DNA SEQUENCING:
FREQUENTLY ASKED QUESTIONS AND TROUBLESHOOTING

Most of the information below has been taken from Appliued Biosystems (http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_041258.pdf) . The pictures have been taken from Roswell Park Cancer Institute (http://www.roswellpark.org/Research/Shared_Resources/Biopolymer_Resource/DNA_Sequencing/TroubleshootingYourData).

1. PRIMER DESIGN AND REACTION SETUP

How should I design my sequence primer?

- Primers should be at least 18 bases long to ensure good hybridisation
- Try to achieve a T_m of approximately 55-60°C
- The GC content should preferably be approximately 50-55%. Higher GC contents might cause the denaturation to be incomplete
- Avoid secondary structures (as hairpins), particularly at the 3' end
- Avoid strings of four or more of one base
- Primers that can hybridise to form dimers should be avoided
- No alternative hybridisation sites present in the template

Can I change the annealing temperature in the cycling protocol?

Yes, it is important to consider the annealing temperature to match your primer. A general rule is that the annealing temperature in the cycling protocol is 3-5°C below the melting temperature (T_m) of the primer.

T_m (approx.) = (no of A/T x 2) + (no of G/C x 4)

Normally an annealing temperature of 50 - 55°C will give good sequence from a standard primer.

Is it possible to reduce the amount of sequencing premix?

Yes, the premix contains a surplus of reagents.

Make a fourth quarter size reaction to start with. If the sequence result is good and the signal peaks are strong, it is possible to make further dilutions of the kit. Always use the dilution buffer that is included in the kit and keep the final volume at 20µl. Keep DNA and primer amounts unchanged. Note that this is recommended only when you sequence with the same type of template and primer. If you use a new primer or make changes in the DNA template preparation, go back to a fourth quarter dilution of the kit and check the outcome of the sequence.

Optimal amounts of sequence primer and template DNA/ 24 cycles reaction

- plasmid DNA; 150-300 ng
- PCR fragments up to 200 bp; 1-3 ng
- PCR fragments 200-500 bp; 3-10 ng
- PCR fragments 500-1000 bp; 5-20 ng
- PCR fragments 1000-2000 bp; 10-40 ng
- PCR fragments ≥2000 bp; 20-50 ng

and

- 3,2 pmol sequence primer

How is the amount of DNA affecting the sequence result?

DNA template quality and the amount of template are very important in order to obtain a good DNA sequence.

Too little DNA:

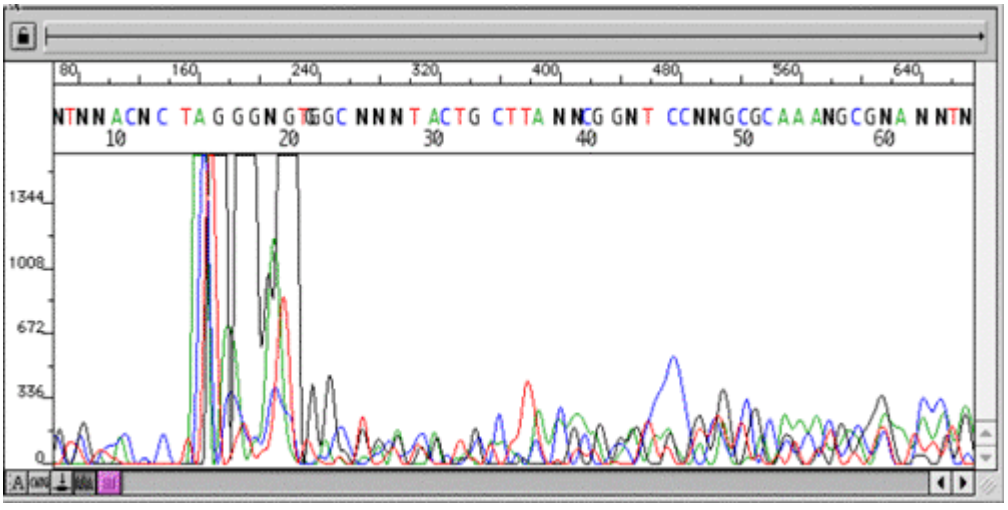
- weak and uncertain sequence with high background

Too much DNA:

