

Long-read sequencing reveals the complex structure of extra dic(21;21) chromosome and its biological changes

Kugui Yoshida-Tanaka¹, Ko Ikemoto¹, Ryoji Kuribayashi¹, Motoko Unoki¹, Takako Takano^{2,3}, and Akihiro Fujimoto¹

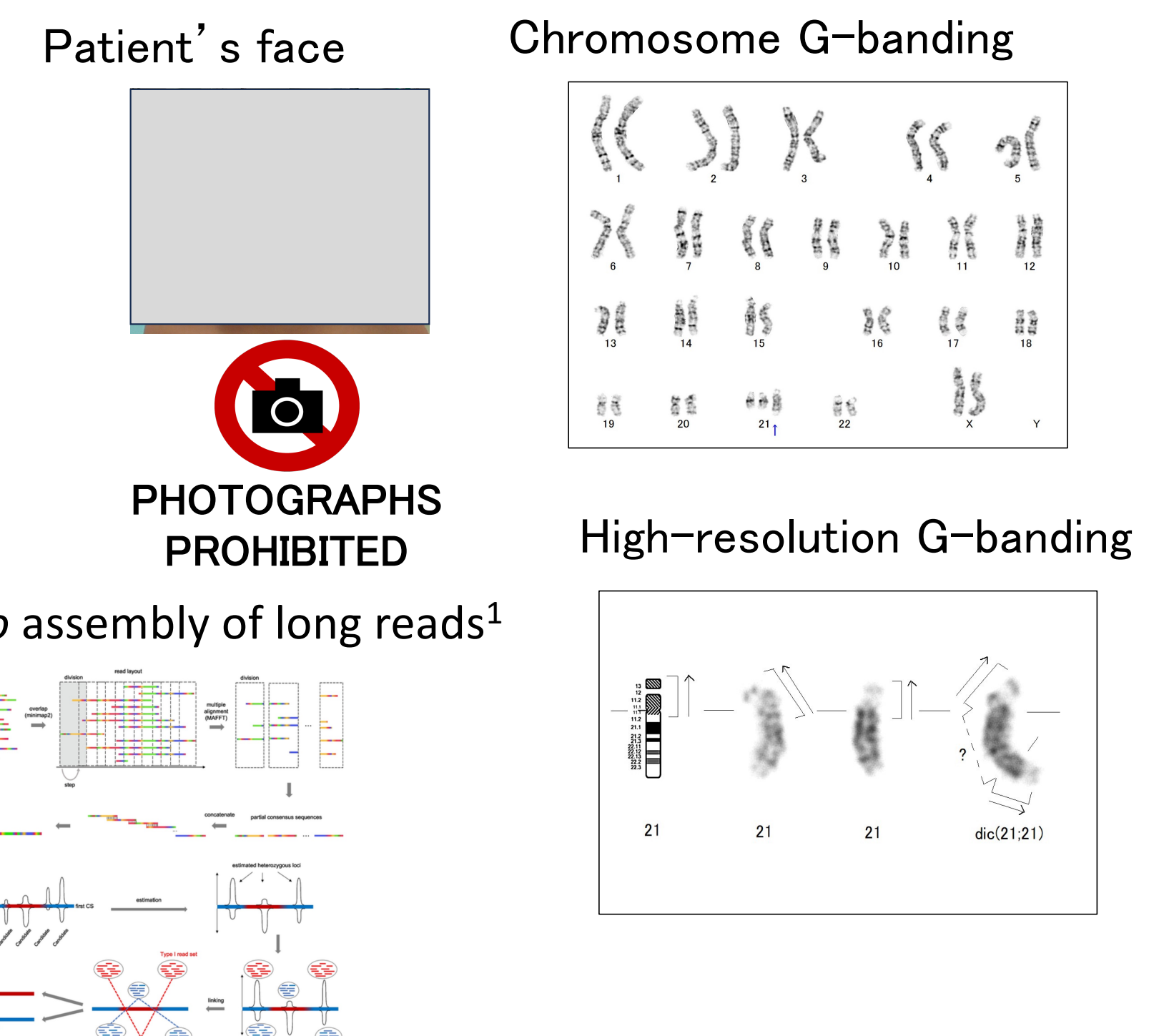
¹Department of Human Genetics, School of International Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; ²Department of Child Health, Tokyo Kasei University, Tokyo, Japan; ³ Tokyo Metropolitan Tobu Medical Center for Children with Developmental Disabilities, Tokyo, Japan
Corresponding author: Takako Takano and Akihiro Fujimoto

Abstract

Complex congenital chromosome abnormalities are rare but often cause severe symptoms. The structures and biological impacts of such chromosome abnormalities have seldomly been analyzed at the molecular level. Previously, we reported a Japanese female patient with severe developmental defects. The patient had an extra dicentric chromosome 21 (chr21) consisting of two partial chr21 copies fused together within their long arms along with two centromeres and many copy number changes. In this study, we performed whole-genome, transcriptional, and DNA methylation analyses, coupled with novel bioinformatic approaches, to reveal the complex structure of the extra chromosome and its transcriptional and epigenetic changes. Long-read sequencing accurately identified the structures of junctions related to copy number changes in the extra chr21 and suggested the mechanism of the structural changes. Our allele-specific transcriptome analysis showed the overexpression of genes in extra chr21. Additionally, allele-specific DNA methylation analysis of the long-read sequencing data suggested that the centromeric region of extra chr21 was hypermethylated, which would be associated with inactivation of one centromere in the extra chromosome. Our comprehensive analysis provides insights into the molecular mechanism underlying the generation of the extra chromosome and its pathogenic roles.

