Sézary syndrome shows whole genome duplication as a late event in tumor evolution

Carsten Hain, Rudolf Stadler, Jörn Kalinowski

PII: S0022-202X(21)02500-8

DOI: https://doi.org/10.1016/j.jid.2021.11.009

Reference: JID 3218

To appear in: *The Journal of Investigative Dermatology*

Received Date: 23 June 2021

Revised Date: 4 November 2021

Accepted Date: 4 November 2021

Please cite this article as: Hain C, Stadler R, Kalinowski J, Sézary syndrome shows whole genome duplication as a late event in tumor evolution, *The Journal of Investigative Dermatology* (2021), doi: https://doi.org/10.1016/j.jid.2021.11.009.

TITLE PAGE

Sézary syndrome shows whole genome duplication as a late event in tumor evolution

SHORT TITLE

Sézary syndrome shows whole genome duplication

AUTHORS

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

AUTHOR AFFILIATIONS

1 Center for Biotechnology (CeBiTec), Bielefeld University, Bielefeld, Germany

2 University Clinic for Dermatology, Johannes Wesling Medical Centre, UKRUB, University of Bochum, Minden, Germany

CORRESPONDING AUTHOR

Name

Jörn Kalinowski

Address

Center for Biotechnology

Universitätsstraße 25

33615 Bielefeld

Germany

Tel.: +49 521 106 8756

Fax: +49 521 106 89041

Email: joern@cebitec.uni-bielefeld.de

ORCID FOR ALL AUTHORS

Carsten Hain: 0000-0002-6206-9906

Rudolf Stadler: 0000-0003-2683-6028

Jörn Kalinowski: 0000-0002-9052-1998

ABBREVIATIONS

SS: Sézary syndrome; CNV: copy number variation; SNP: single nucleotide polymorphism; SNV: single nucleotide variation; WGD: whole genome duplication

TO THE EDITOR

Sézary syndrome (SS) is an aggressive form of cutaneous T-cell lymphoma. Multiple studies analyzed the genetic basis of SS by whole exome sequencing and detected mutations in different pathways, ranging from signal transduction to epigenetic modification (Choi et al., 2015; da Silva Almeida et al., 2015; Ungewickell et al., 2015; Wang et al., 2015). Copy number variations (CNV) were especially abundant and recurrent. For instance, 75 % (Wang et al., 2015) to 92.5 % (Choi et al., 2015) of SS patients showed a TP53 deletion. By contrast, single nucleotide variations (SNV) showed recurrence of around 10 % and lower (Park et al., 2017).

Despite having gathered enough data over the years about individual CNVs and their predicted effects (Shao et al., 2019; Choi et al., 2015; Ungewickell et al., 2015; Wang et al., 2015) – as for example deletion and subsequent loss of function of TP53, amplification of STAT5 and ensuing hyperactivation of the JAK/STAT pathway – there still is a knowledge gap of SS tumor evolution regarding a comprehensive model on a genomic basis.

ABSOLUTE COPY NUMBER ANALYSIS SHOWS WHOLE GENOME DUPLICATION IN SÉZARY SYNDROME

We gathered 76 SS whole exome sequencing datasets from three studies (Choi et al., 2015; Ungewickell et al., 2015; Wang et al., 2015). After filtering for sufficient tumor purity and sequencing quality, we used 72 SS whole exome sequencing datasets for our analysis (Table S1). Exact diagnosis (e.g., primary SS, SS with prior mycosis fungoides) is given only for the samples from Wang et al., 2015. Other SS samples are described as stage IVA-B CTCL with a detectable abnormal population in the blood (Choi et al., 2015) or stage IVA-B SS (Ungewickell et al., 2015).

SNVs and CNVs were called for the selected datasets using respective GATK pipelines (van der Auwera and O'Connor, 2020). SNPs were called using GATK HaplotypeCaller with custom

filtering (Poplin et al., 2017). Based on this initial data, absolute copy numbers per samples were called using ABSOLUTE (Carter et al., 2012). Major and minor allele counts were added to the ABSOLUTE results manually.