- the reagents are used up too early resulting in top heavy data and abrupt signal loss
- pull-up peaks
- bad resolution of sequence peaks, resulting in “noisy” data
- delayed migration into the capillary, resulting in no data at all

2. COMMON SEQUENCE PROBLEMS

Weak or no sequence signals



Many different things might cause weak or failed reactions, for example;

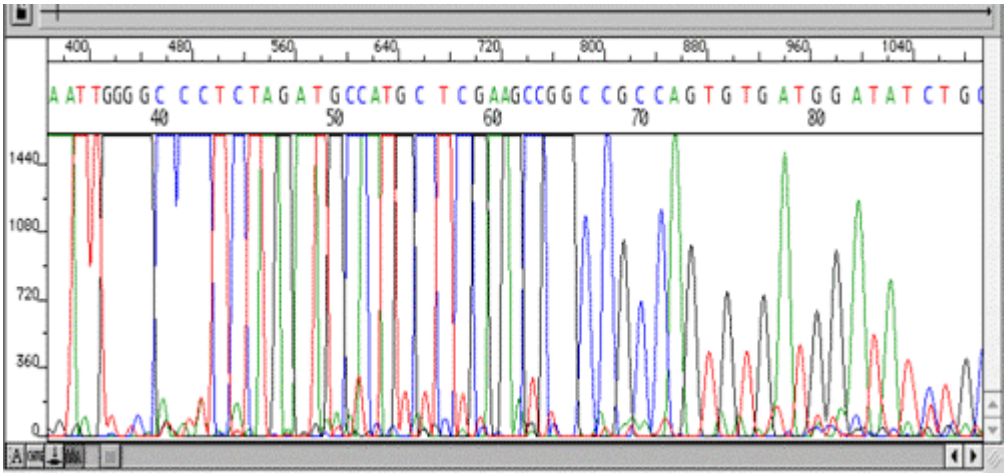
- Insufficient template

- Insufficient primer
- Poor primer design, e.g. low melting temperature
- Primer has no annealing site
- Contamination that inhibits the DNA polymerase (proteins, residual salts, residual organic chemicals, detergents, EDTA)
- Old, mishandled or missing reagents
- Extension products lost during reaction cleanup

Possible solution:
If the signal strengths are very low (under 50), the resulting data are often a mixture of true sequence and background noise. In order to yield more labelled fragments and hence stronger signals;

- Increase the number of cycles (add 5-10 cycles)
- Clean the template DNA
- Add more template and/or primer
- Do not solve the DNA template or primer in a buffer containing more than 0.1mM EDTA

Strong signals/ top-heavy data



Sequence data displaying top-heavy peaks and/or pull-up peaks are characterized by:

- very high peaks in the raw data trace
- very high peaks in the analysed data trace with pull up peaks and poor base calls
- very high signal to noise ratios S/N G:>1000 A:>1000 T:>1000 C:>1000

Pull-up peaks might be caused by:

- Incorrect estimation of template concentration (i.e. too much used)
- Too many cycles used in the sequence reaction

Possible solutions:

- Diluting the sample in deionised formamide and rerunning the sample can often correct this problem and yield good data
- Reduce the amount of template and/or cycles used in the sequencing reaction
- Use less of the sequence kit

Delayed migration

Sequence data displaying delayed migration show:

- Peaks are visable after the usual start point of 1000 to 1500 in the raw data trace
- Poor resolution of peaks and peaks that are not evenly spaced in the raw and analysed traces
- Poor base calls in the analysed electropherogram

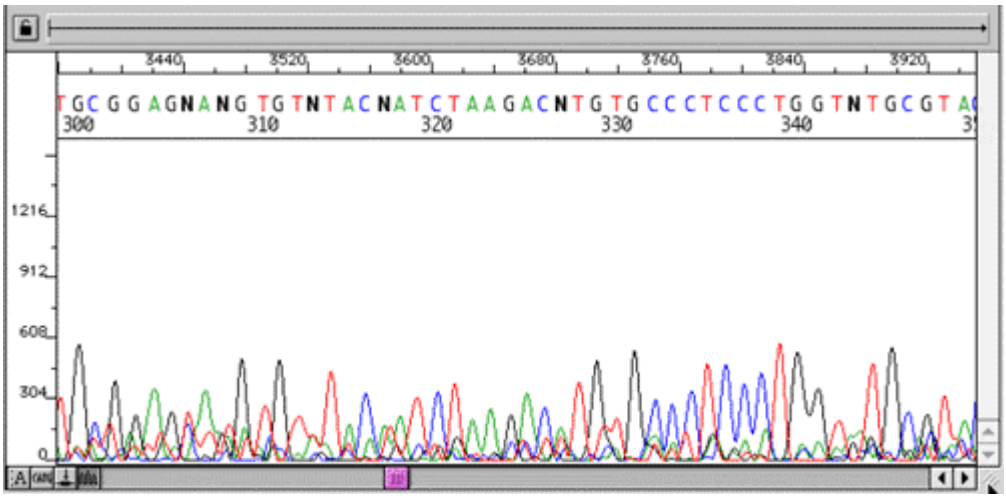
Delayed migration may be the result of:

- Contaminating negative ions (salts or other contaminants) in the sample being preferentially injected to the labeled fragments
- Heavily overloaded samples. Excess of template used during sequencing

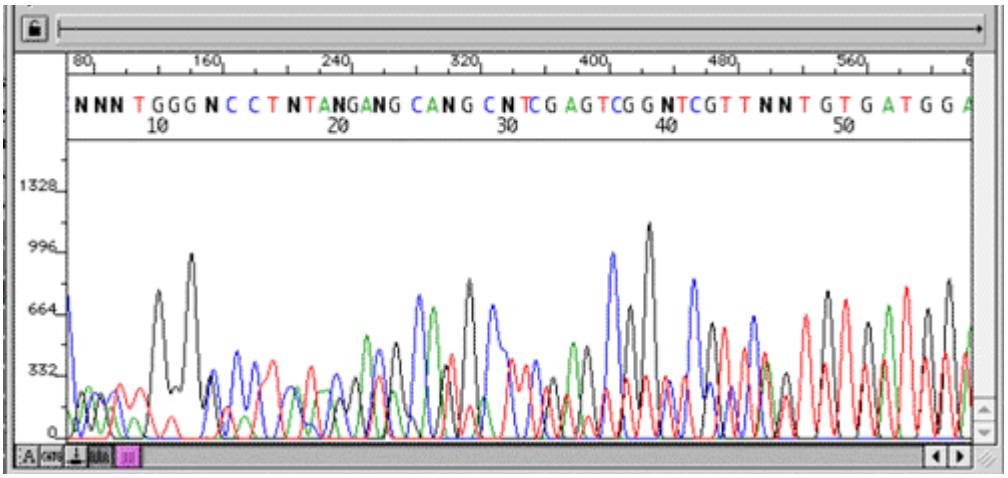
Possible solutions:

- Diluting the sample in deionised formamide and rerunning the sample can often correct this problem and yield good data. Please contact us if you wish to rerun the sample.
- Take care not to use too much template DNA, sequencing chemistry and/or cycles in the sequence reaction
- Clean the template DNA and sequence reaction in order to remove contaminants and excess salt

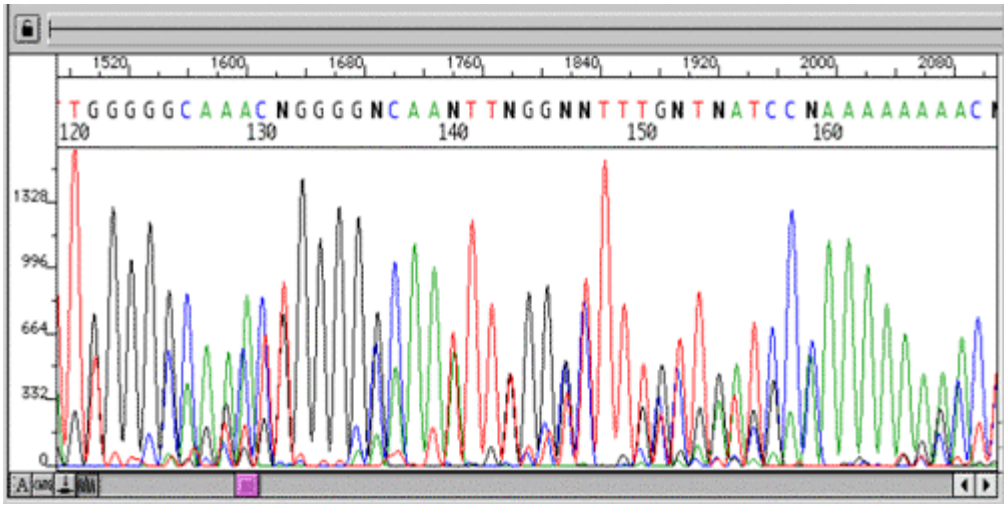
Noisy sequences throughout the sequence



Low signal strength.



