Received: 20 September 2022 | Accepted: 17 October 2023

DOI: 10.1002/jcla.24982

RESEARCH ARTICLE

WILEY

High-throughput RNA sequencing of the T cell receptor alpha and beta chains for simultaneous clonality and biological analyses in Sezary syndrome

Gonzalo Blanco[1,2,3] | Daniel López-Aventín[4,5] | Ramon M. Pujol[4,6] | Andrea Gómez-Llonín[1,2,7] | Anna Puiggros[1,2] | Manuela López-Sánchez[8] | Teresa Estrach[9] | Mª Pilar García-Muret[10] | Ingrid López-Lerma[11] | Octavio Servitje[12] | Beatriz Bellosillo[13,14] | Manuel Muro[8] | Blanca Espinet[1,2] | Raquel Rabionet[15,16,17,18] | Fernando Gallardo[1,4]

Correspondence
Fernando Gallardo, Servei de Dermatologia, Hospital del Mar, Pg. Marítim 25-29, Barcelona 08003, Spain.
Email: fgallardo@parcdesalutmar.cat

Funding information
Generalitat de Catalunya, Grant/Award Number: 2021/SGR025, 2017/SGR738 and 2017/SGR437; Instituto de Salud Carlos III, Spanish Ministry of Economy and Competitiveness, Grant/Award Number: PI21/00390 and PI18/00021

Abstract

Background: Previous investigations pointed out a role for antigen stimulation in Sezary syndrome (SS). High-throughput sequencing of the T cell receptor (TR) offers several applications beyond diagnostic purposes, including the study of T cell pathogenesis.

Methods: We performed high-throughput RNA sequencing of the TR alpha (TRA) and beta (TRB) genes focusing on the complementarity-determining region 3 (CDR3) in 11 SS and one erythrodermic mycosis fungoides (MF) patients. Five psoriasis patients were employed as controls. Peripheral blood $CD4^+$ cells were isolated and RNA sequenced (HiSeq2500). High-resolution HLA typing was performed in neoplastic patients.

Results: Highly expanded predominant TRA and TRB CDR3 were only found in SS patients (median frequency: 94.4% and 93.7%). No remarkable CDR3 expansions were observed in psoriasis patients (median frequency of predominant TRA and TRB CDR3: 0.87% and 0.69%, $p < 0.001$ compared to SS). CDR3 almost identical to the predominant were identified within each SS patient and were exponentially correlated with frequencies of the predominant CDR3 ($R^2 = 0.918$, $p < 0.001$). Forty-six different CDR3 were shared between SS patients displaying HLA similarities, including predominant TRA and TRB CDR3 in one patient that were found in other three patients. Additionally, 351 antigen matches were detected (Cytomegalovirus, Epstein–Barr, Influenza virus, and self-antigens), and the predominant CDR3 of two different SS patients matched CDR3 with specificity for Influenza and Epstein–Barr viruses.

For Affiliation refer page on 7

Conclusions: Besides detecting clonality, these findings shed light on the nature of SS-related antigens, pointing to RNA sequencing as a useful tool for simultaneous clonality and biological analysis in SS.

KEYWORDS
antigen restriction, complementarity determining region, mycosis fungoides, RNA sequencing, Sezary syndrome, T cell receptor

1 | INTRODUCTION

Sezary syndrome (SS) is an aggressive leukemic variant of cutaneous T cell lymphoma (CTCL) defined by erythroderma, generalized lymphadenopathy, and circulating clonal neoplastic T cells with cerebriform nuclei.[1] Mycosis fungoides (MF), the most common type of CTCL, is a clonal epidermotropic malignant T cell proliferation presenting clinically a more indolent behavior, as cutaneous patches or plaques. A few MF patients evolve into more aggressive stages of the disease, with involvement of lymph nodes and other organs. SS and MF represent approximately 65% of CTCLs. Although they are closely related neoplasms, they are currently considered separate entities on the basis of their different clinical behavior and origin.[1,2]

