine spectral calibration should be performed. The MGS monitors this closely, and if ≥20% of samples shows this result across 3 consecutive runs, then a spectral calibration is performed. Align able Q20+ bases are moderate to high depending on the severity of the signal cross talk.

Poor Resolution

The peak width at base position 700, or 90% of the trace read length for traces less than 700 bases, is three times wider than expected for traces collected under this run condition. Align able Q20+ bases are moderate to low. The nucleotide base peaks appear as “blobs” rather than distinct individual base peaks. Refer to figure 10. Excessive peak widths are indicative of poor peak resolution and can result from:

- Overloading of the capillary due to excessive amounts of DNA in the sequencing reaction.
- Contamination with salts, detergents, proteins, RNA, or other organic contaminants such as ethanol.

Fig. 10

Figure 10: Poor resolution early in sequencing read.

From the image above you can see that poor resolution is primarily defined by the progressive broadening of peaks throughout the sequencing read. The peak may be wavy and contain shoulder peaks. The loss of resolution may start at the beginning of the sequence read in extreme cases, or somewhere in the middle of the sequencing read.

Possible solutions to poor resolution sequencing reads:

- Use less template into the sequencing reaction mix.
- Purify the DNA template
  - Carrying out a phenol/chloroform extraction followed by a chloroform extraction can improve the quality of sequencing reactions. Following this, carry out an ethanol precipitation.
  - Purify the DNA by ultra filtration using Centricon-100 columns, or Sephedex G50 spin columns.
  - Prepare new plasmid DNA taking care not to overload the plasmid miniprep columns.

SOLUTIONS TO FAILED AND NOISY SEQUENCING REACTIONS

Poor quality DNA

The problem of poor quality DNA resulting in failed sequencing reactions can be avoided by not sequence plasmid DNA directly and sequencing a PCR amplified fragment of the plasmid insert. If this is not possible then it is recommended that a plasmid miniprep kit is used.

Loss of sequencing reaction during reaction clean up

Loss of sequencing reactions during cleanup can be avoided by not using ethanol precipitation methods. There are a number of commercially available sequencing reaction cleanup kits that work very well, but of course are more expensive. If you are performing your own sequencing reactions and are cleaning up your products using an ethanol precipitation method, one way to avoid loss of the reaction DNA pellet is to add 1µl of a 20 mg/ml solution of glycogen to the sequencing reaction before adding the ethanol. This helps make the pellet visible and the glycogen does not seem to interfere with the injection of the sequencing fragments onto the sequencers capillaries.

Excessive amounts of template

This can be avoided by checking the concentration of the template on a 1-2% agarose gel against a quantification standard of similar size, before sequencing the products. You will be able to visually see the purity of the template DNA and if there is a significant amount of contaminating genomic DNA or RNA present. Do not rely solely on a spectrophotometer reading to calculate the template concentration, as this method is sometimes inaccurate.

Wrong primer

Using the wrong sequencing primer is simple to solve, but can be difficult to detect. Check the sequence of the primer and template to make sure that the primer binding site is present. This can be a particular problem with some "universal" forward and reverse primer sequences which do not work with some common plasmids. Do not trust other people's working stock solutions and always make up your own.

Degraded primer

Do not use old diluted primer stocks. Store the concentrated stock primers in 10mM Tris/ 0.1 mM EDTA (pH 8.5) rather than water. Don't use other peoples stocks, always much up your own. If you are in doubt as to the quality of your primers, then throw them out and make up a fresh working solution from the primer stock.

Failed oligonucleotide synthesis

If you suspect that the primer is of poor quality check it in a polymerase chain reaction (PCR). Alternatively if you have a control template that you know should work with the primer then this can be a good way of identifying primer problems.

NOTE:

Sequence base calling errors do occur at the beginning and at the end of each sequence read. It is the customer's responsibility to check for miss-calls and to truncate the data when the errors get to frequent.

There are a number of other sequencing problems which can be encountered, which are not covered by the QualTrace® Software reporting and these are mentioned below.

IDENTIFICATION OF OTHER SEQUENCING DATA PROBLEMS

Dye Blob Artifacts

The non-removal of excess unincorporated fluorescent labelled ddNTP terminators will cause sequencing "dye blob' artifacts. Dye blobs typically run at 60-90 bp and 100-115 bp. If the dye terminator contamination is severe you may also see peaks or shoulders at 170-190 bp and again at 400-410 bp.

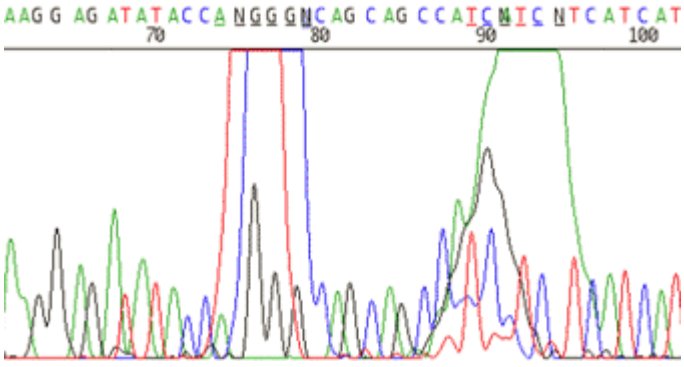

Fig.11

Figure 11: Excessive dye front problem.