Contaminating template or primer



N-1 primer

Noise throughout the sequence might be caused by:

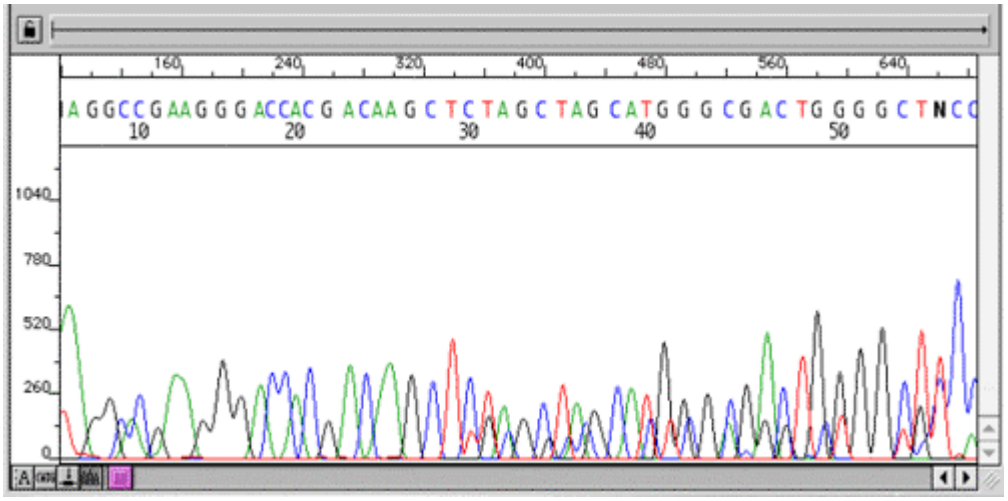
- Multiple primers
- Multiple templates
- Contaminated template
- Multiple primer binding site in the template
- The primer used in the reaction is contaminated with N-1 primer (primer that is one base shorter than the desired sequence primer) due to inefficient synthesis
- Low signal strength due to low template/primer concentration or impurities such as salt or ethanol
- High signal strength

Possible solutions:

- Check and optimize primer and/or template concentration
- Streak bacteria containing the plasmid of interest twice

- When sequencing PCR products, take care to remove excess primers from the PCR reaction
- To decrease the risk of obtaining N-1 primers; use primers that have been purified by HPLC

Noise up to or after a specific point in the sequence

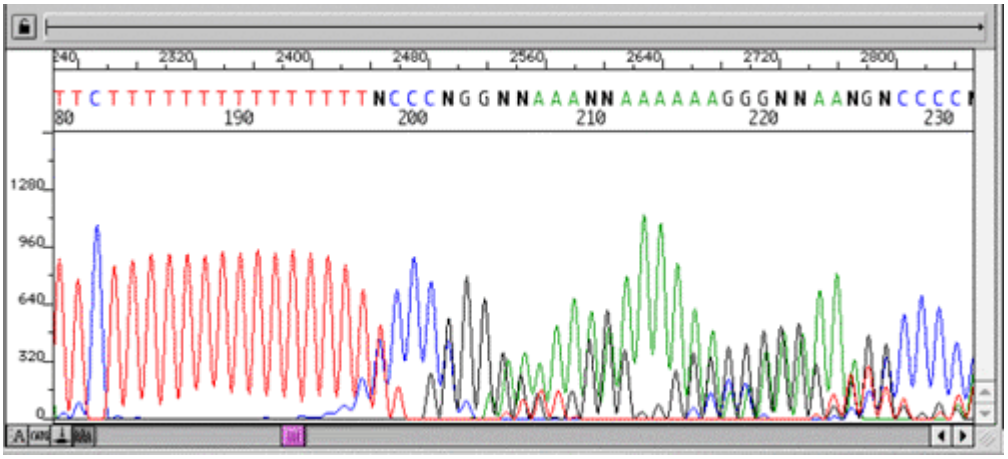


Noise up to or after a specific point in the sequence might be caused by:

- Mixed plasmid prep or PCR fragments
- Frame shift mutation
- “Slippage” of DNA polymerases after homopolymer region in template (see below)
- Low signal strength due to low template/primer concentration or impurities such as salt or ethanol

Possible solution:

- Streak bacteria containing the plasmid of interest twice.



