Nanopore sequencing reveals structural variants in early-stage mycosis fungoides

Carsten Hain[1], Cassandra Cieslak[2], Rudolf Stadler[2], Jörn Kalinowski[1]

chain@cebitec.uni-bielefeld.de

1 Center for Biotechnology, Bielefeld University, Sequenz 1, 33615 Bielefeld, Germany
2 Department of Dermatology, University Hospital Johannes Wesling Klinikum Minden, Hans-Nolte-Straße 1, 32429 Minden, Germany

Introduction

Mycosis fungoides (MF) is the most common form of cutaneous T-cell lymphoma with diverse disease progression, a lack of recurrent point mutations and a high prevalence of copy number variants (CNVs). As such, an outstanding importance of CNVs and the causal structural variants (SV) on tumorigenesis is proposed. Currently, further understanding of this disease is hindered by a lack of sequenced early-stage MF cases due to low tumour purity.

Structural variants in early-stage MF

We used CD3+ cells from an early-stage MF biopsy and performed whole genome sequencing using the PBK004 kit with a N50 of 3.8 kb. Structural variants (SV) were called using Sniffles and filtered with a custom script.

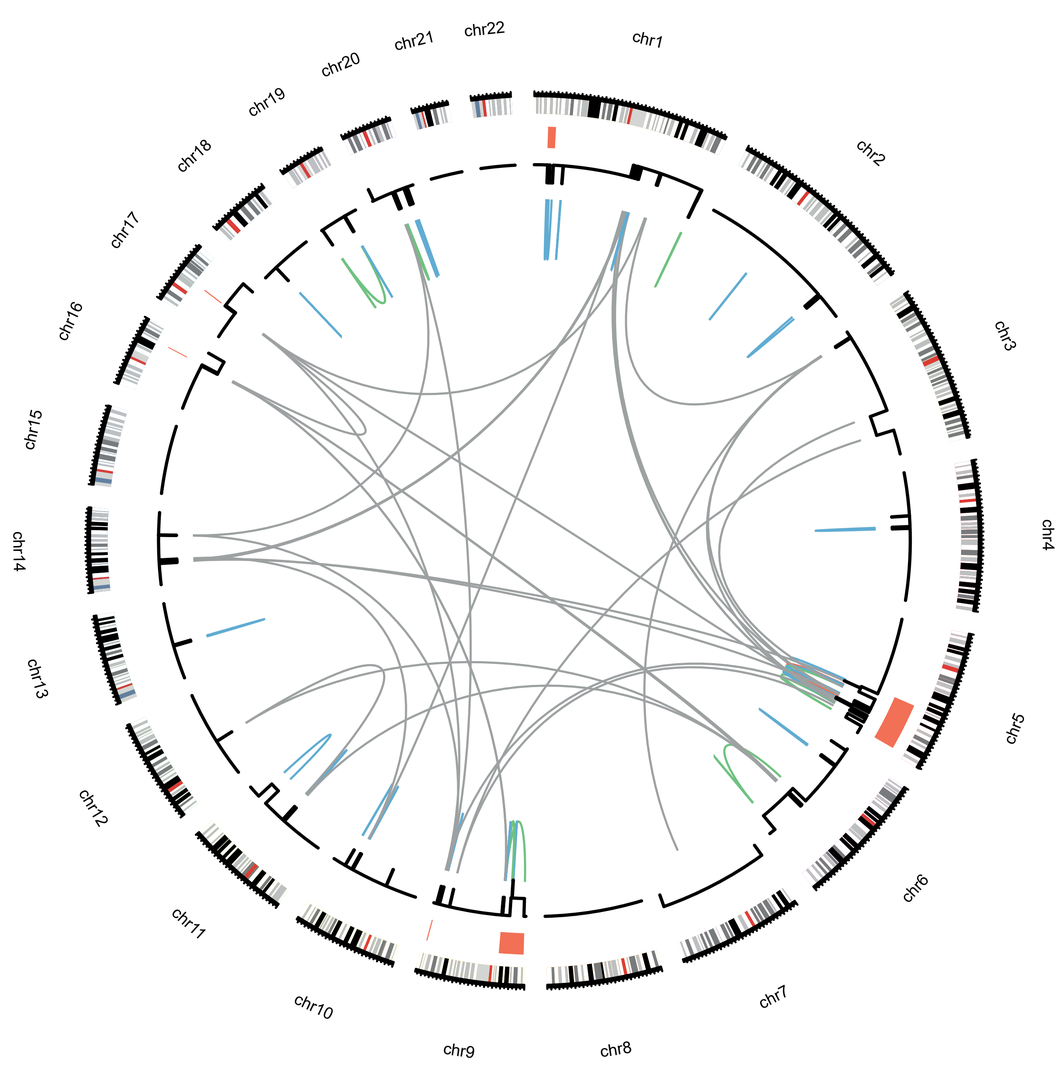

Figure 1: Structural variations in an early-stage MF sample detected by Nanopore sequencing. Circos plot of the copy number (black lines, outer ring) and the SVs detected by Nanopore sequencing (coloured lines, inner ring). Colouring of individual SVs indicates the SV class (deletions: blue, inversions: red, translocations: grey). The areas marked in red are elucidated in Fig. 2.

The sample shows various SVs, ranging from simple deletions to more complex interlocked SVs and to a chromoplexy event. Furthermore, multiple templated insertion SVs, where another genomic segment is inserted into the junction, are validated, with insert sizes up to 4 kb. Integrated analysis suggests further templated insertions with several hundred kb length. This observation is new to MF and might shed light onto the initial mechanism of SV generation in this disease.