For each sample, absolute copy number as well as measured and predicted SNP allele frequencies based of tumor purity, major and minor allele frequency were plotted (Figure S1). Samples showing serious deviations between the measured and modeled SNP allele frequencies were reanalyzed using ABSOLUTE with different parameters for mean ploidy until the SNP allele frequencies matched (detailed description in Supplementary Methods).

Aggregation of CNV data from all samples resulted in the typical CNV landscape for SS (Choi et al., 2015; Wang et al., 2015; Gug et al., 2019) with for example recurrently deleted and amplified arms on chromosome 17 (Figure 1). Out of 72 SS samples, 15 samples showed SNP pattern indicating a whole genome duplication (WGD) (Figure S1). Visual comparison of the CNV landscape of the non-WGD and the WGD samples showed similar pattern between both cohorts with virtually all recurrent SS CNV events present in both cohorts (Figure 1).

WHOLE GENOME DUPLICATIONS ARE A TEMPORAL LATE EVENT IN SÉZARY SYNDROME

Based on the ploidy of SNVs, amplifications and loss-of-heterozygosity events can be assigned a relative molecular time of occurrence during tumor evolution. This enables classification of CNV events into early or late events. This method is applicable only for CNV events with a sufficient number of SNVs. Here, timing of amplifications with three or more SNVs was carried out using MutationTimeR (Gerstung et al., 2020).

Overall, our analysis showed early timing for amplifications in non-WGD samples and late timing for the WGD events (Figure 2a). Detailed analysis for the most recurrent amplified chromosomes in SS (4, 7, 8, 9, 10 and 17) as well as timing for the WGD (Figure 2b) support this finding: WGD (mean molecular timing: 0.7) is a relative late event while amplifications of the chromosomes 4 (0.24), 10 (0.32), 17 (0.33) and 8 (0.35) are early events. The amplification of chromosomes 7 (0.53) and 9 (0.69) are intermediate or even late single chromosome events.

Using our results – the similar CNV landscape of patients with and without WGD, the early occurrence of chromosome-level CNVs and the late occurrence of WGDs – we assume a rough model for tumor evolution in SS: Therein, tumorigenesis in SS starts with one or multiple single chromosome(-arm) level CNV event(s). Afterwards, a variety of additional mutations are accumulated and the typical CNVs for SS emerge. In certain cases, a WGD occurs in a late stage of tumor evolution. This duplication happens in samples with extensive CNVs already present, thus generating the observed CNV profile similar to non-WGD samples, but with twice the ploidy.

This mechanism, albeit quite basic and unrefined, implies CNVs as the initial tumor mutations in SS. This proposed order – initial CNV(s), later SNVs and more CNVs – fits to the previous established lack of recurrent SNV as well as high recurrence and importance of CNVs in SS

(Choi et al., 2015; da Silva Almeida et al., 2015; Park et al., 2017; Ungewickell et al., 2015; Wang et al., 2015).

To further this model, additional questions should be considered: What is the molecular timing of the recurrent deletions (e.g., 10, 17p) as they are not timeable with this method? What CNVs are sufficient to give rise to tumorigenesis into SS? How do (albeit few) samples without CNVs fit into this model? Do patients with prior mycosis fungoides show different genomic pattern than patients with primary SS? To tackle these questions, an analysis of a longitudinal SS cohort, including detailed information about the primary diagnosis, similar to the TRACERx project (López et al., 2020), would be interesting to create evolutionary roadmaps for the development of this disease.

This analysis is important for targeted therapies: Early CNVs are likely clonal and present in all tumor cells. Later CNV events are likely subclonal and therefore not found in all tumor cells and lesions. Thus, therapies targeting genes in late CNVs would not be effective against all tumor cells. In this respect, a longitudinal analysis could identify early and clonally mutated genes which are hence universally targetable.