Background (Takano et al. EJMG (2020))

- A Japanese 12-year-old female patient
- At birth, her mother (40), father (44), and sister (2 years, 5 months) were all healthy.
- The patient exhibited severe psychomotor retardation and multiple dysmorphic features but was not diagnosed with Down syndrome.
- Chromosome analyses revealed that the patient carried an extra chromosome 21 (chr21), which was dicentric and had two stalks on both ends (47,XX,+dic(21;21)).
- An analysis of SNPs showed that extra chr21 was of maternal origin.
- Whole genome sequencing (WGS) using short reads revealed that chr21 had regions with four and six copies



Methods

To reveal accurate structure of extra chr21

- Analysis of allele frequency of SNVs in extra chr21
- Whole genome sequencing (WGS) using long read (Oxford nanopore) of the patient and her parents
- Fluorescence in situ hybridization (FISH)

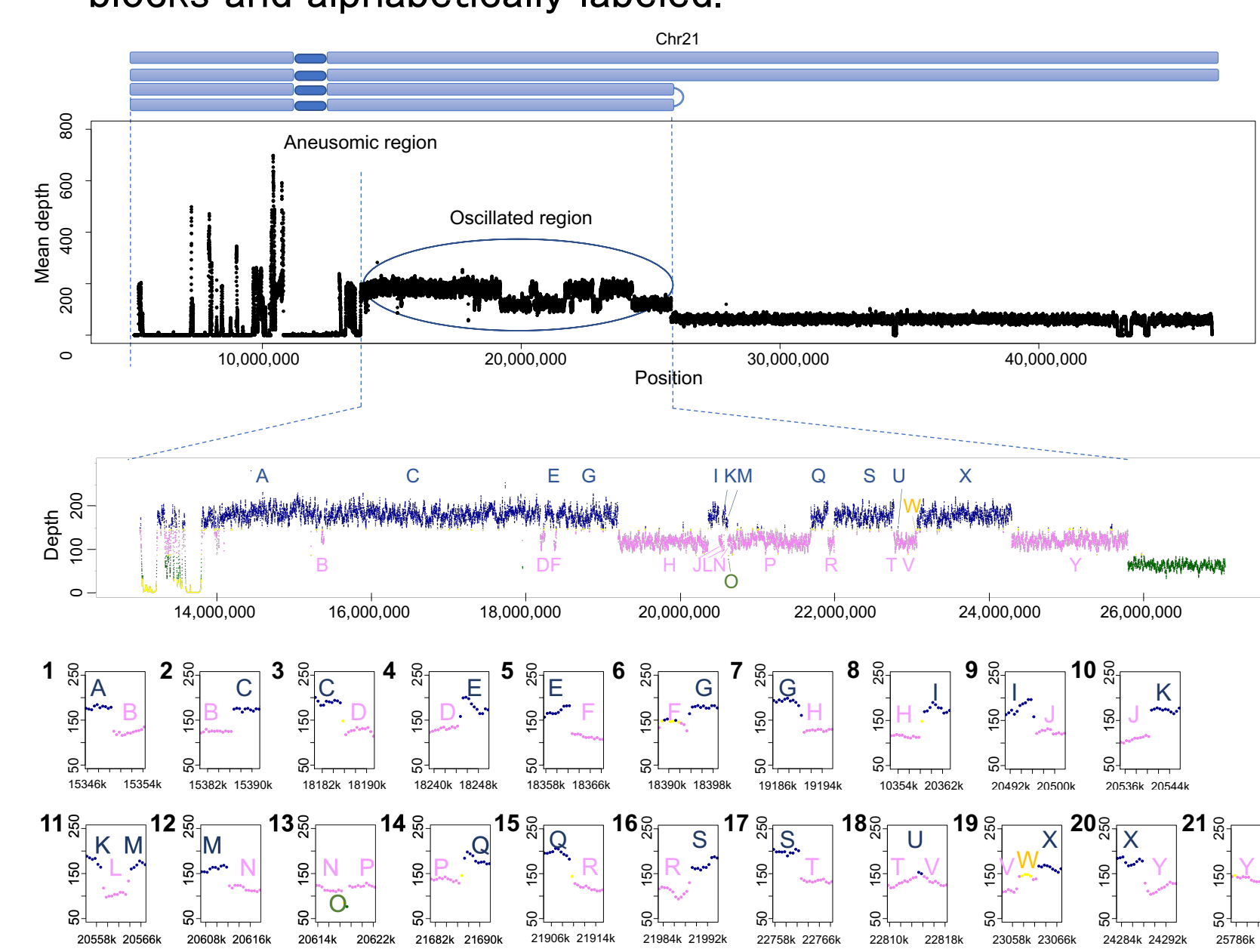
To understand pathogenic changes caused by the extra chr21

