Implementation of nanopore whole-genome long-read sequencing in clinical genetic research

Emilie Boye Lester[1], Steffen Moeller-Larsen[1], Charlotte Brasch-Andersen[1], Christina Fagerberg[1] and Martin Jakob Larsen[1]
[1] Department of Clinical Genetics, Odense University Hospital, Odense, Denmark.

INTRODUCTION: Long-read sequencing combines the advantages of traditional cytogenetic methods with the base-pair-level resolution of newer molecular methods. This study is focusing on different types of structural variants to explore the capability of long-read sequencing to detect structural variants and fine map the involved breakpoints. A broad spectrum of variants has been sequenced using the PromethION platform, including different complex chromosomal rearrangements. Samples with known balanced and unbalanced translocations, tandem duplications, deletions, and inversions is included to test the potential of long-read sequencing in clinical genetic research. The study is ongoing, but preliminary results have demonstrated that long-read sequencing data enables fine mapping of break points that are inaccessible by current methods which enabled a more precise resolution of the mutations.

METHODE:

STEP 1: Extraction

STEP 2: Library prep

STEP 3: sequencing

STEP 4: Data analysis Including breakpoint analysis and genome assemblies

Product	Version
Device	PromethION P24
Flow cells	FLO-PRO112 (R10.4) FlO-PRO114M (R10.4.1)
Kits	Ligation Sequencing Kit (SQK-LSK112) (SQK-LSK114)
DNA extraction	Monarch® HMW DNA extraction kit for Cells

Figure 1 Method step by step. Step 1: Blood sample collection and DNA extraction using Monarch® HMW DNA extraction kit for Cells & Blood. Step 2: Library prep. using ligation sequencing kit. Step 3: DNA whole-genome sequencing on PromethION. Step 4: Data analysis and genome assemblies.

Detection of structural variation

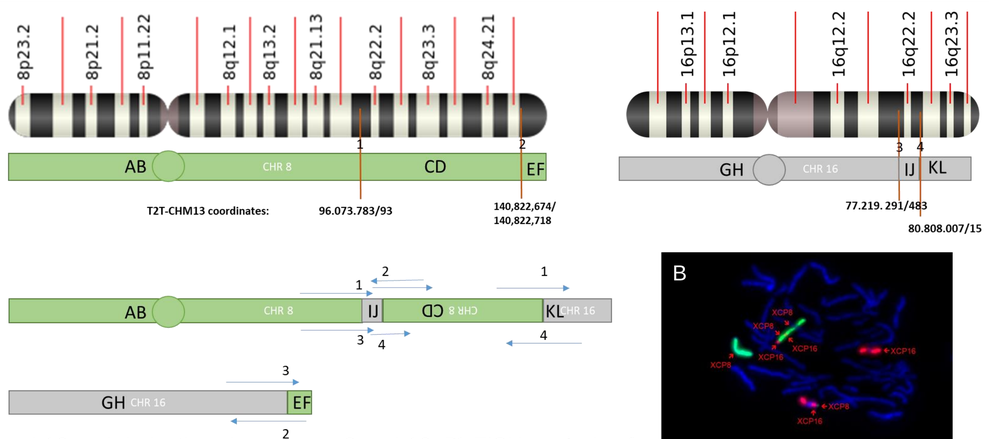
Data analysis	Guppy basecalling (SUP)
	Alignment: winnowmap and minimap2
	Telomere-to-telomere (T2T-CHM13) reference genome
	Mosdepth
	PycoQC
Call SVs	Sniffels2
	In house translocation detection software

Figure 2: In house translocation detection software flow. Visualize largest clusters using IGV.

Figure 2

RESULTS:

Sample 1 A)

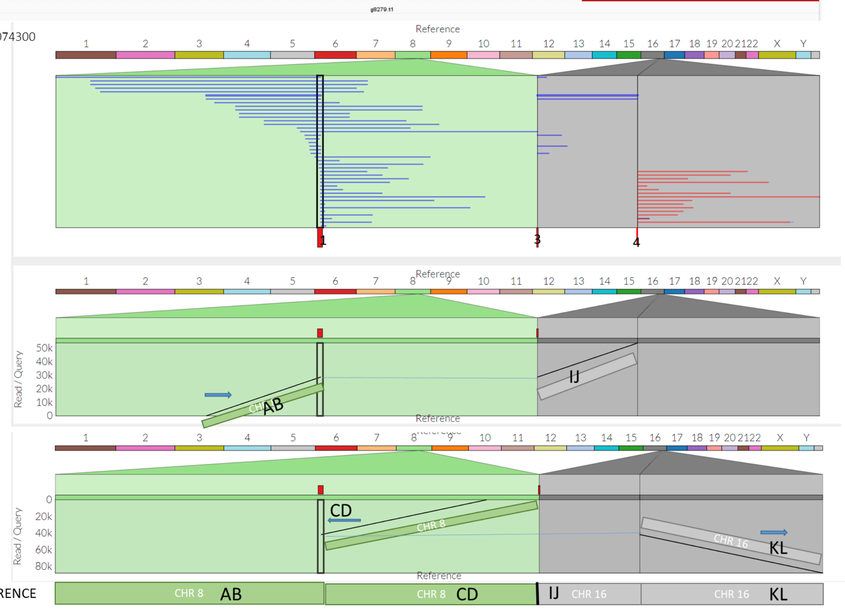


A) Chromosome 8 and 16 and derivative chomosome 8 and 16. B) FISH analysis of sample revealing complex chromosomal rearrangement between chromosome 8 and 16.

Snapshot from IGV: Visualizing the second breack on chromosome 8, including a 45 bp deletion and translocation to chr. 16.