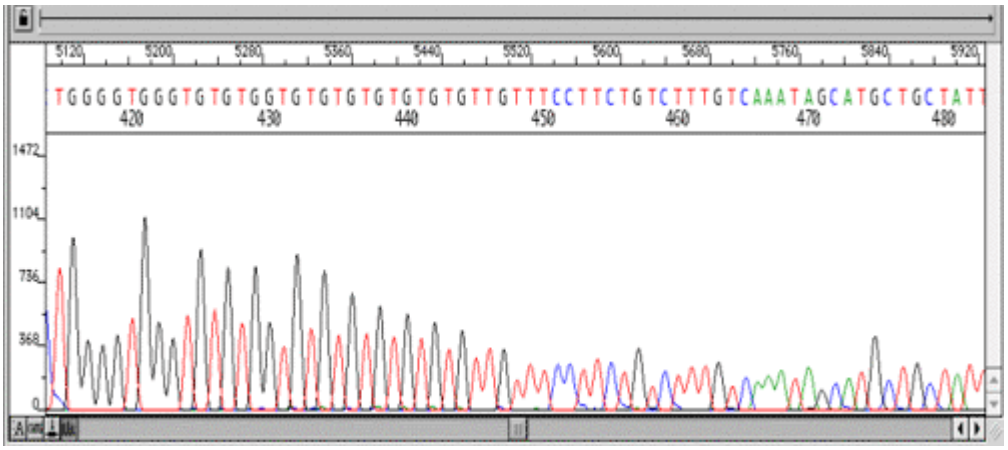
“slippage” of DNA polymerases

PCR amplification of DNA with long polymer regions of a specific nucleotide, such as poly A-tails, might result in a population of fragments with different length of N-stretches (the polymerase “slips” along the N's). When sequencing these fragments, the typical wavelike picture appears.

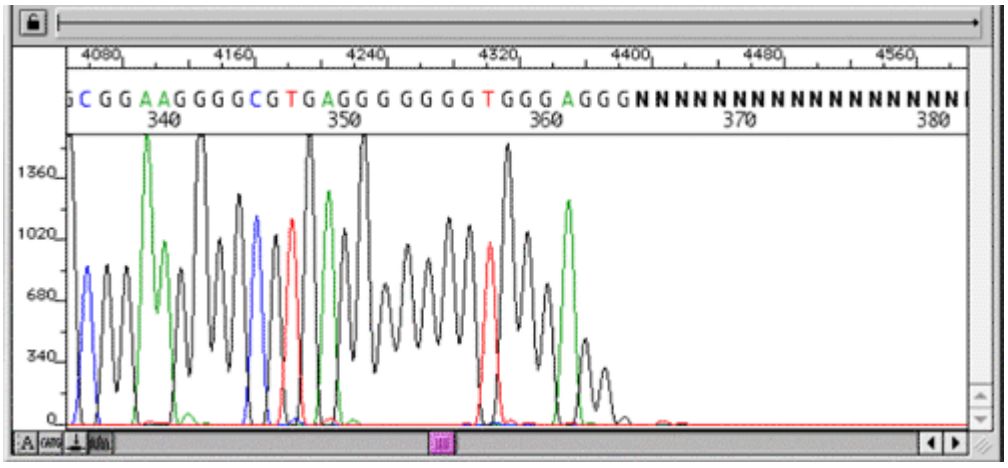
Possible solutions:

- Use an anchored primer, a sequencing primer that is for example polyT containing an A, C or G base at its 3’ end
- Sequencing through a poly-A tail in a plasmid usually works better than in PCR fragments

