UV-induced mutations accumulate during early clonal expansion in aneuploid subtypes of pediatric B-cell precursor acute lymphoblastic leukemia

Lianne C. Suurenbroek,[1*] Cédric G. van der Ham,[1*] Judith M. Boer,[1] Lucca L.M. Derks,[1,2] Rico Hagelaar,[1,2] Markus J. van Roosmalen,[1,2] Puk Klamer,[1] Edwin Sonneveld,[1,3] Monique L. den Boer,[1] Ruben van Boxtel,[1,2] Frank N. van Leeuwen,[1] Jan L.C. Loeffen[1] and Roland P. Kuiper[1,4]

[1]Princess Máxima Center for Pediatric Oncology; [2]Oncode Institute; [3]Dutch Childhood Oncology Group and [4]Department of Genetics, Utrecht University Medical Center, Utrecht University, Utrecht, the Netherlands

**LCS and CGvdH contributed equally as first authors.*

Correspondence: R.P. Kuiper
r.kuiper@prinsesmaximacentrum.nl

Received: October 29, 2025.
Accepted: March 30, 2026.
Early view: April 9, 2026.

https://doi.org/10.3324/haematol.2025.300118

©2026 Ferrata Storti Foundation
Published under a CC BY license

Abstract

Prolonged exposure to ultraviolet (UV) light leads to DNA damage, causing accumulation of mutations and cancer, particularly in the skin. Recent studies revealed that patients with aneuploid subtypes of pediatric B-cell precursor acute lymphoblastic leukemia (BCP-ALL) commonly present with single base substitution signature 7a (SBS7a), a hallmark of UV-based DNA damage, but its origin is unclear. We screened a cohort of 191 whole-genome-sequenced BCP-ALL cases and confirmed the presence of SBS7a in high hyperdiploid, iAMP21 and low hypodiploid subtypes. Screening of an in-house pediatric pan-cancer cohort (N=1,033) detected SBS7a, apart from BCP-ALL, only in tumors with proven cutaneous localization. Subsequent characterization of 43 BCP-ALL samples from 27 SBS7a-positive patients revealed no causative DNA repair defects. Furthermore, the mutational characteristics in SBS7a-positive BCP-ALL were very similar to those encountered in SBS7a-positive adult skin cancers (N=273) and pediatric cutaneous anaplastic large cell lymphomas (N=7), suggesting a shared UV-induced etiology. Subtle differences in aneuploid patterns were observed between SBS7a-positive and SBS7a-negative high hyperdiploid and iAMP21 cases. In several cases, we detected SBS7a-positive subclones at primary diagnosis, or relapses with apparently newly acquired SBS7a-associated mutations, together suggesting prolonged UV damage during expansion. Single-cell whole-genome sequencing in two other cases confirmed some level of SBS7a-induced clonal diversity, but no indications for ongoing SBS7a-associated mutagenesis during bone marrow progression. Our data suggest that SBS7a-associated mutations occur in a subset of aneuploid BCP-ALL subtypes due to UV-induced damage during early expansion in extramedullary sites before bone marrow expansion.

Introduction

Acute lymphoblastic leukemia (ALL) is the most common type of childhood cancer, with an estimated one in every 1,500 newborns developing ALL before their 18th birthday.[1] Roughly 85% are of B-cell precursor (BCP) origin. Recently, multiple studies have investigated the mutational processes which play a role during the development of BCP-ALL. Surprisingly, single base substitution signature 7a (SBS7a), which had previously been identified in skin-localized cancers and associated with damage by ultraviolet (UV) light,[2-5] was detected in specific BCP-ALL subtypes associated with aneuploidies.[6-9]

BCP-ALL has a very low mutational load,[4] but a small subset of patients is affected by mutational processes, evidenced by mutational signatures, such as SBS7a, which can increase the mutational load over 10-fold.[8,9] Currently, the presence of SBS7a has been reported in three BCP-ALL subtypes: 15% of high hyperdiploid cases, 30% of near haploid cases and 50% of cases with intrachromosomal amplification of chromosome 21 (iAMP21).[8-10] These subtypes share the amplification or retention of chromosome 21, while the majority of other subtypes are defined by genomic rearrangements, such as gene fusions.[11] Nevertheless, it is not understood why SBS7a in BCP-ALL only occurs in these specific subtypes.

SBS7a has been strongly associated with UV light as it is prevalent in sun-exposed skin cancers and can be

induced in cell lines by exposure to UV light.[3,12] UVB light can induce the formation of pyrimidine dimers, which can lead to cytosine-to-thymine (C>T) mutations, resulting in SBS7a or, in the case of two adjacent cytosine-to-thymine mutations (CC>TT), double base substitution signature 1 (DBS1).[4,13] This DNA damage from UV light has clear characteristics, such as a bias towards untranscribed regions due to transcription-coupled nucleotide excision repair (NER).[14] In addition to various skin cancers, UV light-induced DNA damage has also been reported in cutaneous T-cell lymphoma, while absent in non-cutaneous T-cell lymphoma.[15]

BCP-ALL is a disease primarily occurring in the bone marrow and blood. Although UV light has been shown to cause SBS7a *in vitro*,[3] UV light cannot penetrate through the dermis and is therefore unlikely to reach the bone marrow.[16] Yet, BCP-ALL cells and healthy memory T cells can present with SBS7a.[9,17] As such, it is still unknown if SBS7a in BCP-ALL truly results from UV light or from a different mutational process that mimics UV light-induced DNA damage. Whether there are underlying DNA repair defects or specific expression patterns underlying the subtype-specific presence of SBS7a in BCP-ALL is also unclear, though patients of European ancestry appear to be more susceptible.[9] Other open questions are where, when and how long BCP-ALL cells get exposed to the mutagenic mechanism that causes SBS7a. In this study, we thoroughly characterized the mutational characteristics in SBS7a-positive BCP-ALL in 27 patients in comparison to 273 cancers with confirmed UV exposure. In addition, we performed subtype-specific comparisons in expression profiles and aneuploidy patterns between SBS7a-positive and SBS7a-negative cases and investigated timing and clonal diversity during SBS7a acquisition.

Methods

Cohort selection

We selected SBS7a-positive pediatric BCP-ALL cases from a retrospective whole-genome-sequenced cohort of multiply relapsed ALL patients,[18] a retrospective whole-exome-sequenced cohort of relapsed ALL patients, a retrospective whole-exome-sequenced cohort of high hyperdiploid ALL, iAMP21 ALL and Down syndrome ALL patients and an in-house prospective whole-genome-sequenced group of newly diagnosed ALL patients (*Online Supplementary Table S1*). Patients selected from whole-exome-sequenced cohorts subsequently underwent whole-genome sequencing (WGS) for validation of SBS7a and inclusion in the study. The WGS was performed at Novogene, Hartwig Medical Foundation or Princess Máxima Center for Pediatric Oncology using Illumina Novaseq 6000 (Illumina, 20012850) to generate 150 base-pair paired-end reads (*Online Supplementary Table S1*). Samples and data were obtained from the biobank of the Dutch Childhood Oncology Group and the Princess Máxima Center. In accordance with the Declaration of Helsinki, informed written consent was obtained from all patients and/or their legal guardians before enrollment in the study and the Princess Máxima Center institutional review board approved the use of excess diagnostic material and data for this study (PMCLAB2019.054, PMCLAB2020.160 and PMCLAB2021.279).

Mutational signature analysis

Mutational matrices were generated using MutationalPatterns v3.4.1.[19] For single base substitutions, *de novo* signature extraction was performed using NMF v0.26[20] with a rank of six (BCP-ALL) or eight (pan-cancer). MutationalPatterns v3.4.1 was used to refit the extracted signatures with COSMIC signatures, and the seven (BCP-ALL) or 11 (pan-cancer) resulting signatures were used to perform a refit of the mutational matrix (*Online Supplementary Tables S2* and *S3*). For samples with a positive contribution for SBS7a, a bootstrap with 100 iterations was performed with MutationalPatterns v3.4.1.[19] Samples were marked as SBS7a-positive if they: (i) showed a positive contribution of SBS7a; (ii) could be reconstructed with a cosine similarity of at least 0.9; and (iii) had a positive SBS7a contribution in at least 90% of bootstraps.

Single-cell whole-genome sequencing

Cells were stained using APC-CD10 (Biolegend, 312209) and DAPI to select living $CD10^+$ lymphoblasts, and sorted for sequencing according to Derks *et al.*[21] Primary template-directed DNA amplification was performed using the ResolveDNA® Whole Genome Amplification Kit – 96 Reactions (BioSkryb 100136).[21] Mesenchymal stromal cells were cultured in bulk, and DNA from these cells was extracted with the QIAamp DNA Micro Kit and used as germline control. Libraries were prepared using a TruSeq Nano kit (Illumina, 20015965), and sequenced with the Novaseq 6000 (Illumina, 20012850) for 2 x 150 base-pair paired-end reads.