Ribbon: BED file
chr8 96073000 96074300

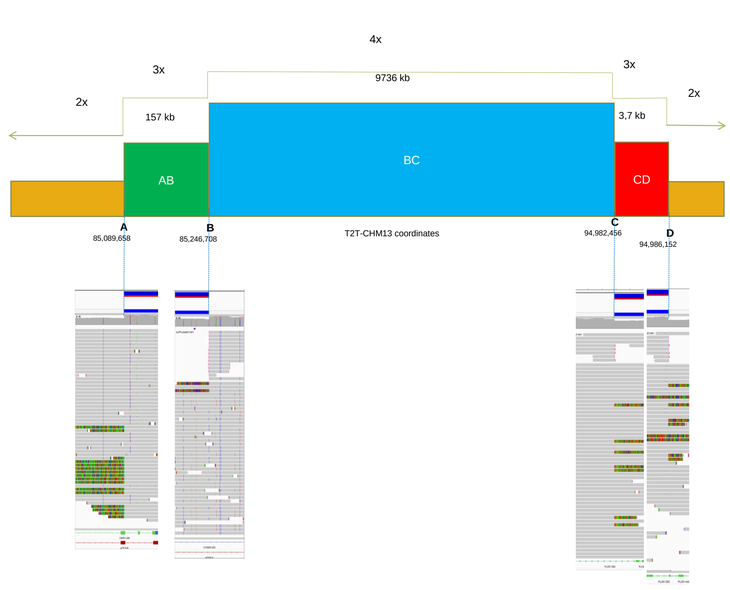
Dotpolt from genomeribbon.com Zoomed in on first break on chromosme 8 using Bed file that include the breakpoint.



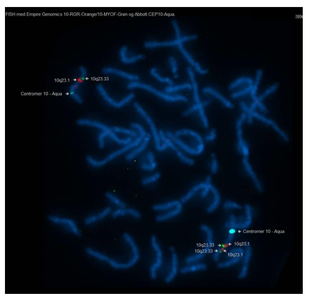
Results from sample 1; Balanced translocation (8;16). Visualisation of breakpoints using Integrative Genomics Viewer (IGV) and Genome Ribbon.

Sample	Flow Cell	Output Gb Q score >10	mosdepth (total)	PHREAD score	N50	Involved chromosome	SV type
1	R10.4	83 Gb	25.24	17.8	27 kb	t(8,16)	Balanced translocation
2	R10.4.1	101 Gb	30.78	20.24	17 kb	10	Triplication

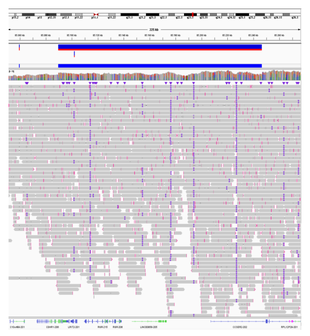
Sample 2



Sample 2: Sample with inversion duplication-triplication events. IGV screenshots from the 4 breakpoints and inversions called by Sniffels2 (vcf in blue). The BC block not identified as inverted by Sniffels2. Derivative chromosome 10 shown in the bottom of the figure.

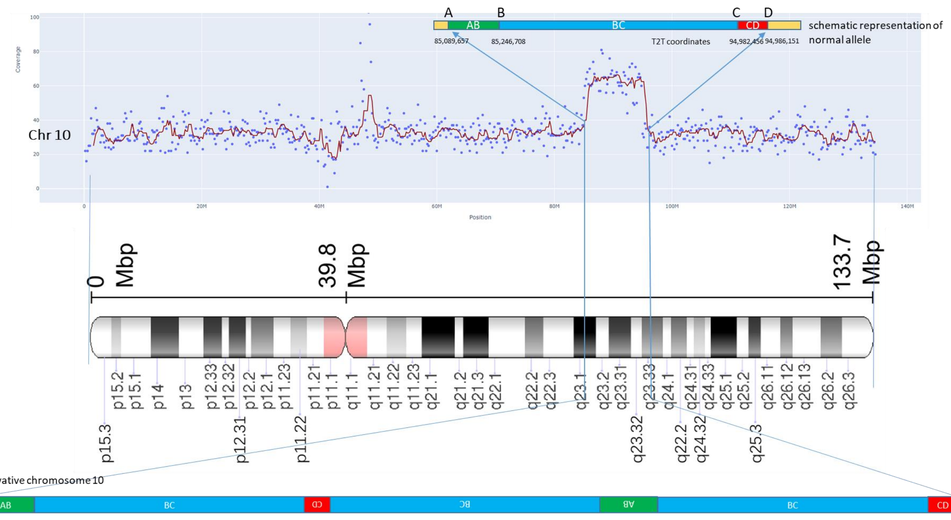


Inversion AB: 85089658-85246708



Inversion CD: 94982456-94986152

Sample 2: Seq[T2T-CHM13] der(10)(pter→q23.33::q23.33→q23.1::q23.1→q23.33::q23.33→qter)



sdu.dk

CONCLUSION AND PERSPECTIVE:

Using Nanopore Long-read sequencing for detection of structural variants and fine mapping of breakpoints has great potential. To optimize the new translocation tool we are working on filtering normal variation and we are also experimenting with local reassembly as this may increase and improve the detection and characterization of structural variants from areas that are challenged by repeat structures or multiple duplications. However, a successful reassembly is still dependent on correct locality assessment for the implicated reads, which may to various degrees depend on read length (the longer, the better) and the fidelity of the used alignment tool.

Print: Grafisk Center, SDU

Contact information: Emilie.boye.lester@rsyd.dk
Emilie Boye Lester