Early signal loss



Gradual signal loss



Abrupt signal loss

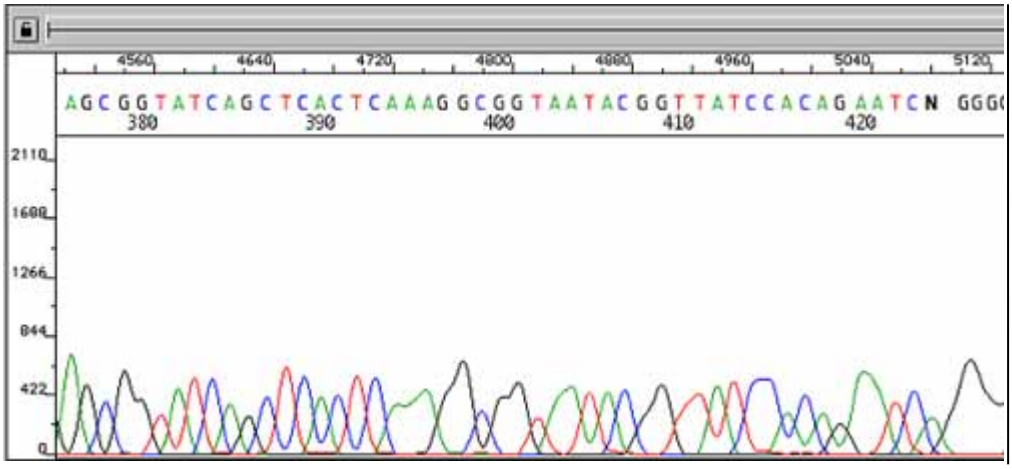
Signal losses have many causes, including the following:

- Region of secondary structure in the template
- G-C rich or A-T rich regions
- Template sequence idiosyncrasies

- Poor quantification of primer and/or template, leading to top heavy data
- Contamination with salt, ethanol, EDTA or other impurities

Possible solutions:
- Sequence opposite strand
- Try a primer that anneals in a different position
- Add DMSO to a final concentration of 5% or Betain to a final concentration of 1M
- Increase denaturation time and/or temperature
- Purify the template

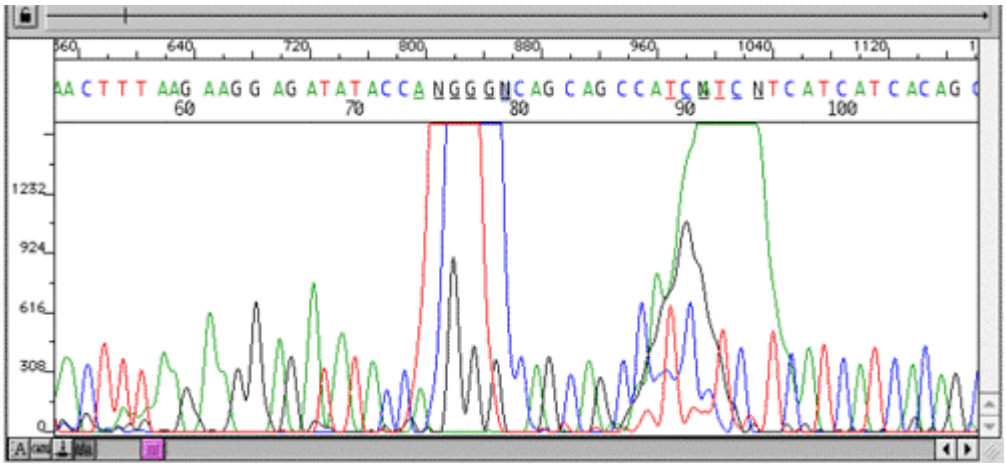
Poor resolution of sequence peaks



Poor resolution might be caused by
- overloaded lanes due to too much DNA in the sequence reactions (signals over 1000)
- contaminants in the template DNA

Possible solutions:
- Use less DNA
- Purify the DNA
  - Perform a phenol-chloroform extraction followed by a chloroform extraction. Continue with ethanol precipitation. This procedure usually improves sequence quality.
  - Pass the plasmids through a Sephadex G50 spin column
- Prepare new plasmids and take care not to overload the plasmid miniprep columns. When these columns are overloaded, many impurities from the bacteria might be left in the DNA
  - Make sure not to overgrow the bacteria culture.
  - Never use larger bacterial cultures than recommended for your kit. Preferably use less.