- RNAseq for blood samples (the patient and her mother) using short read (Illumina) and long read (Oxford nanopore)
- Allele specific methylation analysis using long read (Oxford nanopore)

WGS of the patients and her parents

Copy number of the patient's chr21

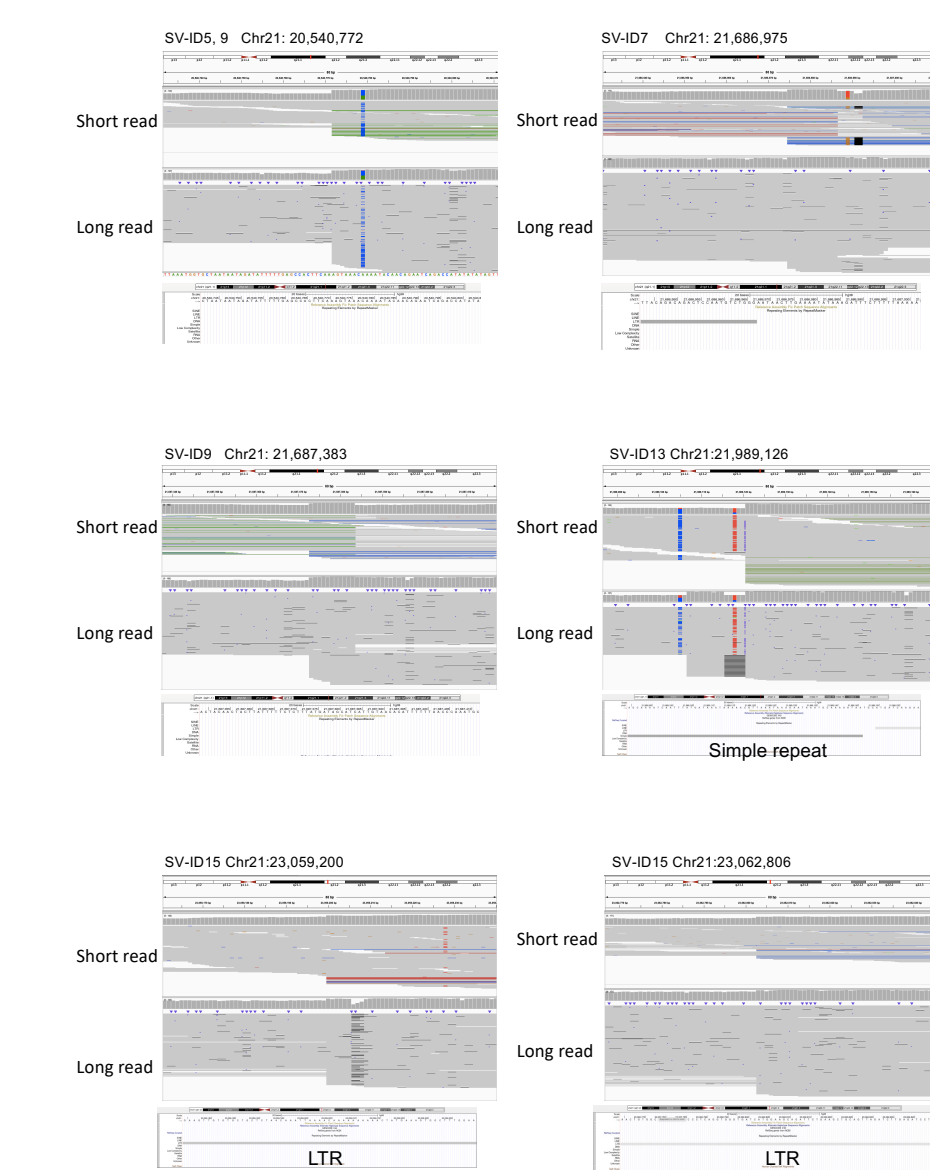
- Copy numbers of the chr21 were estimated from the number of mapped reads within 500bp bins.
- Regions between boundaries were defined as copy number blocks and alphabetically labeled.



Identification of SVs

- Fifteen SVs were detected, in which five were not detected by short reads.