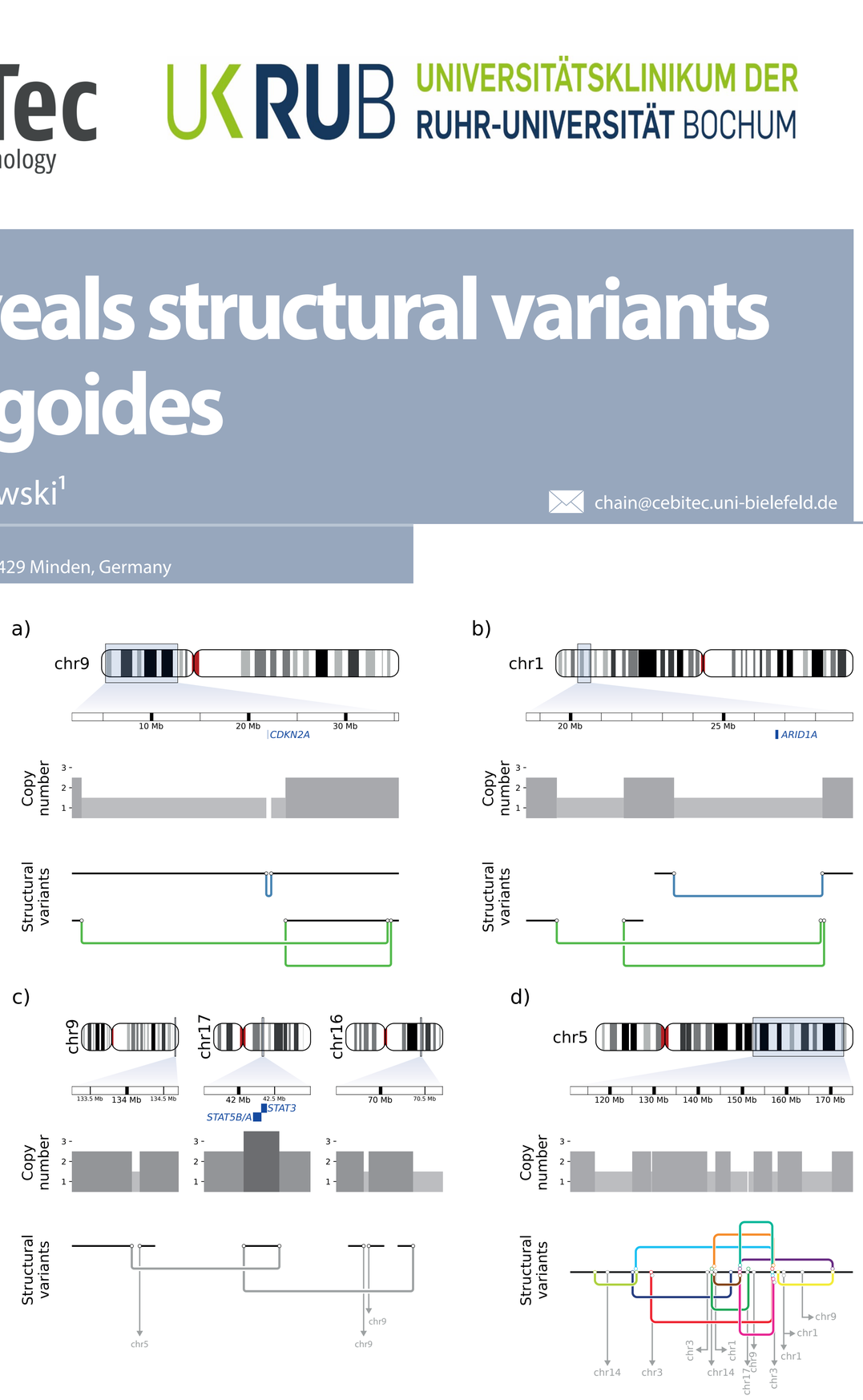

Figure 2: Structural variants leading to (a) homozygous *CDKN2A* deletion, (b) deletion of *ARID1A*, (c) amplification of *STAT3/5* and (d) chromoplexy on chr5. SVs are depicted as junctions between two breakpoints (junction colour dependant on SV type, see Fig. 1, except in (d)). Copy number is displayed as grey bars. Putatively affected genes are shown in blue.

Adaptive sampling

Adaptive sampling in 112 regions, with DNA directly from the unprocessed biopsy, was used for validating SVs even in low tumour purity context. Additional phasing information was used to confirm *in cis* localisation of individual SVs.

Figure 3: Adaptive Sampling detects SVs in low purity tumour samples and enables SV phasing. SVs causative for the *CDKN2A* deletion (Fig. 2a) as detected using PCR-WGS (top) or adaptive sampling (bottom). SV supporting reads are marked. Phasing shows *in cis* localisation of both SVs.

Nanopore sequencing is uniquely well suited for detection of multiple classes of SVs. This includes more complex mutations, like templated insertion SV, and enables phasing for allelic reconstruction of multiple SVs.

References

Choi, et al. Genomic landscape of cutaneous T cell lymphoma. Nat Genet. 2015 Sep;47(9):1011-9.
Martin, et al. WhatsHap: fast and accurate read-based phasing. 2016
Park, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). Blood 2017 Sep 21;130(12):1430-1440.
Sedlazeck, et al. Accurate detection of complex structural variations using single-molecule sequencing. Nat Methods 2018 Jun;15(6):461-468.