Excess dye peaks (“dye blobs”)



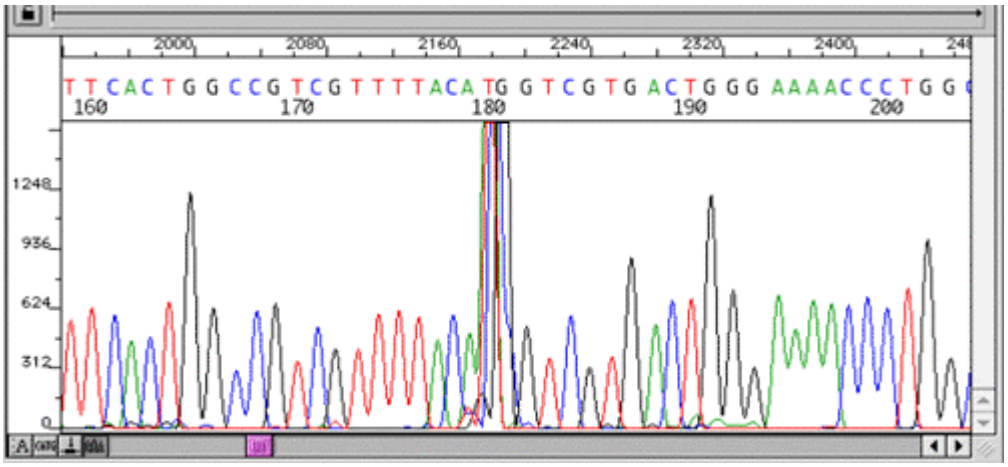
Dye blobs at beginning of sequence are caused by unincorporated labelled ddNTPs.

Solution:

- Remove the unincorporated dye terminators by purifying the sequence reaction according to the manual of the sequence kit or follow the ethanol precipitation protocol for a 20µl sequence reaction below

  - Add 2µl 125 mM EDTA, 2 µl 3M NaOAc pH5,3 and 50 µl 100 % EtOH.
  - Incubate on ice 10 min.
  - Centrifuge 10 min. at maximum speed in room temperature.
  - Remove supernatant.
  - Add 250 µl 70 % EtOH. Prepare fresh 70% EtOH solution weekly from 95 or 96% EtOH. EtOH solutions evaporate easily!
  - Centrifuge 5 min.
  - Remove as much of the supernatant as possible.
  - Dry the pellet (protect from light).

Spikes



"Spikes" might be caused by tiny air bubbles within the liquid polymer or by small pieces of dried polymer that have flaked off and entered a capillary.

Solution:

- Please contact us, and we will rerun the sample without charge.

3. SEQUENCE DATA

How do I interpret the signal strength values on the printout?

Signal strength is a number that indicates the intensity of the fluorescence for each nucleotide in the sample. Signal strength values from a balanced, well functioning sequence reaction lie in the range of about 100-900.

- Lower signals than 100 indicate a sequence reaction that has been inhibited, resulting in weak signals just above background levels.
- Stronger signals than 1000 indicate a sequence reaction that has been working "too well": producing a large amount of DNA fragments. The sample is overloaded, resulting in pull-up peaks and/or poor resolution of sequence peaks.

How long sequences can be read?

If the sequence reaction is optimal, with high purity and correct concentration of the template and optimal conditions for the primer, you can read around 700 bases using the ABI 3730 or 3130 XL machine and the BigDye™ terminator sequence kits (Applied Biosystems).

Why is the beginning of my sequence missing?

There are different explanations for this;

- When you design specific primers, position the primer at least 30 nucleotides upstream from where you want to start reading the sequence. If you need to read

short sequences and closer to your primer, use BigDye v. 1.1, that is specifically designed for sequencing short PCR products.

- The shortest fragments can sometimes be missing in the pellet, depending on the efficiency of the clean-up method, for example ethanol precipitation.
- The start point calculated by the analysing software is not optimal (usually depending on weak signals). The start point can be manually adjusted to retrieve the beginning of the sequence. Contact us about this.
- There are too many free dye-labelled ddNTP's present in the DNA pellet. These free nucleotides will show up as “blobs” in the beginning of the sequence, and we have to remove this part in order to get a readable sequence onwards. It is important to perform the ethanol precipitation and washing step according to the manual for the sequence kit. It is also important to remove all ethanol with a tip after the centrifugation steps.

How long time does it take before I get the results?

If we get the full-service samples (template DNA and primer) before 10.00 o’clock and the pre-made sequence samples before 13.00, you will get the sequence files later the same day or at the latest before noon the following working day.

Occasional delays may occur if we experience unusually heavy loads or equipment failure.