Additional descriptions of the methods are given in the *Online Supplementary Data*.

Results

Cases of B-cell precursor acute lymphoblastic leukemia with SBS7a

To study SBS7a in BCP-ALL we collected SBS7a-positive BCP-ALL samples from 27 patients treated at the Princess Máxima Center for Pediatric Oncology (Figure 1A). Seventeen of these were identified by mutational signature analysis on an unselected cohort of 191 BCP-ALL samples diagnosed between 2019 and 2023, revealing a prevalence of 9% and a relative contribution of SBS7a to the mutational spectrum from 10% to 100% (median 36%).

In line with previous studies,[8-10] SBS7a was enriched in three specific subtypes; specifically, 50% of the iAMP21 cases (4 out of 8), 20% of the low hypodiploid cases (1 out of 5) and 16% of the high hyperdiploid cases (12 out of 74) presented with SBS7a (Figure 1B). These subtypes typically present with copy number alterations of chromosome 21. The remaining ten samples were from multiple retrospective BCP-ALL cohorts (*Online Supplementary Table S1*).[18,22] In total, we included 16 high hyperdiploid, nine iAMP21, one low hypodiploid and one B-other case (Figure 1A). SBS7a was usually detected at initial diagnosis, but in two cases the first appearance of SBS7a was in a relapse sample. We did not find significant differences in age at initial diagnosis between SBS7a-negative and

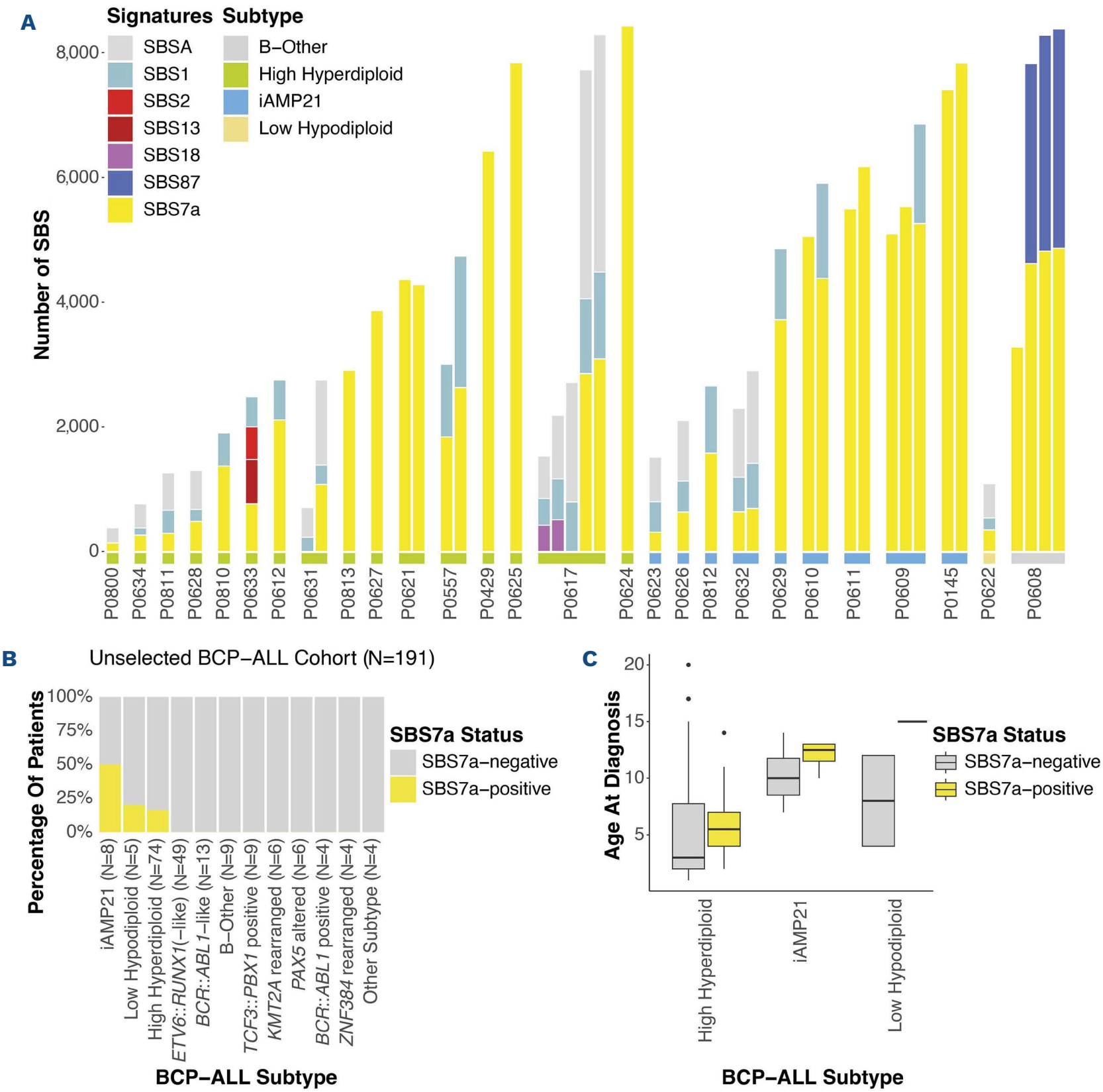

Figure 1. The mutational signature SBS7a can be detected in multiple subtypes of B-cell precursor acute lymphoblastic leukemia. (A) Bar plot depicting the total cohort of patients with B-cell precursor acute lymphoblastic leukemia (BCP-ALL) with SBS7a-associated mutations. A colored bar was added to the x-axis that displays the subtype of each sample. If a patient relapsed, multiple bars are shown to indicate each diagnosis in chronological order. (B) Percentage bar plot presenting the prevalence of SBS7a in BCP-ALL. The x-axis shows the included BCP-ALL subtypes, with *DUX4*-rearranged, *MEF2D*-rearranged, *MYC*-rearranged and near haploid (all N=1) collected in a "Other subtype" group for visual clarity. (C) Boxplot comparing age at diagnosis for patients with and without SBS7a-associated mutations. Boxplots are colored to indicate the presence (yellow) or absence (gray) of SBS7a-associated mutations. Statistical analysis was performed with two-way analysis of variance. SBS: single base substitution.

SBS7a-positive cases for any of the BPC-ALL subtypes (*P*=0.35), indicating that the increased tumor mutational load is not related to aging (Figure 1C).

SBS7a in other pediatric cancers

Next, we investigated the prevalence of SBS7a-associated mutations in pediatric cancers other than BCP-ALL. To this end, we analyzed a pediatric pan-cancer cohort of 1,033 whole-genome-sequenced tumors from the Princess Máxima Center for Pediatric Oncology (Table 1). Except for one melanoma, all SBS7a-positive malignancies were of hematologic origin (Table 1) which, apart from BCP-ALL, included one B-cell lymphoblastic lymphoma (iAMP21 subtype) and seven anaplastic large cell lymphomas (ALCL) (Table 1). The melanoma and B-cell lymphoblastic lymphoma had a cutaneous localization. The ALCL cases often presented with multiple lesions, of which some were cutaneous (*Online Supplementary Table S4*). One BCP-ALL relapse, which presented with SBS7a, was found to be located in the patient's eye. Therefore, whereas in the majority of cases of SBS7a-positive BCP-ALL an association with UV exposure was unclear, other pediatric cancers with SBS7a-associated DNA damage were exclusively found in sun-exposed lesions.