Clonally rearranged T cell receptor (TR) genes are a molecular hallmark of CTCL. The PCR- based BIOMED-2 protocol has been the most widely used approach for T cell clonality assessment in clinical practice.[3–5] Additionally, high-throughput sequencing of the TR is increasingly becoming an important tool in investigative dermatology, with several applications beyond diagnostic purposes, including the study of the immune response or the pathogenesis of T cell disorders.[6] RNA sequencing has also been proposed as a useful tool for detecting TR rearrangements in T cell lymphoma.[7] One of the best approaches is the analysis of the complementarity-determining region 3 (CDR3), since it displays an extremely high variability resulting in highly antigen-specific CDR3 motifs. TR antigen recognition also occurs in the context of HLA molecules, intimately involved in immune responsiveness.[8]

Previous investigations pointed out a role for antigen stimulation and T cell restrictions in CTCL[9–11] and HLA associations among patients.[12–15] Herein, we report our experience with high-throughput RNA sequencing of the TRA and TRB genes to evaluate its use for simultaneous clonality and biological characterization of SS.

2 | MATERIALS AND METHODS

2.1 | Patients and samples

A total of 11 patients with SS (45% males, median age: 73 years) were evaluated (Table S1). Additionally, one patient with erythrodermic MF (male, 74 years) not fulfilling B2 burden in peripheral blood was studied, together with 5 psoriasis patients employed as negative controls for clonality (TRβ/γ polyclonal by BIOMED-2, Table S1). No case had evidence of infection at sampling. The study was performed in accordance with National and International Guidelines (Professional Code of Conduct, Declaration of Helsinki) and approved by the Ethics Committee of our institution (2010/3824/I). The SS cohort studied here was previously characterized[16]; correspondence between patient identifiers is detailed in Table S1. Fresh peripheral blood samples (25 mL) were obtained from all patients.

2.2 | Cell isolation and RNA extraction

Peripheral blood was subjected to Ficoll centrifugation. Purified $CD4^+$ cells were obtained from mononuclear cells employing positive selection immunomagnetic methods (CD4 MicroBeads) and AutoMACS technology (Miltenyi Biotec, Bergisch Gladbach, Germany). RNA was extracted using the miRNAeasy kit (QIAGEN, Venlo, The Netherlands).

2.3 | High-throughput RNA sequencing

Extracted RNA was sequenced and data were analyzed as previously described[16] (Appendix S1). Raw sequencing data of the SS patients included in this study can be found at https://ega-archive.org/datasets/EGAD00001001998, hosted at the European Genome-Phenome Archive (EGA, https://ega-archive.org/). In this dataset, patient identifiers correspond to those assigned by Prasad et al.;[16] correlations between patient identifiers from both studies are detailed in Table S1. Samtools was used to obtain the sequencing reads aligning to the TRA and TRB gene regions in SAM format, which were then transformed to fastq/fasta with bedtools.

2.4 | Sequence analysis, definitions, and interpretation

The obtained FASTA sequences were entered into IMGT/V-QUEST (http://www.imgt.org). Only CDR3 regions corresponding to productive V(D)J rearrangements were employed. CDR3 amino acid regions were considered expanded within a sample when corresponding to ≥2 sequences (reads). The most expanded CDR3 within each sample

was referred to as the predominant (immunodominant) CDR3, and its frequency was assessed as the number of sequences corresponding to the predominant CDR3 divided by the total number of CDR3 sequences obtained for that sample. CDR3 sequences were considered almost identical to the predominant if they matched in length and contained only one amino acid mismatch. Their frequencies were assessed as the number of sequences corresponding to CDR3 regions almost identical to the predominant divided by the total number of CDR3 sequences or, alternatively, divided by the number of non-predominant CDR3 sequences for that sample. For TRAV and TRBV gene analysis, only sequences with a V-gene score ≥200 were considered. Clonotypes were defined as unique rearrangements carrying identical TRAV-TRAJ or TRBV-TRBJ genes and CDR3 amino acid sequences. Those CDR3 or clonotypes found in different individuals were referred to as "public," whereas those only found in one specific individual as "private."

2.5 | Comparison of CDR3 sequences

Comparison of CDR3 amino acid sequences across all patients was performed. Additionally, comparison to public data was carried out using VDJdb (https://vdjdb.cdr3.net).[17]

2.6 | High-resolution HLA typing

High-resolution typing of the HLA-A/-B/-C/-DRB1/-DQA1/-DQB1 loci was performed for all SS and MF patients. Sequence-specific oligonucleotide-PCR by microbeads array (LABType XR®, 3D Luminex Technology) was employed. HLA data from 282 healthy unrelated individuals from Spain previously assessed by our collaborators[18] was used to compare HLA gene frequencies between SS and healthy subjects from the same population.