Dye blob artifacts are typically seen when using ethanol based sequencing reaction cleanup protocols. If you are setting up your own sequencing reactions and you are having problems with dye fronts in your sequencing data, it would be a good idea to consider changing to a commercial non-ethanol based cleanup protocol. The MGS has made experience with the following commercial sequencing reaction cleanup kits, and both methods are effective in the removal of dye fronts:

- X-Terminator™: Supplied by Applied Biosystems Inc.
- Agencourt Cleanseq™: Supplied by Beckman Coulter

Causes of dye blob artifacts:

- Ineffective post-sequencing reaction cleanup. Usually occurs when a higher than recommended concentration of ethanol is used with ethanol based cleanup methods.
- Pellet is lost during the sequencing reaction clean up. This typically occurs at the 70% ethanol step when removing the ethanol.

Solutions to dye blob artifacts:

- Take special care when using ethanol based sequencing reaction cleanup methods:
  - Ensure the correct ethanol concentrations are used and replace ethanol solutions regularly with fresh solutions, as ethanol can draw moisture in from the air.
  - Ensure the correct concentration of sodium acetate is used.
  - Ensure the correct incubation time is used.
  - Ensure the correct centrifugation speed and time is used.
- To avoid drawing off the loose pellet after the 70% ethanol wash, consider carrying out a second centrifugation step.
- 100% ethanol will draw water from the air as soon as the lid is opened. Consider using 95% ethanol.

Sequencing Trace Signal Hard Stop

Good quality trace signal suddenly stops, or declines rapidly, with the trace peak signal intensity falling to less than 150 FU in the raw data channel after the stop point. The quality score after the "hard stop" is low. Refer to figures 12A and 12B.

Fig.12A

Figure 12A: Example of a "Hard Stop" – Processed channel

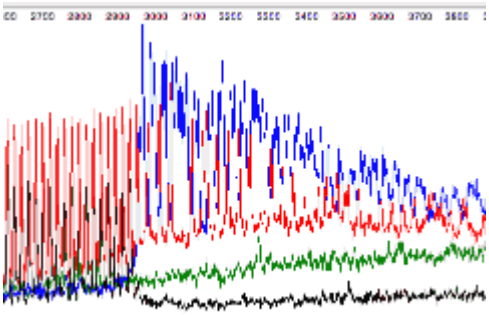

Fig. 12B

Figure 12B: Example of a "Hard Stop" – Raw Data channel

Causes of sequencing trace signal suddenly stopping:

- Secondary structures can fold back on themselves to form hairpin loops with prevent the sequencing DNA polymerase enzyme from passing through this region, hence the sequencing suddenly stops.
- Long strings of G and C bases in the template, which form strong secondary structures with hairpin loops and can also form single strand conformational structures, which the DNA polymerase has difficulty passing through. Long template regions of guanidine (G runs) are particular problematic because of the substitution of dITP for dGTP in the BigDye™ sequencing mix. The BigDye™ sequencing DNA polymerase does not efficiently extend runs of multiple inosine residues causing the polymerase to stop in the G run regions.

Solutions for overcoming sequencing trace "hard stops":

- Add 5%DMSO to the sequencing reaction before PCR amplification. The MGS can add 1µl of DMSO to the sequencing reaction upon request. DMSO helps in the degradation of secondary structures. Please request this in the customer notes section of the online sequencing request submission.
- Use the dGTP BigDye™ Terminator 3.0 sequencing chemistry instead of the BigDye™ Terminator 3.1 sequencing chemistry, for high GC rich templates and for template with high GC, GT and GA rich regions. Please request this by selecting the dGTP BigDye™ Terminator

3.0 chemistry from the “Reaction Type” pull-down menu of the online sequencing request submission.

Mononucleotide Sequencing Slippage

The sequencing data becomes mixed after a mononucleotide string of bases in the DNA template. This problem also occurs with strings of dinucleotide bases indicative of stutter bands seen in microsatellite polymerase chain reaction (PCR) amplification. Refer to figures 13A and 13B.