Characteristics of ultraviolet light damage in subtypes of B-cell precursor acute lymphoblastic leukemia

UV exposure introduces pyrimidine dimers that can result in cytosine-to-thymine mutations, which, apart from the single base substitution signature SBS7a, includes an enrichment of CC>TT double base substitutions (DBS1 signature). Furthermore, as observed in sun-exposed tissues, SBS7a and DBS1 present with transcriptional strand asymmetry.[4,14] We therefore performed a comprehensive comparison of mutational characteristics in SBS7a-positive BCP-ALL with those in skin cancer, where UV is well-established as the

Table 1. Prevalence of SBS7a-associated mutations in a pediatric pan-cancer cohort (N=1,033)

Tumor type	N of samples	N of SBS7a-positive samples (%)
Hematologic malignancies	**389**	**24 (6.2)**
Anaplastic large cell lymphoma	9	7 (77.8)
B-cell lymphoblastic lymphoma	3	1 (33.3)
B-cell progenitor acute lymphoblastic leukemia	207	16 (7.7)
Acute myeloid leukemia	60	0
T-cell acute lymphoblastic leukemia	35	0
Burkitt lymphoma	22	0
T-cell lymphoblastic lymphoma	21	0
Diffuse large B-cell lymphoma	11	0
Hodgkin lymphoma	8	0
Chronic myeloid leukemia	7	0
Primary mediastinal large B-cell lymphoma	4	0
Intestinal T-cell lymphoma	1	0
Mature T-cell lymphoma	1	0
Solid tumors	**428**	**1 (0.2)**
Melanoma	4	1 (25)
Neuroblastoma	81	0
Nephroblastoma	71	0
Rhabdomyosarcoma	54	0
Osteosarcoma	48	0
Ewing sarcoma	43	0
Sarcoma other	36	0
Gonadal germ cell	18	0
Hepatoblastoma	16	0
Fibrosarcoma	13	0
Carcinoma other	8	0
Gonadal carcinoma	6	0
Gonadal other	6	0
Histiocytic neoplasm	5	0
Adrenocortical carcinoma	4	0
Nasopharyngeal carcinoma	4	0
Thyroid carcinoma	4	0
Germ cell other	3	0
Hepatic carcinoma	3	0
Renal carcinoma	1	0
Neurological tumors	**216**	**0**
Astrocytoma	63	0
Medulloblastoma	39	0
Glioma high-grade	29	0
Ependymoma	26	0
Glioma	16	0
Atypical teratoid rhabdoid tumor	12	0
Brain high-grade	10	0
Astrocytoma high-grade	9	0
Brain other	9	0
Germ cell brain	3	0

source of SBS7a (*Online Supplementary Methods*).[12] We used a previously characterized cohort of adults with metastatic skin cancer,[23] and included the seven pediatric ALCL cases from our cohort in this analysis. Although SBS7a-positive skin cancers showed a 27.4-fold higher mutational load than SBS7a-positive BCP-ALL (P=2.2x10^{-15}), mutational load in these BCP-ALL cases was still 2.6-fold higher than that in SBS7a-negative BCP-ALL and similar to that in SBS7a-positive ALCL (P=2.1x10^{-8}) (Figure 2A). Similarly, the load of DBS1 (CC>TT mutations) was significantly higher in skin cancer (P=2.7x10^{-15}), but the ratio of CC>TT compared to single C>T mutations was similar in all SBS7a-positive tumors (Figure 2B). Additionally, the well-established bias for the untranscribed strand in skin cancer, caused by transcription-coupled NER,[14] was also present in SBS7a-positive BCP-ALL and ALCL, but not in SBS7a-negative BCP-ALL (P=2.4x10^{-6})

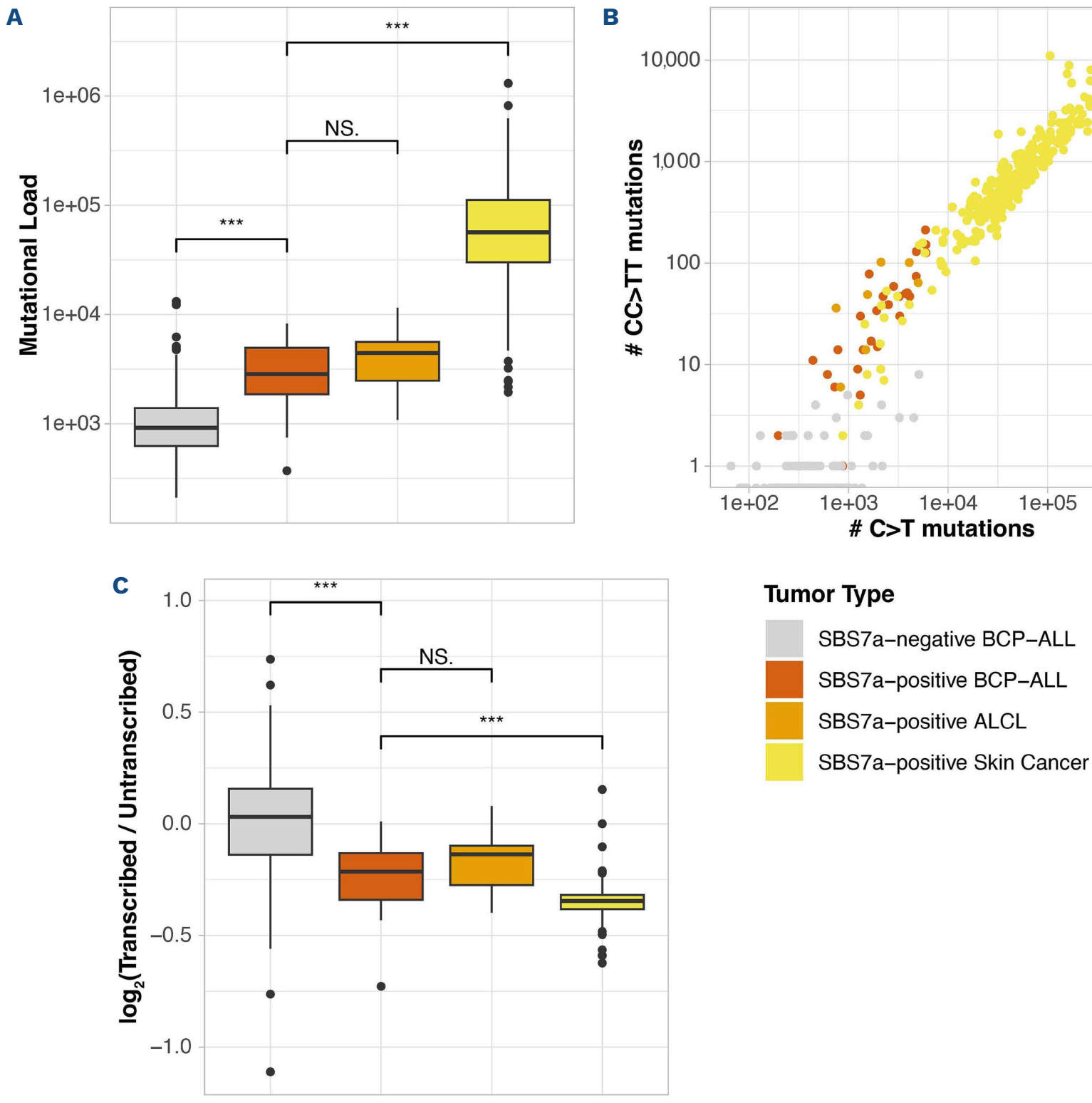

Figure 2. Comparison of SBS7a in pediatric B-cell precursor acute lymphoblastic leukemia, pediatric anaplastic large cell lymphoma and adult skin cancer. (A) Boxplot depicting the median mutational load in each tumor type, with the B-cell precursor acute lymphoblastic leukemia (BCP-ALL) samples split into SBS7a-positive and SBS7a-negative samples. (B) Scatterplot showing the ratio between CC>TT mutations (y-axis) and C>T mutations (x-axis) for each tumor type, with the BCP-ALL samples split in SBS7a-positive and SBS7a-negative samples. (C) Boxplot of the transcriptional strand bias of C>T mutations in each tumor type, with the BCP-ALL samples split in SBS7a-positive and SBS7a-negative samples. On the y-axis, positive values indicate a bias for the transcribed strand and negative values indicate a bias for the untranscribed strand. (A-C) Color corresponds to tumor type. SBS7a-negative BCP-ALL is depicted in grey, SBS7a-positive BCP-ALL is depicted in dark orange, SBS7a-positive anaplastic large cell lymphoma is depicted in orange and SBS7a-positive skin cancer is depicted in yellow. Statistical analysis was performed with the Wilcoxon rank sum test. NS: not significant, *P<0.05, **P<0.01, ***P<0.001. ALCL: anaplastic large cell lymphoma.

(Figure 2C). The stronger transcriptional strand bias in the skin cancer samples (P=2.6x10^{-6}) likely resulted from the far greater proportion of SBS7a-associated mutations relative to mutations associated with other mutational signatures in the C>T spectrum without transcriptional strand bias, like SBS1. Overall, the characteristics of SBS7a in BCP-ALL are indistinguishable from SBS7a in cancers with established UV-induced mutations, with the only notable difference being the number of SBS7a-associated mutations. These findings confirm that mutagenesis in SBS7a-positive BCP-ALL takes place through the formation of pyrimidine dimers, the established direct consequence of UV damage. This is in line with a previous study that reported that SBS7a is more common in BCP-ALL patients of European ancestry than in those of African ancestry.[9]

No recurrent nucleotide excision repair pathway mutations in SBS7a-positive B-cell precursor acute lymphoblastic leukemia

The main DNA repair pathway for UV light-induced DNA damage is NER, and defects in this pathway can exacerbate the acquisition of mutations after UV light exposure.[13,24,25] Therefore, we investigated whether SBS7a-positive BCP-ALL samples harbored NER pathway mutations that could lead to a greater sensitivity to UV-associated DNA damage. To this end, we investigated all genes attributed to the NER pathway in the Kyoto Encyclopedia of Genes and Genomes (*Online Supplementary Table S5, Online Supplementary Methods*) using our genomic and transcriptomic data.[26] We did not detect any pathogenic NER pathway mutations or aberrant expression of NER pathway genes in SBS7a-positive BCP-ALL (*Online Supplementary Figure S1, Online Supplementary Table S6*). This is in line with the presence of transcriptional strand bias, which indicates active transcription-coupled NER. Therefore, NER pathway defects are not responsible for the acquisition of SBS7a in BCP-ALL.

