LEIDEN UNIVERSITY MEDICAL CENTER

SV validation proposal

Jeroen F. J. Laros

Leiden Genome Technology Center

Department of Human Genetics

Center for Human and Clinical Genetics

Introduction

The samples

Around 1000 breakpoints selected by the SV group.

- The breakpoints have been fine mapped, so the location is known up to the nucleotide level.
- Allows for primer design close to the breakpoint.
- Allows for sequencing of the PCR product.

Introduction

The samples

Around 1000 breakpoints selected by the SV group.

- The breakpoints have been fine mapped, so the location is known up to the nucleotide level.
- Allows for primer design close to the breakpoint.
- Allows for sequencing of the PCR product.

We need a high throughput method for validation.

- Multiple samples.

Experimental design

Traditional approach

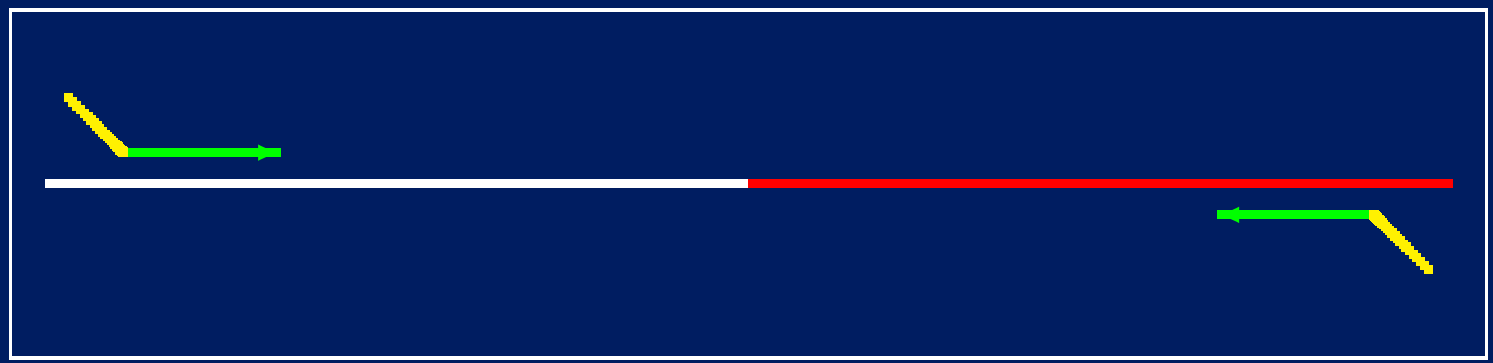

Figure 1: PCR the region of interest

Legend:

- White and red parts represent the regions flanking the breakpoint.
- Green parts represent the primers.
- Yellow parts represent the sequencing adapters.

Experimental design

Traditional approach

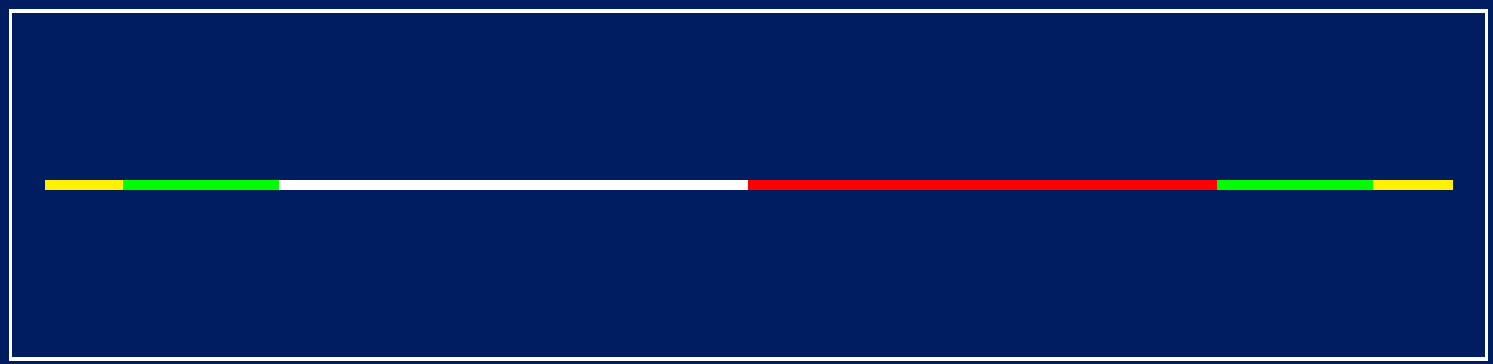

Figure 2: Product ready for sequencing

Not really high throughput or cheap.

- Cost: €2.5 to €5 per primer.
- Work: 1000 *separate* PCR reactions.