2.7 | Statistical analysis

The Mann–Whitney test was employed to compare the frequencies of the predominant CDR3 regions between SS and psoriasis patients. Simple linear regression analysis and Spearman correlation, or exponential regression analysis, were used to analyze the relationship between CDR3 frequencies, whereas the Wilcoxon test was employed to compare the frequencies of shared CDR3 regions between TRA and TRB. The Fisher's exact test was used to compare the frequencies of antigen matches within CDR3 regions shared between patients with those observed for private CDR3, and to compare HLA gene frequencies between SS patients and healthy subjects. Statistical analyses were performed using SPSS v.22 software (SPSS Inc.). *p*-values below 0.05 were considered statistically significant.

3 | RESULTS

3.1 | Highly expanded predominant TRA and TRB CDR3 regions only found in SS

Overall, 78,415 productive CDR3 sequences were obtained from peripheral blood $CD4^+$ T cells of SS patients, 28,586 corresponding to TRA (median: 2,500 sequences/patient, range: 865–4,917) and 49,829 to TRB (median: 2,896 sequences/patient, range: 250–20,932). The total number of distinct CDR3 was 3,670 for TRA (median: 113/patient, range: 28–1,437) and 3,931 for TRB (median: 163/patient, range: 38–1,339). The median number of expanded and single CDR3/patient was 42 (range: 7–354) and 79 (range: 21–1,083) for TRA and 52 (range: 9–310) and 111 (range: 29–1029) for TRB, respectively (Table S2).

Highly expanded predominant TRA and TRB CDR3 regions were observed in SS patients, with a median frequency of the predominant

TABLE 1 TRA and TRB clonotypes corresponding to the most frequent CDR3 region for each SS patient.

	TRA				TRB			
			CDR3				CDR3	
Patient ID	V gene	J gene	Sequence	%	V gene	J gene	Sequence	%
SS01	TRAV6	TRAJ31	ALDNNARLM	94.6	TRBV7-2	TRBJ1-2	ASSPGQPNYGYT	94.9
SS02	TRAV1-2	TRAJ4	AVDGGYNKLI	94.4	TRBV19	TRBJ1-2	ASSPDRGRNYGYT	87.4
SS03	TRAV29/DV5	TRAJ40	AASASSGTYKYI	21.8	TRBV21-1	TRBJ2-7	ASRQGADEQY	26.2
SS04	TRAV9-2	TRAJ36	AHQTGANNLF	55.0	TRBV11-3	TRBJ2-7	ASSGDRGREQY	11.7
SS05	TRAV19	TRAJ27	ALMGNAGKST	95.9	TRBV7-9	TRBJ2-6	ANSFGRSGANVLT	94.2
SS06	TRAV26-1	TRAJ27	IVRVNTNAGKST	96.5	TRBV20-1	TRBJ2-7	SARGLAKIDEQY	93.7
SS07	TRAV13-1	TRAJ29	AARNSGNTPLV	65.5	TRBV7-9	TRBJ1-1	ASSLGQNTEAF	80.0
SS08	TRAV23/DV6	TRAJ50	AASMKTSYDKVI	95.3	TRBV28	TRBJ1-1	ASSLWRRRGTEAF	95.4
SS09	TRAV12-3	TRAJ29	AMSSGNTPLV	96.3	TRBV19	TRBJ2-1	ASSRTGGSYNEQF	96.8
SS10	TRAV17	TRAJ29	ATSRRSGNTPLV	69.8	TRBV29-1	TRBJ2-3	SVGPSGSHTQY	95.0
SS11	TRAV6	TRAJ23	ALAPIYNQGGKLI	87.7	TRBV10-3	TRBJ2-3	AISEPREGPDTQY	90.6

CDR3 of 94.4% for TRA (range: 21.8–96.5) and 93.7% for TRB (range: 11.7–96.8). The detailed frequencies, CDR3 amino acid sequences and V and J genes detected together with the predominant CDR3 regions are detailed in Table 1. Generally, the frequencies of the predominant TRA and TRB CDR3 regions within the same patient were very similar, and a positive correlation between both parameters was observed ($\rho = 0.700$, $p = 0.016$) (Figure 1A).