The presence of SBS7a is not associated with transcriptional changes

We then investigated whether the presence of SBS7a could be explained by, or result in, expression of other specific genes or pathways. To do this we performed differential expression analysis of SBS7a-positive and SBS7a-negative samples in the high hyperdiploid and iAMP21 subgroups separately. For validation, we used a cohort of 118 high hyperdiploid and 27 iAMP21 cases from the St. Jude Children's Research Hospital,[27] in which the SBS7a mutation contribution was determined (*Online Supplementary Methods*). Whereas differentially expressed genes were identified in each of the datasets, none of these genes was recurrently found in all datasets (*Online Supplementary Figure S2*). Two upregulated genes overlapped between the iAMP21 cohorts, but these genes were not differentially expressed in the high hyperdiploid samples.

When comparing iAMP21 and high hyperdiploid BCP-ALL to *ETV6::RUNX1* BCP-ALL, we found no overlapping gene sets which could explain the presence of SBS7a (*Online Supplementary Figure S3*). However, when genes commonly upregulated across aneuploid subtypes were selected, an enrichment of several Gene Ontology terms involved in migration was found (*Online Supplementary Figure S4, Online Supplementary Table S7*). Although speculative, this enrichment may underlie a stronger potential of aneuploid BCP-ALL subtypes to migrate to extramedullary sites, providing an explanation for the subtype-specific presentation of SBS7a.

Aneuploidy patterns in SBS7a-positive B-cell precursor acute lymphoblastic leukemia

The restricted presence of SBS7a in aneuploid BCP-ALL subtypes suggests that specific copy number aberrations may be involved in a temporal skin localization or in greater susceptibility to UV light. We therefore investigated these aneuploidy patterns in more detail. We were able to confirm earlier studies demonstrating that SBS7a-associated mutations in high hyperdiploid and iAMP21 BCP-ALL were generally not duplicated on amplified chromosomes, and thus occurred after the subtype-defining ploidy alterations (*Online Supplementary Figure S5, Online Supplementary Methods*).[9,10]

The high hyperdiploid subtype can present with different combinations of amplified chromosomes. Besides the obligatory amplification of chromosome 21, chromosomes 4, 6, 10, 14, 17, 18 and X are often amplified, while amplification of other chromosomes is much less common.[28] For a comparison of the aneuploidy patterns between SBS7a-positive and SBS7a-negative high hyperdiploid BCP-ALL samples, we expanded our dataset with a cohort of high hyperdiploid patients from the St. Jude Children's Research Hospital[27] to a cohort of 219 cases of which 40 were SBS7a-positive (Figure 3A). Clustering analysis did not reveal a globally different pattern between the two groups, but chromosome-specific analyses showed that amplification of chromosomes 4 and 17 was slightly less common in SBS7a-positive patients in both cohorts (Figure 3A, *Online Supplementary Table S8*, *Online Supplementary Methods*). This difference was not due to a general difference in the number of whole chromosome gains (*Online Supplementary Table S9*). Since chromosomal amplification patterns in high hyperdiploid BCP-ALL involving chromosomes 5, 17, 18, and 20 were previously shown to impact prognosis,[29] we tested whether an association could be found between these risk groups and the presence of SBS7a, but such an association was not found (Figure 3B) (P=0.60). We also analyzed other known recurrent alterations in high hyperdiploid BCP-ALL, specifically 1q gain, 6q deletion, isochromosome 7q and isochromosome 17q,[28] and again found no association with SBS7a (*Online Supplementary Table S8*). As the copy number of chromosome 21 in high hyperdiploid BCP-ALL can vary between three and five copies,[30] we analyzed the amount

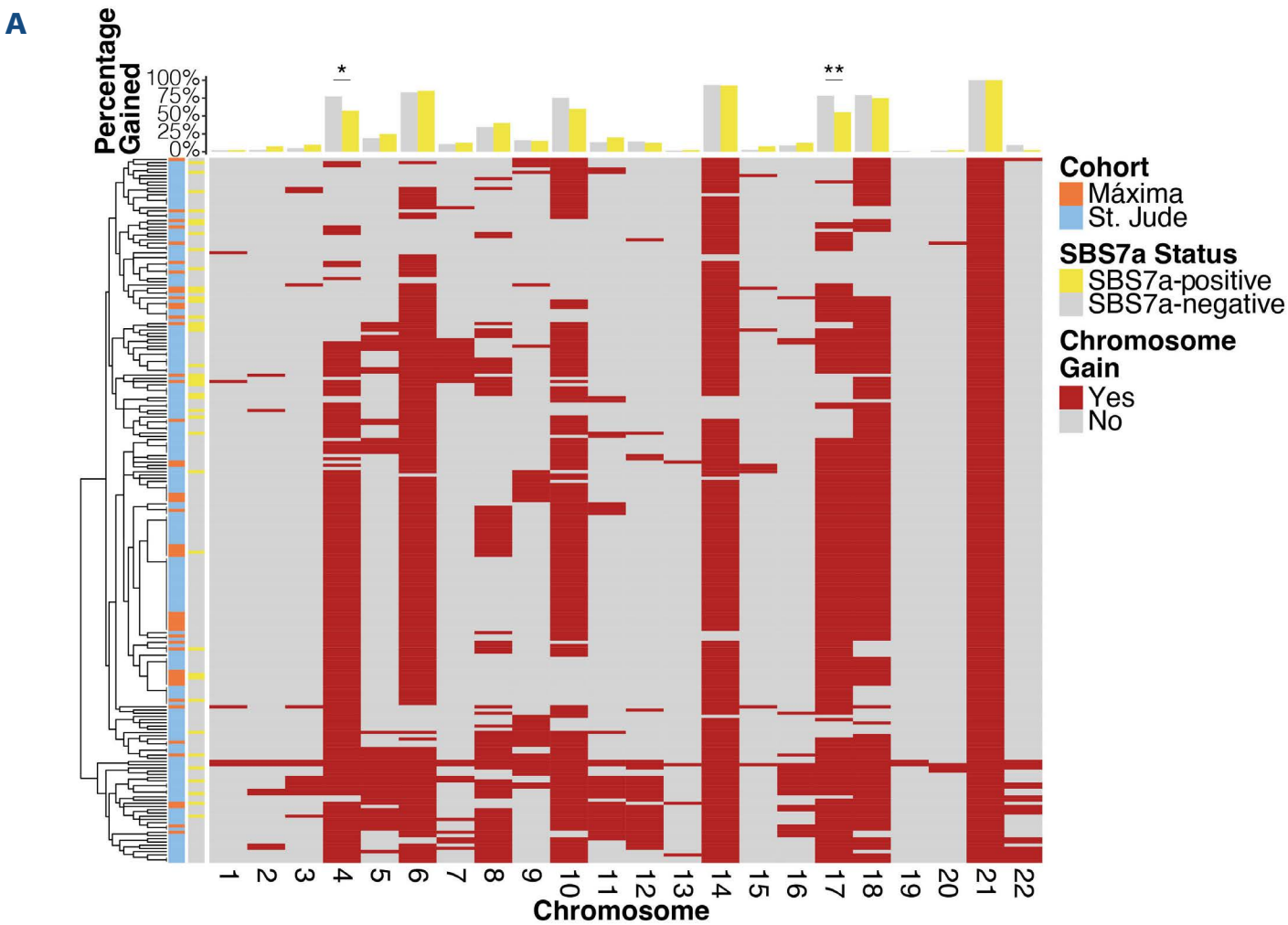

Figure 3. Copy number alterations in SBS7a-positive patients. (A) Plot that depicts the amplified chromosomes in each patient with high hyperdiploid B-cell precursor acute lymphoblastic leukemia (BCP-ALL). Chromosomes are shown on the x-axis and colored red if amplified. Patients are clustered on the y-axis depending on their set of amplified chromosomes. The y-axis also shows if patients are SBS7a-positive (yellow) or SB7a-negative (gray), and if the patient belongs to the Princess Máxima Center cohort (orange) or St. Jude Children's Research Hospital cohort (blue). Additionally, above the plot the percentage of patients with an amplification of each chromosome is shown, for SBS7a-positive (yellow) and SBS7a-negative (gray) patients. (B) Percentage bar plot showing the percentage of SBS7a-positive and SBS7a-negative high hyperdiploid BCP-ALL patients who had a good or poor prognosis according to the criteria established by Enshaei *et al.*[29] (C) Percentage bar plot that indicates the number of chromosome 21 copies detected in SBS7a-positive and SBS7a-negative patients. (D) Graphical representation of chromosome 21 with a line graph of the mean amplified region of chromosome 21 in iAMP21 patients. Patients were separated by cohort and presence of SBS7a. Two regions with higher variability between the four groups are highlighted and accompanied with a boxplot showing the average copy number in that region for SBS7a-negative and SBS7a-positive patients. Statistical analysis was performed with the Fisher exact test (A, B), χ^2 test (C) or Wilcoxon rank sum test (D). NS: not significant, **P*<0.05, ***P*<0.01, ****P*<0.001.