Experimental design

Experimental approach

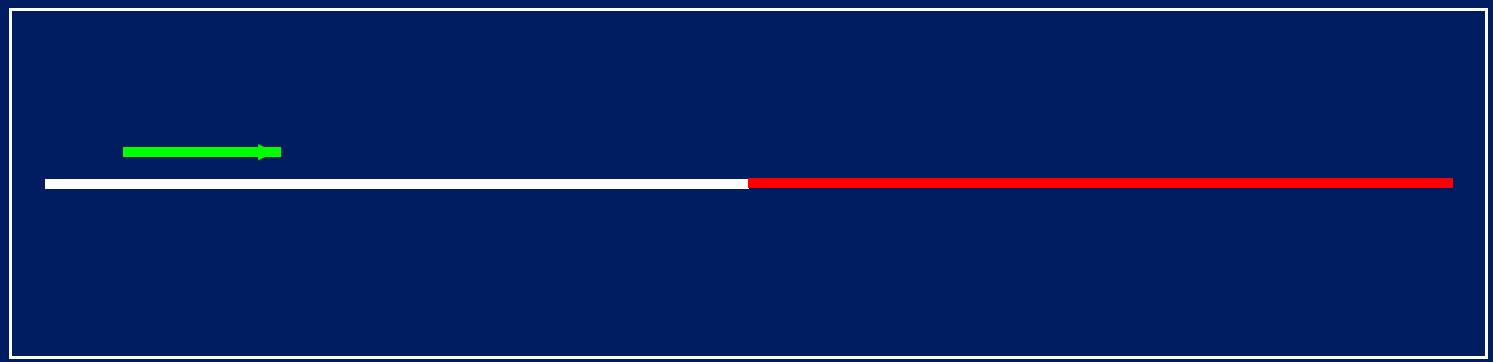

Figure 3: Primer extension

Step one.

- Design one primer at either side of the breakpoint.
- Extend the primer with a normal polymerase.

H.P.J. Buermans

Experimental design

Experimental approach

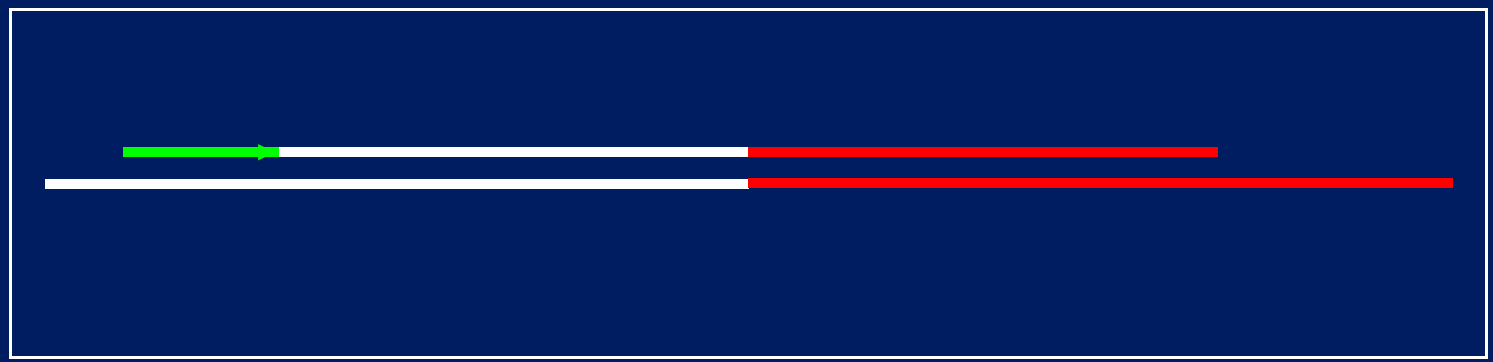

Figure 4: Extended primer

Step two.

- Do DNA end repair.
- Ligate adapters.

H.P.J. Buermans

Experimental design

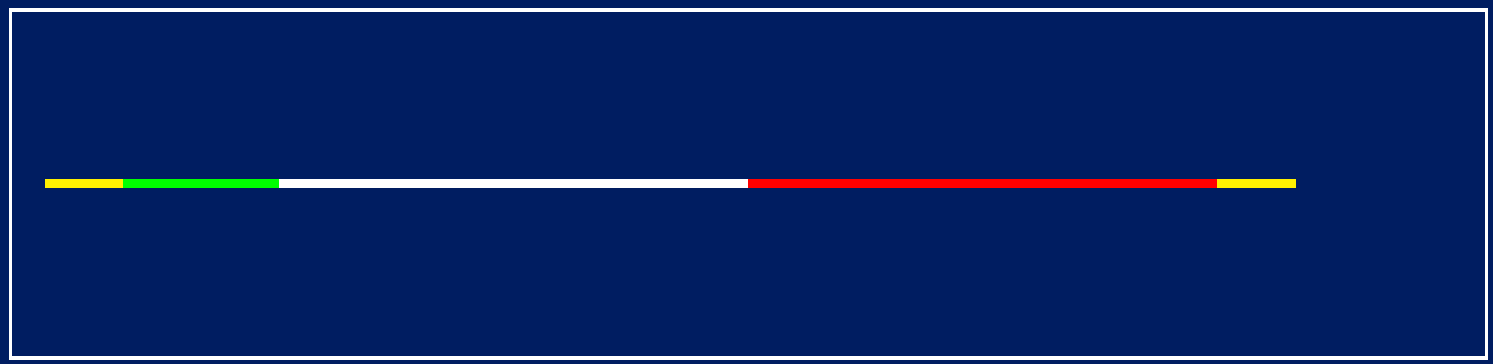

Figure 5: Product ready for sequencing

The primer design is done only once.

- Primers can be made with FlexGen.
  - Around 6000 primers can be made.
  - Multiple primers can be used per locus.
  - Cost around €1400.
- Sequencing on the Ion Torrent.
  - Around 10 samples can be multiplexed per run.
  - Cost around €600.

Data analysis

Combined analysis

For the data analysis, there are several options.

- Map the sequences and use Pindel.
- Do analysis on the raw data with a tool designed for STR analysis.
  - Find two anchors in the flanking regions.
  - Classify the region between these regions.
  - No mapping needed.
  - Capable of finding heterozygous events.
  - Capable of finding mosaicism.

K. Ye, J.W.F. van der Heijden

Questions?

Acknowledgements:

Henk Buermans
Jaap van der Heijden
Kai Ye
Johan den Dunnen