Concerning the MF patient, 5,469 productive CDR3 sequences were obtained, 2,514 corresponding to TRA and 2,955 to TRB. The number of distinct CDR3 was 1,719 for TRA and 2,082 for TRB, whereas the number of expanded/single CDR3 was 491/1,228 for TRA and 556/1,526 for TRB, respectively. As expected, since this MF patient did not present blood involvement, the frequencies of the predominant TRA and TRB CDR3 regions were very low (1.7% and 2.6%, respectively).

Finally, 12,563 productive CDR3 sequences were obtained for psoriasis patients, 4,433 corresponding to TRA (median: 830 sequences/patient, range: 384–1,579) and 8,130 to TRB (median: 1558 sequences/patient, range: 980–2455). The total number of distinct CDR3 was 3053 for TRA (median: 584/patient, range: 220–1083) and 5540 for TRB (median: 1079/patient, range: 544–1652). No remarkable CDR3 expansions were found in psoriasis patients (median frequency of the predominant TRA and TRB CDR3: 0.87% and 0.69%, $p < 0.001$ compared to SS) (Figure 1B).

3.2 | CDR3 regions almost identical to the predominant in the same SS patient

CDR3 regions almost identical to the predominant (same length and only one amino acid mismatch) were frequently detected within the same SS patient (Table S3). When the total number of sequences (reads) of each patient were considered, the median frequency of CDR3 regions almost identical to the predominant was 1.3 for TRA (range: 0.5–2.7) and 1.9 for TRB (range: 0.4–3.6). When only non-predominant sequences were considered, these values were increased up to a median frequency of 21.4 for TRA (range: 0.6–64.7) and 28.9 for TRB (range: 0.5–81.4) (Table S2).

The frequencies of the predominant CDR3 regions and those of almost identical CDR3 regions in the same SS patient were exponentially correlated. This relationship was less evident when the frequencies of almost identical CDR3 regions were assessed considering the total number of productive sequences for each patient ($R^2 = 0.654$, $p < 0.001$) (Figure 1C), although it became very clear when these frequencies were calculated after removing the sequences of the predominant CDR3 region ($R^2 = 0.918$, $p < 0.001$) (Figure 1D).

3.3 | HLA similarities between SS patients

As detailed in Table 2, HLA similarities were observed between SS patients, mainly regarding class II genes (HLA-DRB1, HLA-DQA1 and HLA-DQB1). Thus, DRB1*01:01-DQA1*01:01-DQB1*05:01 was detected in four SS patients and DRB1*11-DQA1*05:05-DQB1*03:01 in other four SS patients, two presenting the DRB1*11:01 allele and the other two the DRB1*11:04 allele. Besides, A*03:01-B*35:01-C*04:01 was shared between two SS patients. The HLA allele frequencies of SS patients were compared to those of a healthy Spanish cohort previously evaluated by our collaborators[18] (Table S4). Different allele frequencies were identified between both cohorts, although statistical significance was not