of additional copies of chromosome 21 in SBS7a-positive and SBS7a-negative patients and found no significant difference (Figure 3C) (P=0.31). Finally, we compared the chromosome 21 amplification patterns between SBS7a-negative and SBS7a-positive iAMP21 cases, again using an extended cohort,[27] which revealed two regions that seemed to differ in average copy number between SBS7a-negative and SBS7a-positive patients (Figure 3D). The second region, at the end of the q-arm, was more amplified in SBS7a-negative cases (P=0.01), although this did not result in differential expression of genes (*Online Supplementary Figure S2*). Altogether, SBS7a in high hyperdiploid and iAMP21 BCP-ALL appears to be associated with very subtle differences in commonly amplified regions.

SBS7a-associated mutations in minor subclones

Whereas the likelihood that a (pre)leukemic cell is exposed to UV light appears to be subtype-dependent, the mutagenic consequences may also depend on the duration of UV exposure. Furthermore, expansion of a preleukemic clone during UV exposure may even result in clonal heterogeneity and relapse. To obtain more insight into this process, we investigated whether these mutations accumulate during clonal evolution of BCP-ALL, resulting in subclones with newly acquired SBS7a mutations, or whether they would be found only in the major clone. We selected six patients from our cohort that showed a bimodal split in allele frequency distribution, indicating the presence of a major clone and one or more subclones, each harboring at least 200 mutations (*Online Supplementary Figure S6*). In three patients, SBS7a contribution was high in the major clone, but the minor clone(s) carried no additional SBS7a mutations (Figure 4A, *Online Supplementary Figure S7*). In contrast, the other three cases showed high levels of subclonal SBS7a-associated mutations, suggesting prolonged UV exposure (Figure 4B, *Online Supplementary Figure S7*). These findings suggest that UV-associated mutagenesis is not restricted to a single leukemia-initiating cell but contributes to clonal heterogeneity in at least a subset of cases.

New appearance of ultraviolet light-induced mutations in relapses of B-cell precursor acute lymphoblastic leukemia

In line with the above observation, we noticed that seven out of ten relapsed BCP-ALL cases presented with additional SBS7a-associated mutations at the time of relapse (Figure 1A). To determine whether these mutations could be traced back to initial diagnosis, we selected three cases for in-depth analysis. Each case presented with SBS7a at initial diagnosis and acquired new SBS7a mutations at first relapse (Figure 5, *Online Supplementary Methods*). We then selected relapse-specific mutations with a high likelihood of being caused by SBS7a, and performed amplicon-based

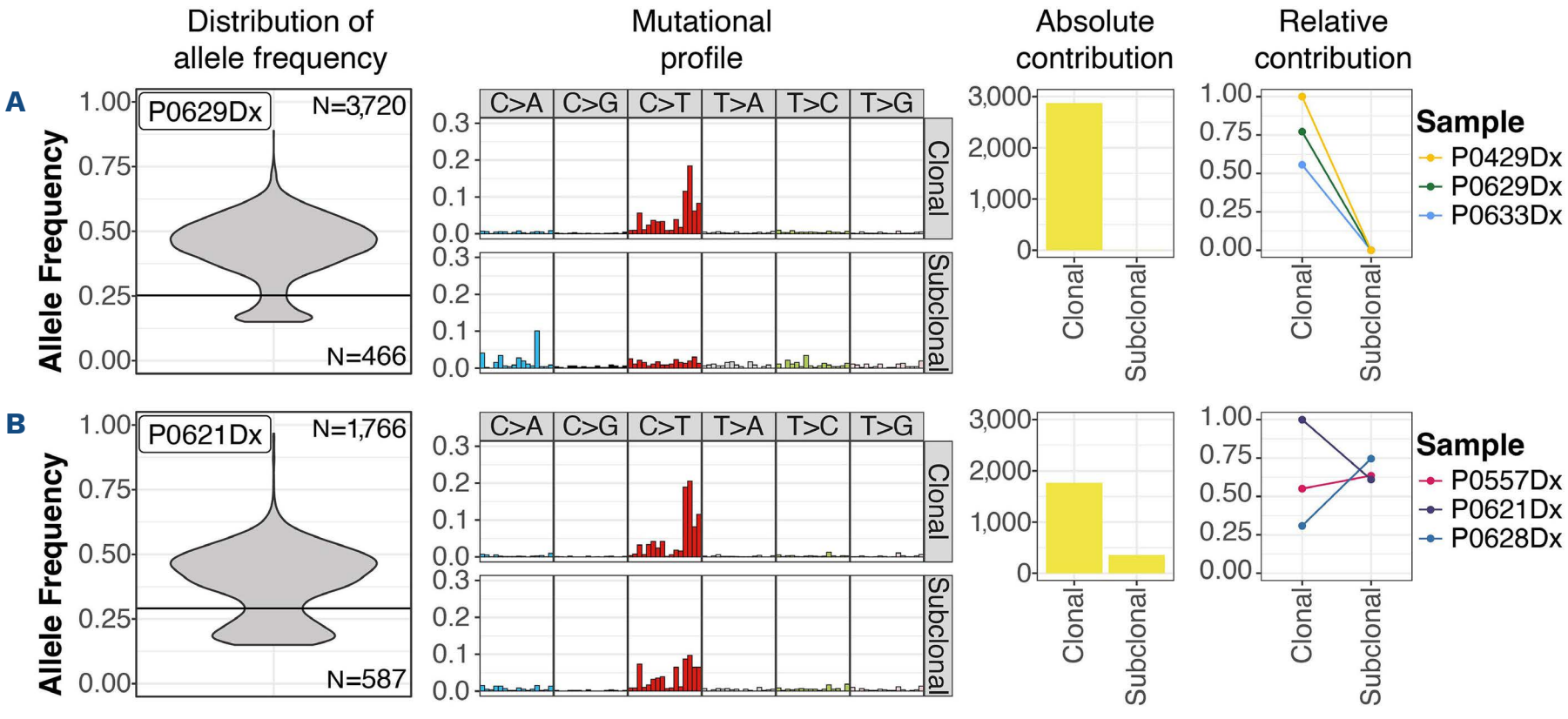

Figure 4. Subclonal SBS7a mutations. (A) In three patients, subclonal mutations that were distinct from the major clone could be identified but were not associated with SBS7a. (B) Three patients had subclonal mutations which were distinct from the major clone and were associated with SBS7a. (A, B) The first column depicts a violin plot of the allele frequency of single base substitutions. Lines were added to indicate the split between mutations belonging to the major clone and subclonal mutations. The second column shows the mutational profile of the clonal and subclonal mutations identified in the first column. The third column shows the absolute contribution of SBS7a-associated mutations to the clonal and subclonal mutations identified in the first column. The fourth column shows the relative number of SBS7a-associated mutations in the clonal and subclonal fraction for all patients belonging to that category. Dx: initial diagnosis.

deep sequencing on the initial diagnosis sample (median depth ~20,000) and first relapse sample (median depth ~6,000) of the three cases. In all three cases, mutations were only detected in the relapse samples. At initial diagnosis, the allele frequency of all mutations was indistinguishable from that of the negative control samples (Figure 5) (P=0.84, P=0.42, P=0.63). These results show that SBS7a mutations found in relapse samples either originated from a subclone at initial diagnosis that was undetectable in the sample due to clone size or location in the bone marrow, or were truly newly acquired.