Finally the WGD as the terminal feature of the model, should be brought into biological context for SS. A review of other cancer types showed WGD timing ranges from early (e.g. in NSCLC (López et al., 2020)) to late (e.g. in melanoma (Birkeland et al., 2018; Gerstung et al., 2020)). WGD is often associated with tumor cell diversity, accelerated cancer genome evolution and worse prognosis (López et al., 2020). Additional work in characterization of the WGD group in SS samples with respect to their pathology, prognosis and their previous therapies is needed to infer the molecular role of WGDs in SS.

DATA AVAILABILITY STATEMENT

Datasets related to this article are hosted at dbGaP under the accessions phs000913, phs000725 and SRP058948.

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

ACKNOWLEDGMENTS

The authors acknowledge a research grant from Bielefeld University ("Fonds zur Förderung transdisziplinärer, medizinrelevanter Forschungskooperationen in der Region OWL"). We thank Marina Simunovic for critical reading of the manuscript.

AUTHOR CONTRIBUTIONS STATEMENT

Conceptualization: CH; Data curation: CH; Formal analysis: CH; Funding acquisition: JK, RS; Investigation: CH; Methodology: CH; Visualization: CH; Supervision: JK; Writing - original draft: CH; Writing - review & editing: CH, JK, RS

Figure 1: Copy number variation landscape of Sézary syndrome. The CNV landscape of 72 SS cases, separated in non-whole genome duplicated (WGD) and WGD, plotted against the genomic coordinates. Neutral copy number is 2 in the non-WGD and 4 in the WGD samples. Red areas above this level indicate amplifications and blue regions below this level indicate deletions.

Figure 2: Molecular timing of amplification in Sézary syndrome patients. The amplifications of 72 SS patients were timed with MutationTimeR (Gerstung et al., 2020). This included all amplification with 3 or more SNVs as well as all whole genome duplications. Results are plotted against their genomic coordinates (a) or aggregated for the WGD event and the most recurrently amplified chromosomes (b). Initial diagnosis for individual patients (prior mycosis fungoides, primary SS) is indicated if this data was available.

References

Birkeland E, Zhang S, Poduval D, Geisler J, Nakken S, Vodak D, et al. Patterns of genomic evolution in advanced melanoma. Nat Commun 2018; 9(1): 2665.

Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 2012; 30(5): 413–21.

Choi J, Goh G, Walradt T, Hong BS, Bunick CG, Chen K, et al. Genomic landscape of cutaneous T cell lymphoma. Nat Genet 2015; 47(9): 1011–9.

da Silva Almeida AC, Abate F, Khiabanian H, Martinez-Escala E, Guitart J, Tensen CP, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. Nat Genet 2015; 47(12): 1465–70.

Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, Rosebrock D, et al. The evolutionary history of 2,658 cancers. Nature 2020; 578(7793): 122–8.

Gug G, Huang Q, Chiticariu E, Solovan C, Baudis M. DNA copy number imbalances in primary cutaneous lymphomas. J Eur Acad Dermatol Venereol 2019; 33(6): 1062–75.

López S, Lim EL, Horswell S, Haase K, Huebner A, Dietzen M, et al. Interplay between whole-genome doubling and the accumulation of deleterious alterations in cancer evolution. Nat Genet 2020; 52(3): 283–93.

Park J, Yang J, Wenzel AT, Ramachandran A, Lee WJ, Daniels JC, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). Blood 2017; 130(12): 1430–40.

Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, van der Auwera GA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples, 2017.

Shao X, Lv N, Liao J, Long J, Xue R, Ai N, et al. Copy number variation is highly correlated with differential gene expression: a pan-cancer study. BMC Med Genet 2019; 20(1): 175.

Ungewickell A, Bhaduri A, Rios E, Reuter J, Lee CS, Mah A, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. Nat Genet 2015; 47(9): 1056–60.

van der Auwera G, O'Connor BD. Genomics in the cloud. Sebastopol, CA: O'Reilly Media, 2020. 467 p.

Wang L, Ni X, Covington KR, Yang BY, Shiu J, Zhang X, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. Nat Genet 2015; 47(12): 1426–34.