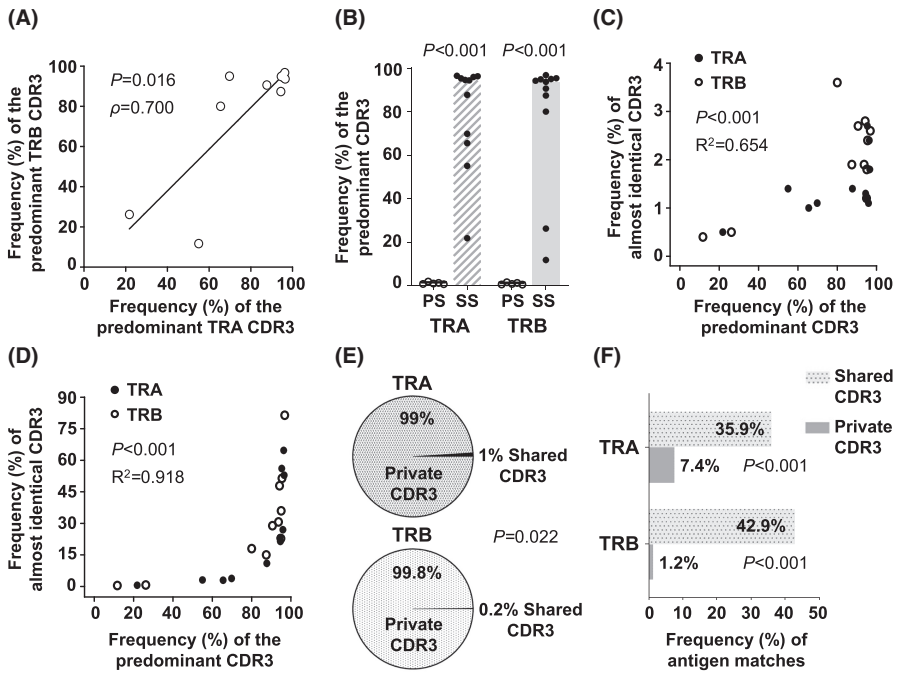

FIGURE 1 TRA and TRB CDR3 analysis. (A) Correlation between the frequencies of the predominant TRA and TRB CDR3 regions of each SS patient. (B) Comparison of the frequencies of the predominant CDR3 regions between SS and psoriasis (PS) patients. Bars represent median frequencies. (C) For SS patients, relationship between the frequencies of the predominant CDR3 regions and the frequencies of CDR3 regions almost identical to the predominant within the total of sequences or (D) within the non-predominant sequences. (E) Frequencies of shared and private CDR3 amino acid sequences across SS patients. (F) Frequencies of antigen matches in shared and private CDR3 regions of SS patients.

TABLE 2 HLA gene usage in the 11 SS and one erythrodermic MF patients.

Patient	HLA-A	HLA-B	HLA-C	HLA-DRB1	HLA-DQA1	HLA-DQB1
SS02	**02:01**, **11:01**	52:01, 55:01	03:03, 12:02	**01:01**, 15:02	**01:01**, **01:03**	**05:01**, 06:01
SS03	**03:01**, **11:01**	**35:01**, **35:01**	**04:01**, **04:01**	**01:01**, 07:01	**01:01**, 02:01	02:02, **05:01**
SS04	**03:01**, **03:01**	18:01, **35:01**	**02:02**, **04:01**	04:03, 14:01	**01:01**, 03:01	03:02, 05:03
SS09	**02:01**, **02:01**	27:05, 51:01	01:01, **02:02**	**01:01**, 13:01	**01:01**, **01:03**	**05:01**, 06:03
SS10	**02:01**, **03:01**	07:02, 40:02	03:04, 15:05	**01:01**, 15:01	**01:01**, 01:02	**05:01**, 06:02
SS06	02:02, 11:01	18:01, 49:01	07:01, 07:01	09:01, **11:04**	03:02, **05:05**	02:02, **03:01**
SS07	23:01, 34:02	14:01, 44:03	04:01, 08:02	11:03, **11:04**	**05:05**, **05:05**	**03:01**, **03:01**
SS08	01:01, **02:01**	44:02, 57:01	05:01, 06:02	07:01, **11:01**	02:01, **05:05**	**03:01**, 03:03
SS11	**02:01**, **02:01**	38:01, 51:01	02:02, 12:03	**11:01**, 13:02	01:06, **05:05**	**03:01**, 06:04
SS01	**29:02**, 31:01	**44:03**, 53:01	04:01, **16:01**	**07:01**, 13:02	**01:02**, **02:01**	**02:02**, 06:04
SS05	02:01, **29:02**	13:02, **44:03**	06:02, **16:01**	**07:01**, 15:01	**01:02**, **02:01**	**02:02**, 05:02
MF01	11:01, **29:02**	18:01, **44:03**	05:01, **16:01**	03:01, **07:01**	**02:01**, 05:01	02:01, **02:02**

Note: Highlighted in the boxes: three groups established according to HLA similarities, which are also highlighted in bold.

achieved due to the reduced number of SS patients. Nonetheless, they showed a trend for a higher frequency of DRB1*01:01 (0.182 vs. 0.064, $p=0.057$), DQA1*01:01 (0.227 vs. 0.108, $p=0.091$) and DQB1*06:04 (0.091 vs. 0.015, $p=0.055$).