Relapse-specific mutations can arise because they are positively selected for, for example by conferring drug resistance.[6,7] We found that two mutations in *NR3C1* and three mutations in *CREBBP* have a high probability of being caused by SBS7a (*Online Supplementary Figure S8*). Since both genes are implicated in glucocorticoid resistance, SBS7a might have played a role in disease progression.[7]

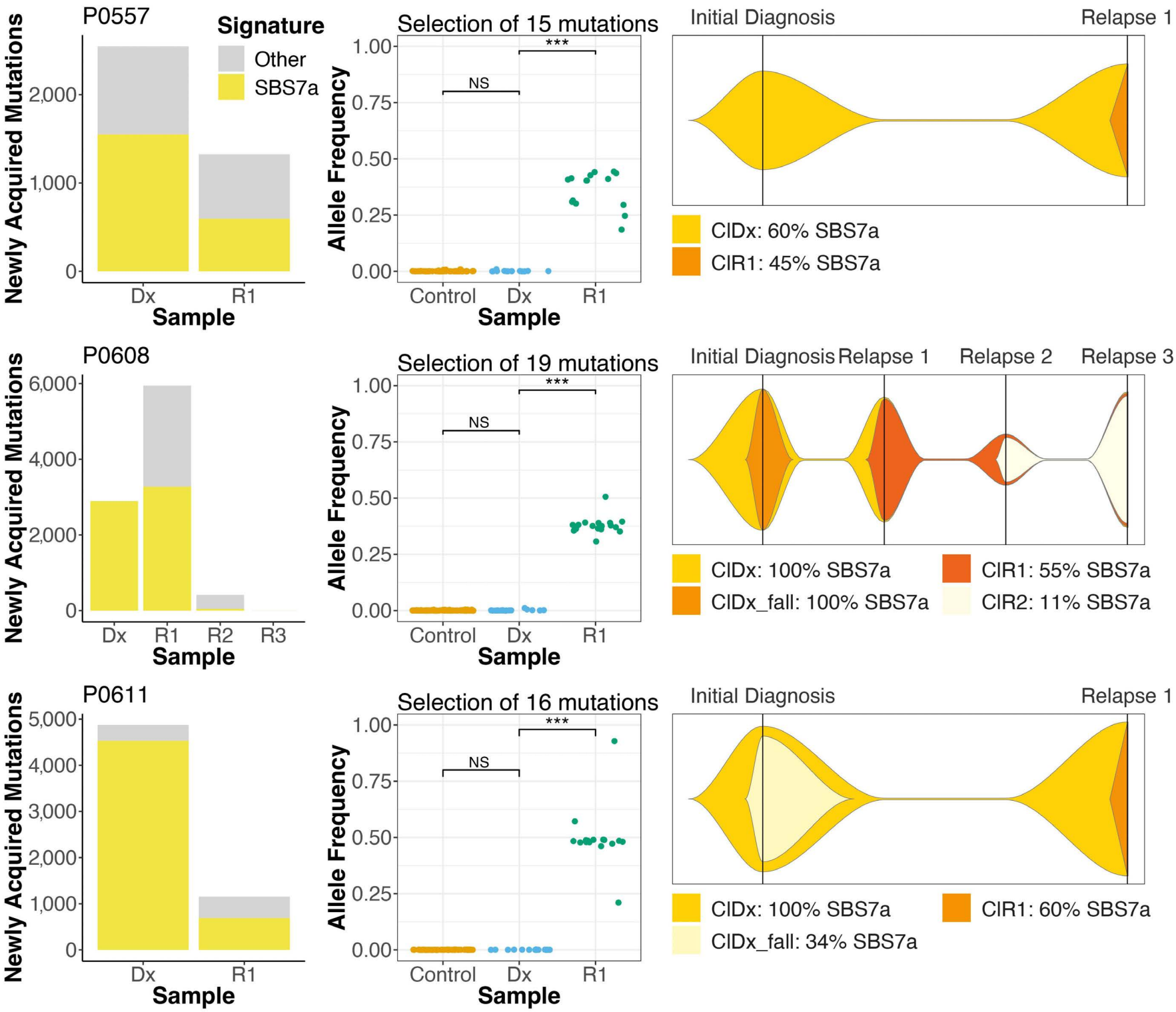

Figure 5. SBS7a-associated mutations found in relapse samples cannot be detected at initial diagnosis. Each row presents data of one patient. The first column shows the number of mutations acquired at each subsequent diagnosis. The yellow coloring represents the number of mutations that were attributed to SBS7a. The second column shows the results of deep sequencing of selected mutations that were very likely introduced by SBS7a and unique to the relapse sample. The x-axis and the color of the dots show if the mutation was sequenced in control samples (orange), the initial diagnosis sample (blue) or the relapse sample (green). Each dot represents a unique mutation. The third column depicts fish plots that represent the evolution of each tumor. Each clone is given a different color and the percentage of SBS7a-associated mutations in each clone is shown in the legend. Statistical analysis was performed with the Wilcoxon rank sum test. NS: not significant, *P<0.05, **P<0.01, ***P<0.001. Dx: initial diagnosis; R: relapse.

SBS7a stops accumulating in the final stage of disease progression

Finally, considering the observation that UV-associated damage proceeds over a prolonged period of time, we addressed the question of whether SBS7a-associated mutations encountered in BCP-ALL represent a past mutational event or a process that is still ongoing during expansion in the bone marrow. We therefore performed single-cell WGS on the initial diagnosis samples of two patients with high hyperdiploid BCP-ALL (P0624 and P0625). For each of these samples, 13 single cells were sorted, and primary template-directed amplification followed by WGS was performed. The single-cell data corresponded well with the bulk WGS data, as for each of these samples, 92% and 91% of the mutations identified in the bulk WGS data were also identified in the single-cell WGS data. We then built a lineage tree to infer clonal evolution and identify subclonal populations. Both lineage trees showed a trunk that contained the mutations we previously detected in the bulk WGS data. For both patients the trunk of the tree, containing mutations identified in all cells, showed a large contribution of SBS7a. In P0625 (Figure 6), two cells individually branch off during early leukemic development. We attribute these branches to lower data quality for these two cells (*Online Supplementary Figure S9*), which suggests that branches X, F, and G actually all belong to the main trunk. However, branch P, which contained mutations shared by four cells, was of good quality. The mutations shared in this branch had a reduced allele frequency in the tumor bulk ($P<2.2x10^{-16}$), which indicates that the four cells originating from branch P are part of an SBS7a-positive subclone (Figure 6). In contrast, none of the unique mutations in each of these cells (mean 219; range, 175-284) showed SBS7a contribution. In P0624 (*Online Supplementary*