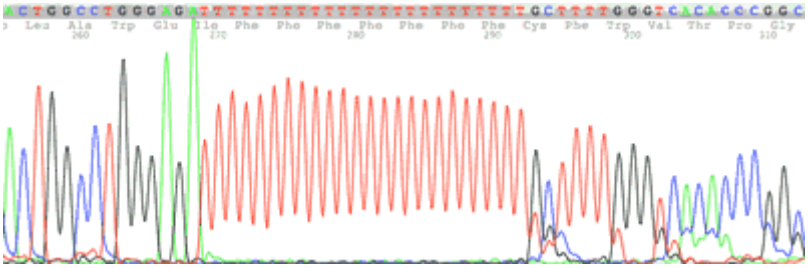

Fig.13A

Figure 13A: Sequence slippage on a mononucleotide T run

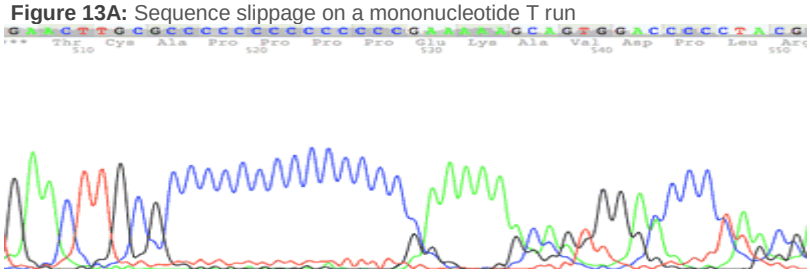

Fig.13B

Figure 13B: Sequence slippage on a mononucleotide C run

Causes of sequencing slippage problems:

- Long strings of mononucleotide bases cause the DNA to slip on the template, where the extension product loops over and rehybridizes, results in sequencing products of varying lengths. These products appear to have mixed signals in the traces downstream on the mononucleotide bases.
- An insertion or deletion in one allele. This can occur with PCR products where two alleles are sequenced together in the one reaction and one of the two mononucleotide strings has an “insertion or deletion” polymorphism.

Solutions for solving sequencing slippage problems:

- Sequence the product in both directions.
- Use a custom primer designed to hybridize just outside the mononucleotide or dinucleotide run region.

Template Insertions and Deletions

The chromatogram trace will appear as clear sequence before the indel but will become mixed after the indel. Often small peaks will be found before or after large peaks, which is the results of leading or trailing peaks (shifted peaks). Refer to figures 14A and 14B below.

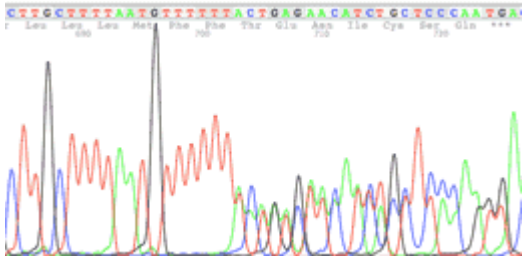

Fig.14A

Figure 14A: Chromatogram showing indel

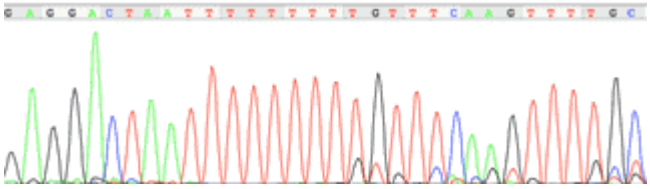

Fig. 14B

Figure 14B: Chromatogram showing shifted peaks due to indel

Causes of sequencing insertion and deletion problems:

- Direct sequencing of amplicons derived from diploid templates containing heterozygous regions.
- Random mutations occurring during cloning or the colony selection stage of the plasmid preparation procedure.

Solutions for solving sequencing insertion and deletion problems:

- If the template is heterozygous, then clone the PCR product before sequencing and sequence from the cloned plasmid.
- Sequence the other strand of the DNA using a reverse primer.

Chimeric Rearrangements

The chromatogram trace will appear as clear sequence up to the chimeric region, and the trace up to this point will align to its expected reference, but will become mixed after with no homology.

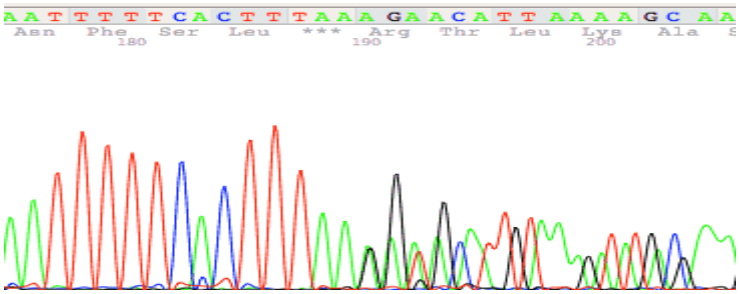

Fig. 15

Figure 15: Chromatogram showing a chimeric rearrangement

REFERENCES

- http://www.ki.se/kiseq/KIGene%20troubleshooting.pdf
- http://www.nemoursresearch.org/cores/bcl/forms/BCL_Sequencing_Troubleshooting_Guide.pdf
- http://www.nucleics.com/DNA_sequencing_support/DNA-sequencing-troubleshooting.html