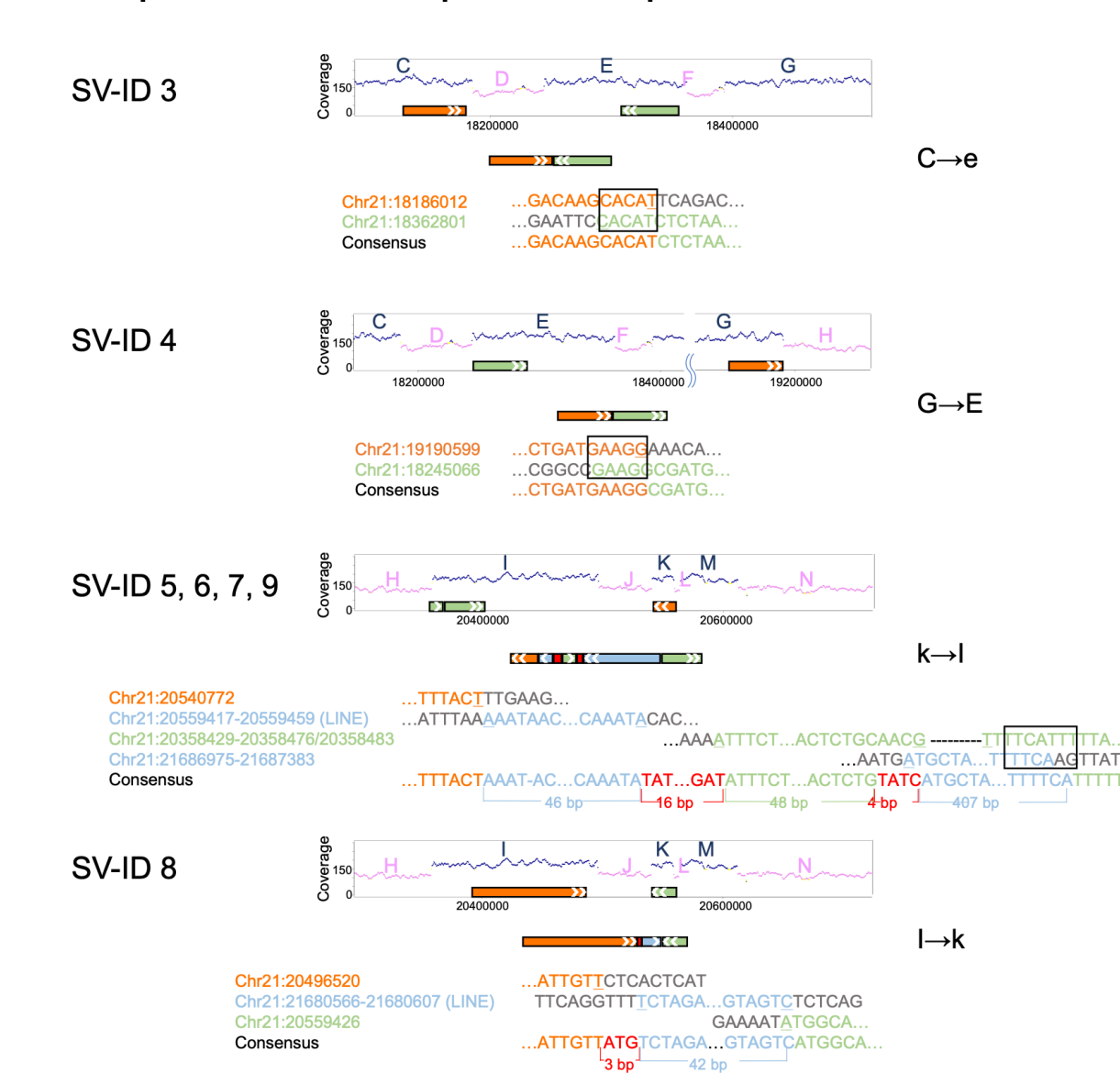
Visualization of novel SVs



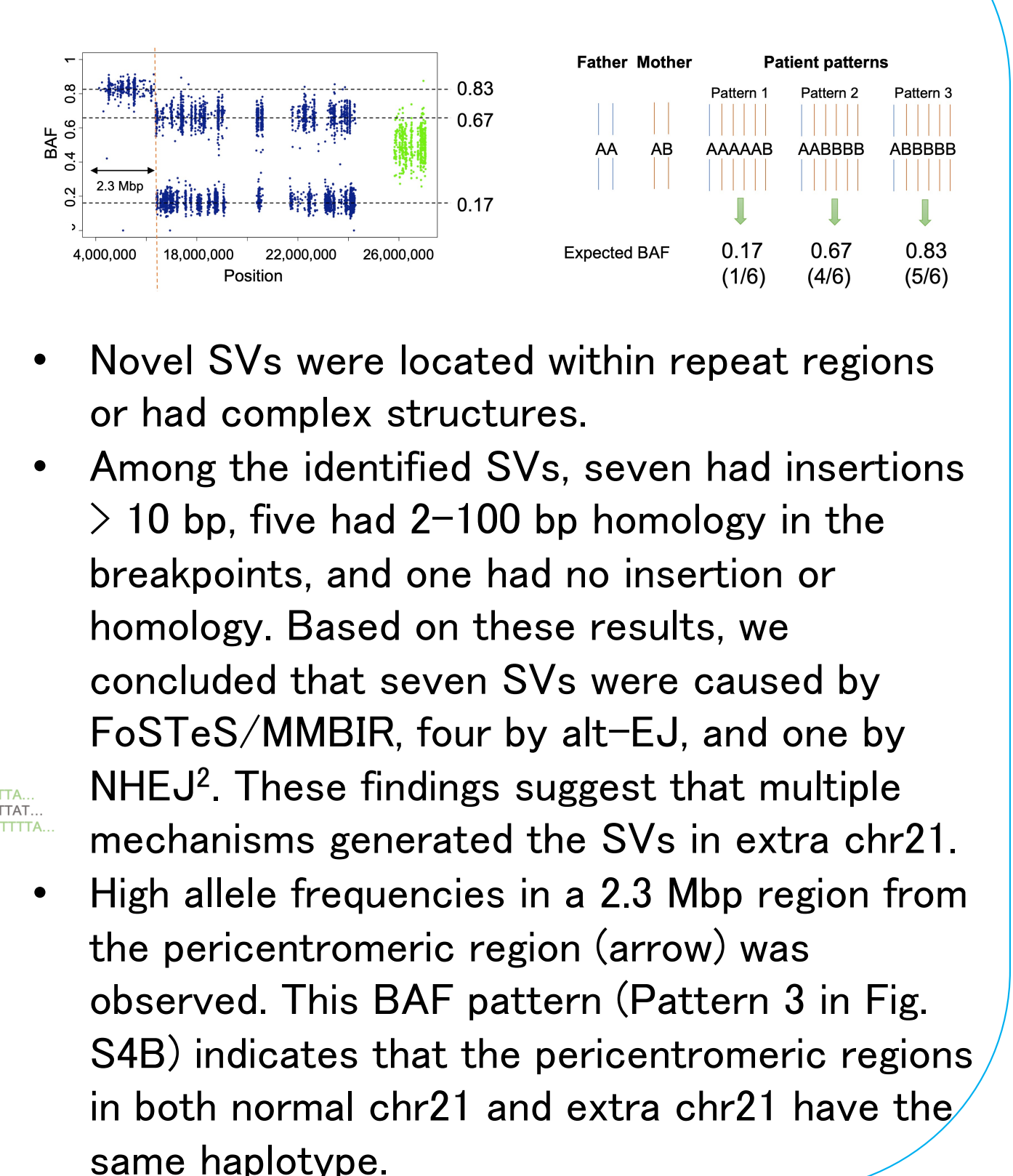
Structure of breakpoints

- Breakpoint sequences were reconstructed by *de novo* assembly¹.