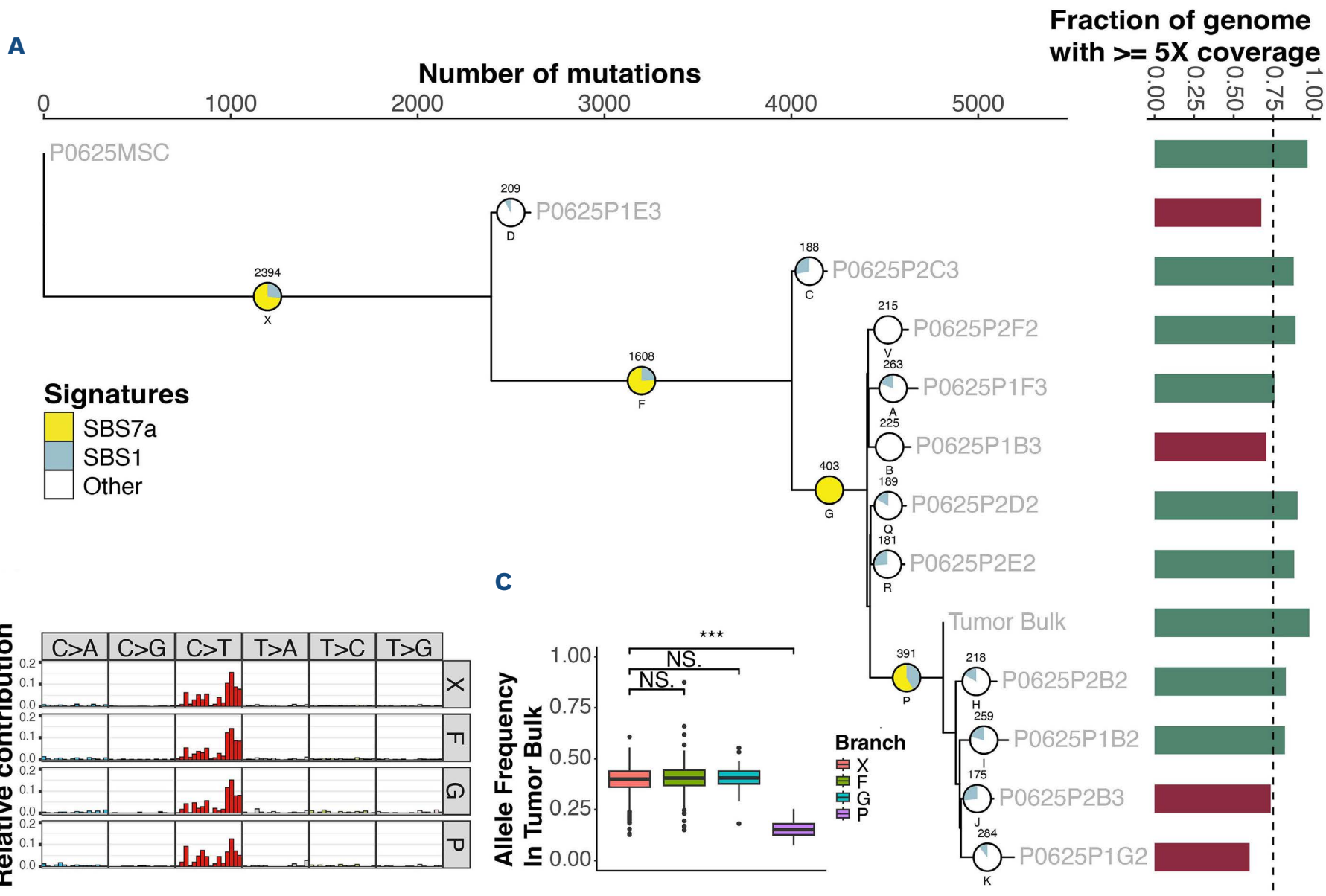

Figure 6. Lineage tree based on the single-cell whole genome sequencing data of patient P0625. (A) Lineage tree. The trunk of the lineage tree shows the number of shared mutations between all following clones. If a clone branches off, it means that it shares no more mutations with other clones. The number of mutations is depicted above each branch along with a pie chart that shows the percentage of mutations associated with each mutational signature. SBS7a is shown in yellow, SBS1 is shown in blue and other mutational signatures are shown in white. To the right of the lineage tree is a bar plot corresponding to each clone in the lineage tree. This bar plot shows whether over 75% of the genome is covered with a depth of at least 5 (green bars). (B) Mutational profiles of each group of shared mutations that had a contribution of SBS7a. (C) Boxplot that shows the allele frequency for each group of shared mutations that contained SBS7a-associated mutations. Statistical analysis was performed with the Wilcoxon rank sum test. NS: not significant, *$P<0.05$, **$P<0.01$, ***$P<0.001$.

Figures S10), we identified two subclonal branches with sufficient quality, branch U and branch W, which do not contain SBS7a (*Online Supplementary Figures S9* and *S10*). Furthermore, similar to P0625, the private branches for all cells in patient P0624 contain only unique mutations (mean 260; range, 190-298) which are not associated with SBS7a (*Online Supplementary Figure S10*). Since these mutations represent the latest stages of leukemic development, these data suggest that SBS7a is not acquired during progression in the bone marrow.

Discussion

The presence of SBS7a, a well-characterized mutational signature linked to UV light-induced DNA damage, was recently reported in aneuploid subtypes of BCP-ALL[31] and then confirmed by several studies.[6,8,9] In this study we identified UV light as the most likely source for SBS7a mutations in BCP-ALL. Additionally, we provide indications that SBS7a-positive BCP-ALL subtypes are exposed to UV light for a prolonged period during early clonal expansion, resulting in subclonal diversity. These results suggest that preleukemic B lymphoblasts with aneuploid karyotypes tend to migrate to and reside in sun-exposed tissues, where they may accumulate UV-induced mutations while undergoing clonal expansion.

A role for UV light in mutation accumulation in BCP-ALL is unexpected in a disease that predominantly develops in the bone marrow. Apart from BCP-ALL, SBS7a has been found in skin cancer, T-cell lymphomas and, as we show in this study, pediatric anaplastic large cell lymphoma,[4,12,15] in which UV exposure as a source of these mutations is plausible. Furthermore, we identified one case in which SBS7a-associated mutations appeared only at relapse in the eye. Consistent with established UV-induced malignancies,[24,32] BCP-ALL presents with pyrimidine dimer formation, followed by mutation incorporation through translesion synthesis.[24,32] So far, the formation of pyrimidine dimers has only been described as a consequence of UV exposure,[24] with no other known inducers of pyrimidine dimers.[3] Although there are chemicals that can induce the formation of DNA adducts, such as psoralen and alkylating agents, these are known to induce different mutational signatures.[3,33] As for endogenous processes, reactive oxygen species can also cause DNA adducts, but result in another mutational signature.[2] These findings are all in line with the previously observed trend that SBS7a occurs more frequently in BCP-ALL patients of European ancestry than in patients of African ancestry, further suggesting that there is an association with sensitivity to UV light.[33-37] Therefore, UV exposure remains the most likely explanation for SBS7a in BCP-ALL.

The most likely scenario by which BCP-ALL can be exposed to UV light is localization in the skin prior to or during (early stages of) leukemia development. Skin lesions in BCP-ALL are rare,[34-38] but may remain small or temporary, and thus underreported. Hematopoietic stem and progenitor cells can circulate in the peripheral blood, home to sites of injury, including the skin, and return to the bone marrow.[39-41] Additionally, a release of large numbers of hematopoietic stem and progenitor cells from the bone marrow can also be triggered by numerous stress conditions.[40] Perhaps, B lymphoblasts can also migrate to extramedullary sun-exposed tissues under certain conditions.

Migration of B lymphoblasts to the skin could be triggered by early driver events. SBS7a in BCP-ALL shows a striking subtype specificity, being mostly restrained to aneuploid BCP-ALL subtypes with amplification or retention of chromosome 21,[9,10] and these amplifications always precede the accumulation of SBS7a. However, not all such cases show SBS7a, suggesting that aneuploidies might infer a risk instead of directly causing SBS7a mutation accumulation. We found minor differences in copy number profiles, which could affect the expression of specific genes that influence the likelihood for temporal localization and/or expansion in sun-exposed areas, resulting in a higher chance of SBS7a-accumulation. Since we found no overall differences in gene expression profiles between SBS7a-positive and -negative BCP-ALL cases, these cases do not appear to be intrinsically different leukemias and thus may in general behave similarly. Additionally, we found some indications that subtype-specific expression might increase the capability of cell migration. This suggests that SBS7a accumulation is a stochastic process, but has a higher likelihood of occurring in the presence of certain primary subtype-specific aberrations.

The increased tumor mutational load in SBS7a-positive BCP-ALL compared to SBS7a-negative cases and the detection of relapse-specific SBS7a mutations indicate prolonged exposure to UV light. Additionally, using both bulk and single-cell WGS, we found that SBS7a mutations can be present in subclones. In later stages of development, however, no new SBS7a mutations seem to be acquired. These results indicate that SBS7a mutations accumulate over time during initial stages of clonal expansion, before further progression in the bone marrow.

Altogether, we propose a model in which (pre)malignant aneuploid BCP-ALL cells temporarily end up in the skin, where they get exposed to UV light while they expand as a (premalignant) lesion. The total amount of UV exposure, and the subsequent mutational consequences, may vary from case to case. Preleukemic cells from such a lesion may return to the blood stream and home to the bone marrow where they potentially progress to overt leukemia. Continued release of cells from such a lesion, which may home to different sites in the bone marrow, could explain the clonal heterogeneity we detected in BCP-ALL. Such a model would also be consistent with our observation that some relapses present with SBS7a mutations that were not detected in the sample taken at initial diagno-

sis, although re-exposure after initial diagnosis cannot be excluded. So far, research into the motility of ALL cells has mainly focused on the differences between B- and T-lineage ALL,[42,43] and it would be interesting to see future studies explore the motility differences between the various BCP-ALL subtypes. Based on our model, we expect that aneuploid subtypes of BCP-ALL show increased motility, leading to the possibility of SBS7a mutation acquisition in sun-exposed tissues.

A key limitation of our study is that it does not resolve why SBS7a arises almost exclusively in aneuploid BCP-ALL subtypes. Whether this reflects increased vulnerability of these cells to UV-induced mutagenesis, or a biological tendency to reside or persist in environments where UV exposure is possible, remains unknown. Addressing this would require experimentally demanding approaches, including subtype-specific UV exposure models, migration or skin-homing assays, and *in vivo* systems that capture transient or pre-diagnostic extramedullary phases. Such models are currently limited or not physiologically representative. Consequently, while our data support UV exposure during early clonal evolution, the upstream cause of subtype specificity remains unresolved.

In conclusion, we show that temporal exposure to UV light is the most likely cause of SBS7a in BCP-ALL. It remains unknown whether SBS7a mutations can alter the development of BCP-ALL. We have found that SBS7a acquisition can introduce pathogenic driver mutations,[18] so SBS7a acquisition might play a role in the onset of BCP-ALL, or even induce therapy resistance in individual cases. However, this remains very speculative, and further studies on large cohorts will need to prove whether SBS7a mutations can influence the prognosis of patients with BCP-ALL. Although the clinical relevance of SBS7a in BCP-ALL remains unknown, the mechanisms of SBS7a mutation accumulation give insight into the origins and development of BCP-ALL.

Disclosures

No conflicts of interest to disclose.

