

Sequencing Troubleshooting Guide

DNA sequencing 반응은 가장 흔한 분자생물학적인 실험 방법임에도 불구하고 시료의 준비 상태나 DNA base composition 특이성으로 인하여, 기대치와는 다른 결과를 보이기도 합니다. 이는 동일한 sequencing 반응 상의 어떤 조작으로도 개선 가능성이 없으며, 가장 효율적인 방안으로 다음 실험을 계획하는 것이 합리적입니다.

저희 Takara Korea에서는 아래와 같은 Standard Program으로 단일조건 하에 sequencing 반응을 수행합니다. 제공해 주시는 primer의 Tm 값이 50~58 °C를 벗어나지 않도록 17~25 mer내로 디자인하여 주시기 바랍니다.

[Standard Program]

Denaturation : 95 °C 0:20, Annealing: 50 °C 0:15, Extension: 60 °C 1:30 (35 cycles)

Sequencing data에 대한 이해를 돕고자 흔히 발생하는 문제점들에 대한 양상을 정리하였습니다.

- Blank Lanes
- Weak Signal Sample
- Short Reads
- BAC or large template
- High GC contents/secondary structure
- Homopolymer Regions
- Slippage with PCR fragments
- Repetitive Elements
- Compression
- Template Contamination
- Primer Contamination(N-1mer)
- Multiple Binding

Blank Lanes

전혀 반응이 일어 나지 않은 경우로 초기 dye peak가 나오며, background를 peak으로 인식하여 noise 처럼 나타나게 되며, 이 경우는 데이터로서의 신뢰도는 전혀 없습니다. 간혹 basecalling이 되지 않는 경우도 있으며, 이런 반응 결과를 나타내는 주된 원인은 다음과 같습니다.

- Template의 양이 충분하지 못할 때
- Template의 primer annealing site가 없거나 부적당한 primer일 때
- Template나 primer의 정제상태가 좋지 못할 때
- Extremely high GC content나 template secondary structure

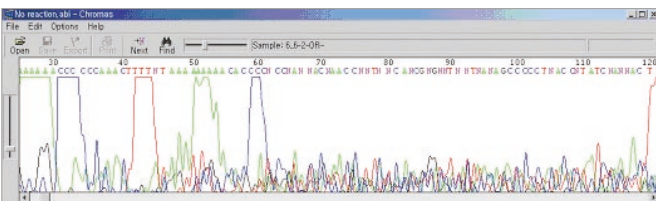


Fig 1. Blank lanes

Weak Signal Sample

Blank lanes와 동일한 원인에 의해 발생하는 것이나 그 원인에 의한 영향이 약한 경우 나타납니다. 또는 너무 적은 양의 primer나 annealing site binding activity가 낮은 primer로 반응 시 나타날 수 있습니다. 일반적인 signal 강도보다 미약하게 나오는 경우가 대부분입니다.

Short Reads

가장 일반적 원인으로는 dirty sample입니다. 만약 peak가 초기에 높게 시작하였으나 급격히 감소한다면 Sequencing PCR reaction이 salt나 유기물, EtOH 등 불순물로 인하여 저해를 받아 정상적으로 진행 되지 못하였거나, 과도한 template의 초기 유입으로 볼 수 있습니다.

이런 경우 template을 새로 정제하거나 primer를 정제하고 그것들을 다시 정량함으로써 개선 시킬 수 있습니다.

BAC or large template

Very large template의 sequencing에는 2가지 주요한 문제가 존재합니다. 한가지는 충분한 signal 강도를 얻을 수 없고 다른 하나는 완벽하게 template를 denature할 수 없다는 점입니다. BAC clone이나 cosmid의 경우 1 ug정도의 template이 필요합니다.

High GC contents/secondary structure

GC rich한 서열을 갖거나, secondary structure를 가지는 template의 경우 특정 부위에서 signal이 낮아지며, 보통의 경우 더 이상 polymerization 되지 못하거나 현저하게 signal이 줄어드는 양상입니다.

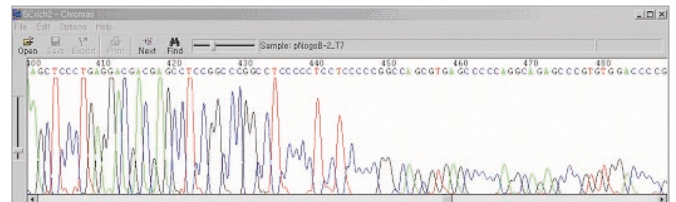


Fig 2. GC rich sequence

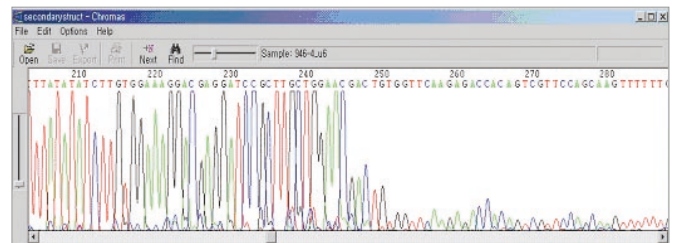


Fig 3. Secondary structure 형성으로 abrupt signal loss

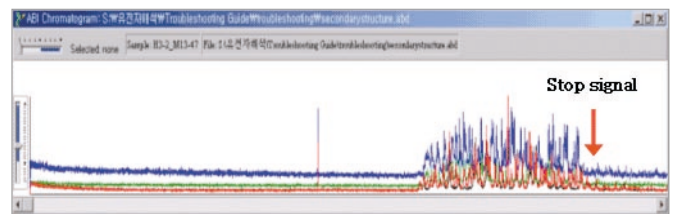


Fig 4. Raw data

Homopolymer Regions

Long poly A, T 서열이 존재하는 경우 "Slippage" 현상을 나타낼 수 있습니다. Slippage의 기작에 대해서는 정확하게 알려진 바가 없으나, homopolymer region이 존재하는 경우, polymerization시 두가지 DNA strand의 pairing이 정확하지 않아 나타날 가능성이 많다는 가설이 지배적입니다. 이 영역을 피하여 sequencing용 anchor primer를 제작하거나, 반대방향으로 sequencing하여 결과를 호전 시킬 수도 있습니다.