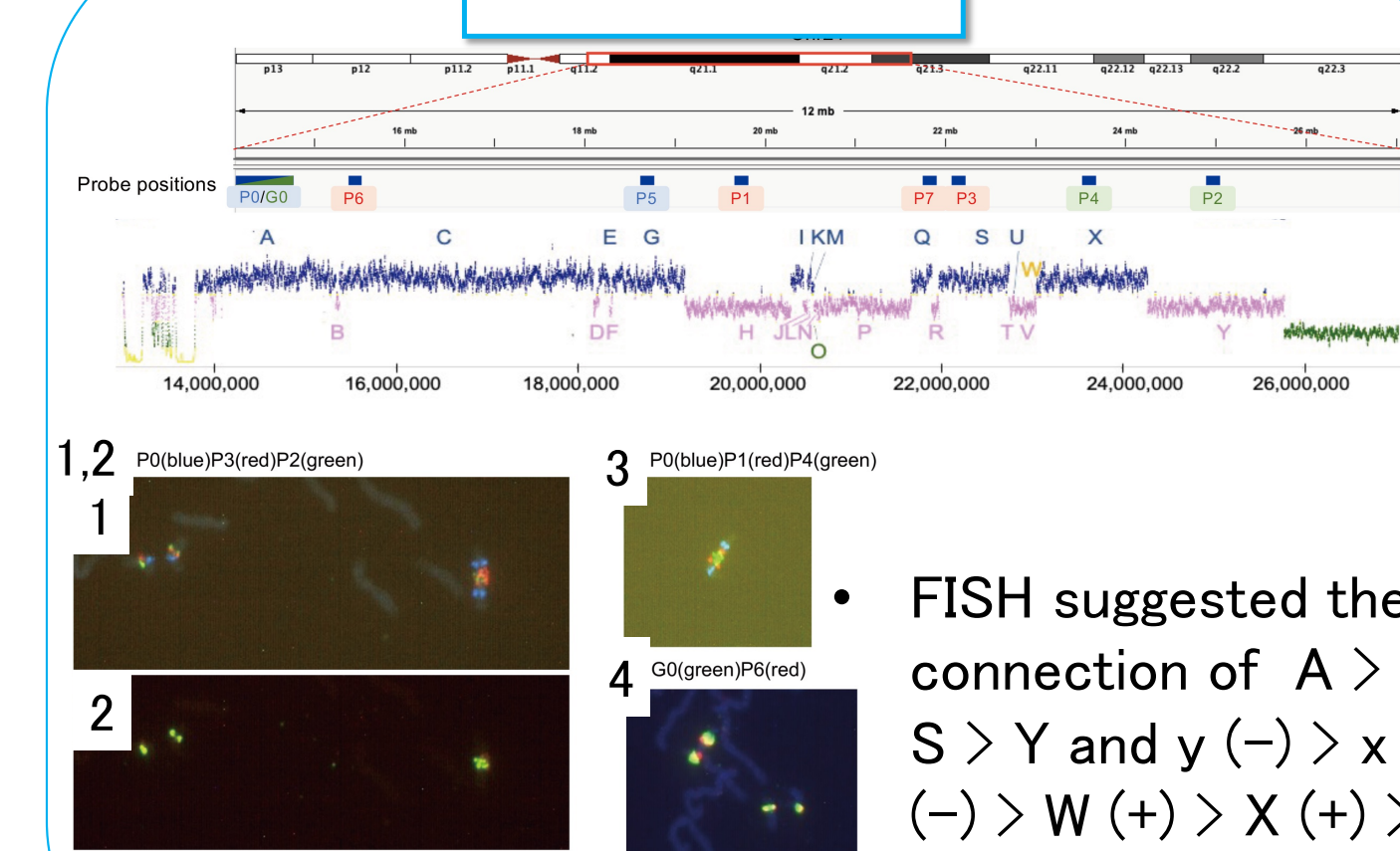
Examples of breakpoint sequence



Analysis of allele frequency of SNVs in chr21

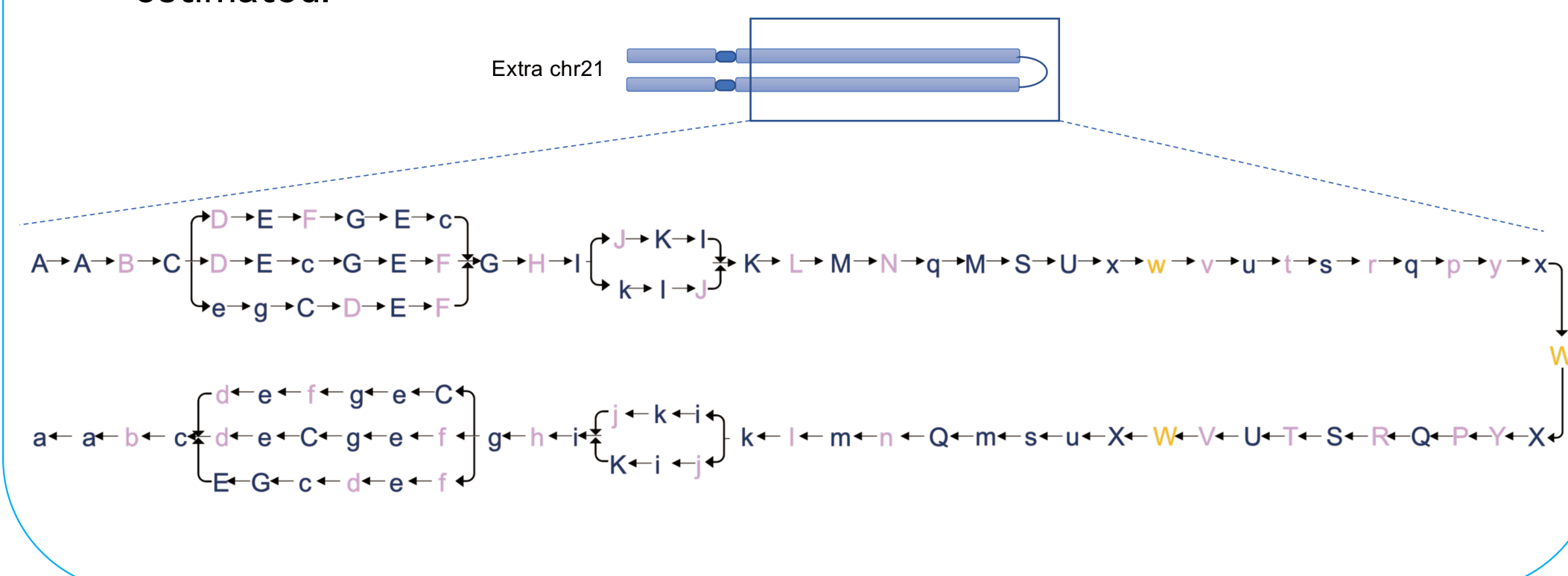