Contributions

LCS conceived the study, curated data, conducted the formal analysis, was responsible for investigations, methodology, software, validation and visualization, wrote the original draft and reviewed and edited the manuscript. CGvdH conceived the study, curated data, conducted the formal analysis, was responsible for investigations, methodology, validation and visualization, wrote the original draft and reviewed and edited the manuscript. JMB, LLMD, ES, MLdB, RvB and JLCL provided resources. RH and MJvR were responsible for resources and software. PK conducted investigations. FNvL acquired funding and wrote, reviewed and edited the manuscript. RPK conceived and supervised the study, acquired funding, was project administrator, wrote the original draft and reviewed and edited the manuscript.

Funding

This study was funded by a grant from the Dutch Cancer Society (grant n. 12482)

Data-sharing statement

The RNA sequencing (EGAD00001015598), whole-genome sequencing (EGAD00001015599, EGAD00001015600) and single-cell whole-genome sequencing data (EGAD00001015601) generated in this study are available on EGA (study ID EGAS00001008128). The code used in this study is publicly available on Zenodo at https://doi.org/10.5281/zenodo.18622757.

References

1. Pieters R, Mullighan CG, Hunger SP. Advancing diagnostics and therapy to reach universal cure in childhood ALL. J Clin Oncol. 2023;41(36):5579-5591.
2. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415-421.
3. Kucab JE, Zou X, Morganella S, et al. A compendium of mutational signatures of environmental agents. Cell. 2019;177(4):821-836.e16.
4. Alexandrov LB, Kim J, Haradhvala NJ, et al. The repertoire of mutational signatures in human cancer. Nature. 2020;578(7793):94-101.
5. Mata DA, Williams EA, Sokol E, et al. Prevalence of UV mutational signatures among cutaneous primary tumors. JAMA Netw Open. 2022;5(3):e223833.
6. Li B, Brady SW, Ma X, et al. Therapy-induced mutations drive the genomic landscape of relapsed acute lymphoblastic leukemia. Blood. 2020;135(1):41-55.
7. Waanders E, Gu Z, Dobson SM, et al. Mutational landscape and patterns of clonal evolution in relapsed pediatric acute lymphoblastic leukemia. Blood Cancer Discov. 2020;1(1):96-111.
8. Studd JB, Cornish AJ, Hoang PH, Law P, Kinnersley B, Houlston R. Cancer drivers and clonal dynamics in acute lymphoblastic leukaemia subtypes. Blood Cancer J. 2021;11(11):177.
9. Brady SW, Roberts KG, Gu Z, et al. The genomic landscape of pediatric acute lymphoblastic leukemia. Nat Genet. 2022;54(9):1376-1389.
10. Gao Q, Ryan SL, Iacobucci I, et al. The genomic landscape of acute lymphoblastic leukemia with intrachromosomal amplification of chromosome 21. Blood. 2023;142(8):711-723.
11. Iacobucci I, Kimura S, Mullighan CG. Biologic and therapeutic implications of genomic alterations in acute lymphoblastic leukemia. J Clin Med Res. 2021;10(17):3792.
12. Hayward NK, Wilmott JS, Waddell N, et al. Whole-genome landscapes of major melanoma subtypes. Nature. 2017;545(7653):175-180.
13. Mullenders LHF. Solar UV damage to cellular DNA: from mechanisms to biological effects. Photochem Photobiol Sci.

2018;17(12):1842-1852.
14. Otlu B, Díaz-Gay M, Vermes I, et al. Topography of mutational signatures in human cancer. Cell Rep. 2023;42(8):112930.
15. Jones CL, Degasperi A, Grandi V, et al. Spectrum of mutational signatures in T-cell lymphoma reveals a key role for UV radiation in cutaneous T-cell lymphoma. Sci Rep. 2021;11(1):3962.
16. D'Orazio J, Jarrett S, Amaro-Ortiz A, Scott T. UV radiation and the skin. Int J Mol Sci. 2013;14(6):12222-12248.
17. Machado HE, Mitchell E, Øbro NF, et al. Diverse mutational landscapes in human lymphocytes. Nature. 2022;608(7924):724-732.
18. van der Ham CG, Suurenbroek LC, Kleisman MM, et al. Mutational mechanisms in multiply relapsed pediatric acute lymphoblastic leukemia. Leukemia. 2024;38(11):2366-2375.
19. Manders F, Brandsma AM, de Kanter J, et al. MutationalPatterns: the one stop shop for the analysis of mutational processes. BMC Genomics. 2022;23(1):134.
20. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics. 2010;11:367.
21. Derks LLM, van Leeuwen AJCN, Steemers AS, et al. Protocol for genome-wide analysis of somatic variants at single-cell resolution using primary template-directed DNA amplification. STAR Protoc. 2025;6(1):103499.
22. Hormann FM, Østergaard A, van den Broek S, et al. Secondary lesions and sensitivity to signaling inhibitors in iAMP21 acute lymphoblastic leukemia. Hemasphere. 2025;9(1):e70069.
23. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. Nature. 2019;575(7781):210-216.
24. Ikehata H, Ono T. The mechanisms of UV mutagenesis. J Radiat Res. 2011;52(2):115-125.
25. Lehmann AR, McGibbon D, Stefanini M. Xeroderma pigmentosum. Orphanet J Rare Dis. 2011;6:70.
26. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30.
27. McLeod C, Gout AM, Zhou X, et al. St. Jude Cloud: a pediatric cancer genomic data-sharing ecosystem. Cancer Discov. 2021;11(5):1082-1099.
28. Haas OA, Borkhardt A. Hyperdiploidy: the longest known, most prevalent, and most enigmatic form of acute lymphoblastic leukemia in children. Leukemia. 2022;36(12):2769-2783.
29. Enshaei A, Vora A, Harrison CJ, Moppett J, Moorman AV. Defining low-risk high hyperdiploidy in patients with paediatric acute lymphoblastic leukaemia: a retrospective analysis of data from the UKALL97/99 and UKALL2003 clinical trials. Lancet Haematol. 2021;8(11):e828-e839.
30. Abbasi MR, Nebral K, Haslinger S, et al. Copy number changes and allele distribution patterns of chromosome 21 in B cell precursor acute lymphoblastic leukemia. Cancers (Basel). 2021;13(18):4597.
31. Ma X, Liu Y, Liu Y, et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. Nature. 2018;555(7696):371-376.
32. Nieto Moreno N, Olthof AM, Svejstrup JQ. Transcription-coupled nucleotide excision repair and the transcriptional response to UV-induced DNA damage. Annu Rev Biochem. 2023;92(1):81-113.
33. Olafsson S, Rodriguez E, Lawson ARJ, et al. Effects of psoriasis and psoralen exposure on the somatic mutation landscape of the skin. Nat Genet. 2023;55(11):1892-1900.
34. Cho-Vega JH, Medeiros LJ, Prieto VG, Vega F. Leukemia cutis. Am J Clin Pathol. 2008;129(1):130-142.
35. Andriescu EC, Coughlin CC, Cheng CE, et al. Pediatric leukemia cutis: a case series. Pediatr Dermatol. 2019;36(5):658-663.
36. Afzal A, Mingola P, Farooque U, Shabih S, Thomas CA. B-cell acute lymphoblastic leukemia presenting as leukemia cutis: a case report. Cureus. 2020;12(10):e11032.
37. Yin Y-T, Tseng J-H, Liu Y-L, Miser JS, Chen S-H. Neonatal acute lymphoblastic leukemia (MLL-AF9) with leukemia cutis. Pediatr Neonatol. 2021;62(6):676-678.
38. Kanoje PK, Khullar G, Yadav AK, Saxena AK, Gupta DK. Generalized leukemia cutis as the initial manifestation of precursor B-cell acute lymphoblastic leukemia in a young male. Dermatol Ther. 2021;34(1):e14594.
39. Khalil S, Ariel Gru A, Saavedra AP. Cutaneous extramedullary haematopoiesis: Implications in human disease and treatment. Exp Dermatol. 2019;28(11):1201-1209.
40. Yang X, Chen D, Long H, Zhu B. The mechanisms of pathological extramedullary hematopoiesis in diseases. Cell Mol Life Sci. 2020;77(14):2723-2738.
41. Kavanagh DPJ, Kalia N. Hematopoietic stem cell homing to injured tissues. Stem Cell Rev. 2011;7(3):672-682.
42. Barz MJ, Behrmann L, Capron D, et al. B- and T-cell acute lymphoblastic leukemias evade chemotherapy at distinct sites in the bone marrow. Haematologica. 2023;108(5):1244-1258.
43. Kanari M, Jimenez Garcia I, Steffen FD, et al. A three-dimensional ex vivo model recapitulates in vivo features and drug resistance phenotypes in childhood acute lymphoblastic leukemia. Leukemia. 2025;39(12):2881-2894.