FISH

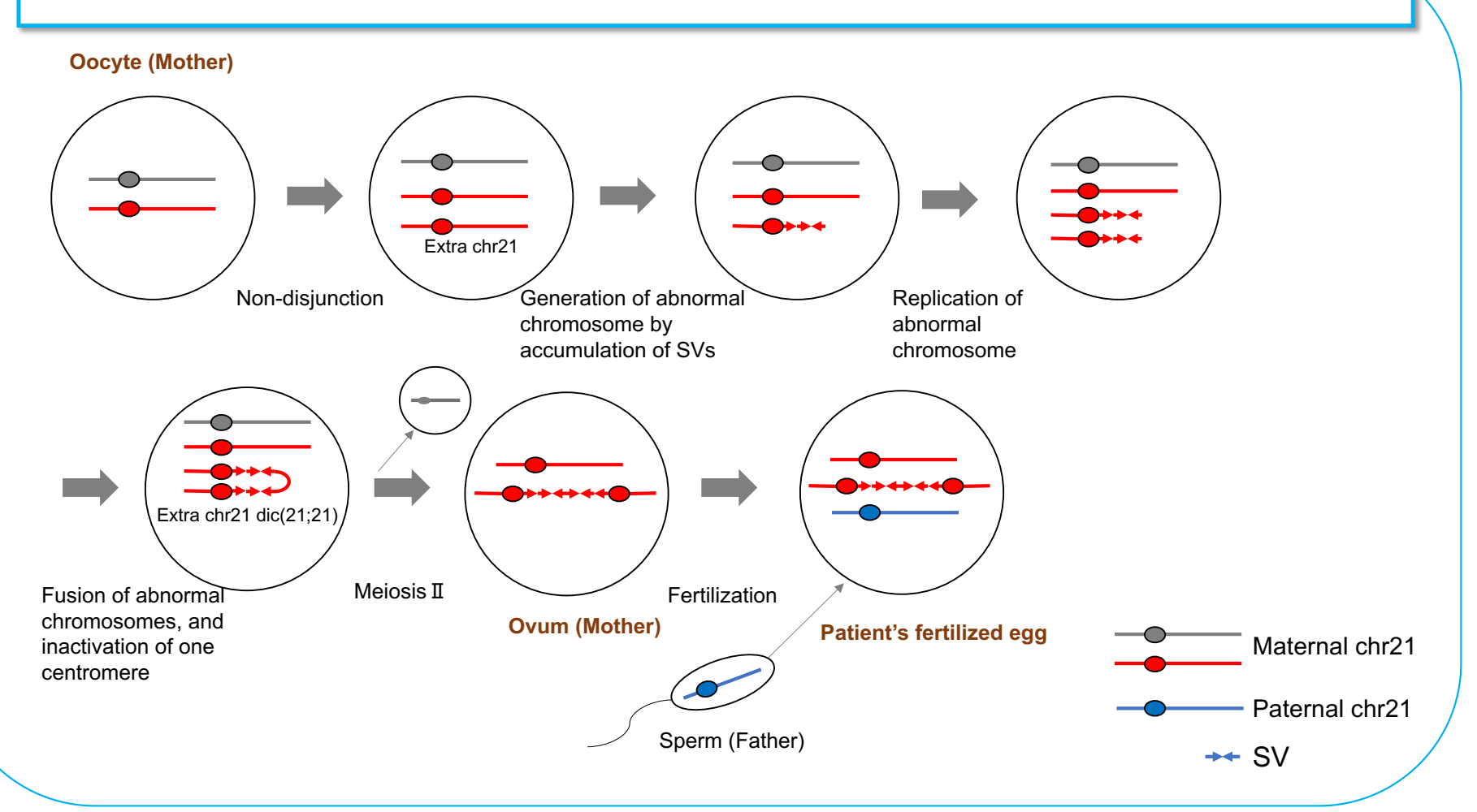


Estimated structure of extra chr21

- Based on the connections of SVs and FISH, structure of extra chr21 was estimated.

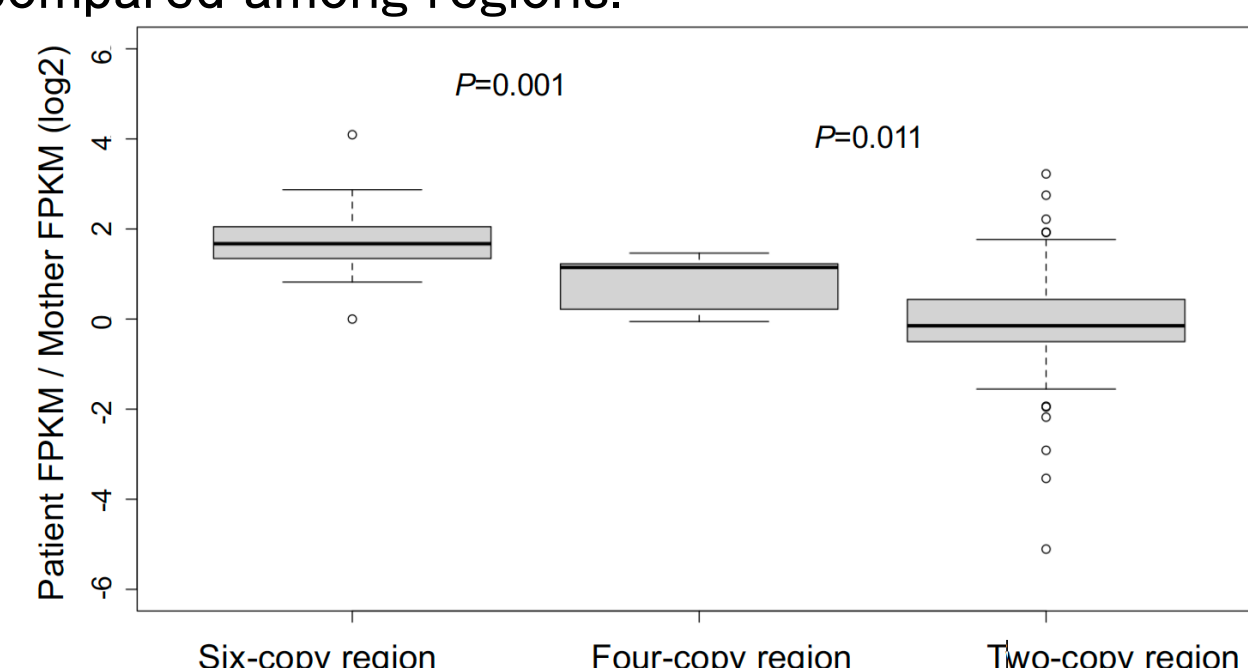


A model of the extra chr21 formation



RNAseq

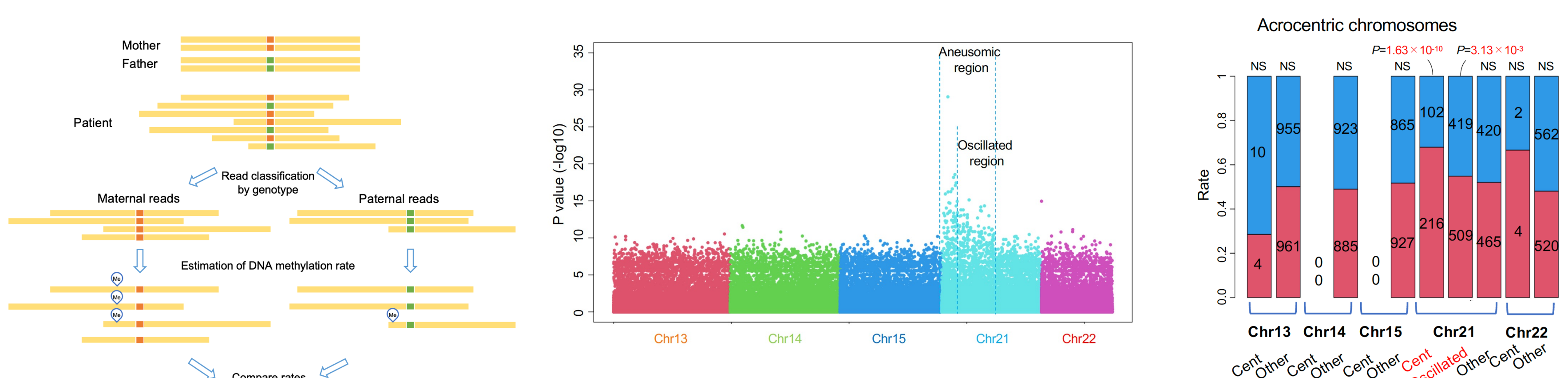
- RNA was obtained from blood samples of the patient and her mother.
- Ratio of expression levels (patient/mother) were compared among regions.



- The ratio was significantly higher in the six-copy regions than in the four-copy regions and in the four-copy regions than in the two-copy regions

Haplotype specific methylation analysis using long reads

- We classified long reads into paternal and maternal using heterozygous SNVs and compared the methylation rate of CpG sites.



- In chr21, many CpG sites exhibited low P values compared to those in other chr.
- A significantly higher number of maternally hypermethylated CpG sites were in the centromeric and oscillated regions.

Conclusion

- Long-read sequencing identified the structures of junctions related to copy number changes in extra chr21 and suggested the mechanism of the structural changes.
- The transcriptome analysis showed the overexpression of genes in extra chr21.
- Allele-specific DNA methylation analysis of the long-read sequencing data suggested that the centromeric region of extra chr21 was hypermethylated, which would be associated with inactivation of one centromere in extra chr.
- Our comprehensive analysis provides insights into the molecular mechanism underlying the generation of the extra chromosome and its pathogenic roles.