Finally, haplotype analysis revealed the presence of the Western European HLA haplotype A*29:02-B*44:03-C*16:01-DRB1*07:01-DQA1*02:01-DQB1*02:02 in three patients, two of them diagnosed with SS and the third with erythrodermic MF.

3.4 | Shared CDR3 regions between distinct patients

Next, we compared all the obtained TRA and TRB CDR3 amino acid sequences across the SS patients included in the study. In all, 46 different CDR3 regions were shared by at least two SS patients, 39 for TRA and 7 for TRB (1.0% and 0.2% of all different CDR3 regions respectively, $p=0.022$) (Figure 1E). Interestingly, AARNSGNTPLV was detected as the predominant TRA CDR3 sequence in patient SS07 (with a frequency of 65.5%) and was also identified in patients SS03 and SS04, in both cases as single TRA CDR3 sequences (frequencies: 0.04% and 0.03%, respectively). Notably, the nucleotide sequences of the aforementioned CDR3 regions were different in the three patients. Similarly, ASSLGQNTEAF was identified as the predominant TRB CDR3 region in patient SS07 (frequency: 80%) and was also observed in patient SS11 as a single TRB CDR3 region (frequency: 0.03%), also displaying different nucleotide sequences. Most of the cases sharing CDR3 regions presented HLA similarities (Table S5).

The CDR3 regions of the erythrodermic MF patient were also compared with the CDR3 regions of all SS patients. A total of 60 matches were observed, 50 for TRA and 10 for TRB (2.9% and 0.5% of all different CDR3 regions of the MF patient, respectively). Except for one case, HLA similarities were detected between the MF patient and those SS patients sharing identical CDR3 regions (Table S5).

3.5 | Antigen matches mainly detected for shared CDR3 regions

VDJdb analysis of the CDR3 amino acid sequences in SS patients revealed 351 matches, 300 for TRA and 51 for TRB. These included a match between the predominant TRA CDR3 region of patient SS09 and the CDR3 region of a T cell clone with specificity for Influenza A virus, and between the predominant TRB CDR3 region of patient SS07 and the CDR3 region of an Epstein–Barr virus-specific T cell clone. This same TRB CDR3 region was also observed as a single CDR3 in patient SS11. Nonetheless, in all these cases, the employed V genes were discordant. When considering only CDR3 regions sharing the same V gene (public clonotypes) and with ≥3 sequence reads, CDR3 of specific T cell clones associated with CMV were highlighted (5 out of 6 CDR3 regions of SS patients fulfilling these criteria, 83%). Additionally, a match with the TRB CDR3 region ASSLGETQY of patient SS02 was identified, which, together with the TRBV5-1 gene, was associated with the recognition of a self-epitope present in the protein encoded by *IGF2BP2* (Table 3). Interestingly, the frequency of antigen matches within the CDR3 regions shared between distinct SS patients was higher than that observed for private (non-shared) CDR3 regions, both for TRA (35.9% vs. 7.4%, $p<0.001$) and TRB (42.9% vs. 1.2%, $p<0.001$) (Figure 1F).

VDJdb analysis of the CDR3 regions identified in the MF patient provided 198 matches, 162 for TRA and 36 for TRB. A match between the predominant TRB CDR3 region of this MF patient (ASSKQGATEAF) and the CDR3 region of a CMV-specific T cell clone was detected, although they exhibited discordant V genes. Considering only public clonotypes with ≥3 sequence reads, two antigen matches were revealed. The first concerned a match with the TRA CDR3 region AGRTGNQFY, which together with the TRAV27 gene was related to a T cell clone with specificity for Influenza A virus. The second match consisted of the TRA CDR3 region AVNGGGADGLT, which together with the TRAV12-2 gene was associated with the recognition of a self-epitope present in the protein coded by *MLANA*, an antigen found on the surface of melanocytes (Table 3).

TABLE 3 Antigen matches for the CDR3 regions of the 11 SS and one erythrodermic MF patients, according to VDJdb.

			Shared clonotype		
Gene	Patient	Antigen match	V gene	J gene	CDR3 sequence
TRA	SS03	CMV	TRAV8-6	TRAJ43	AVSSYNNNDMR
	SS03	CMV	TRAV17	TRAJ23	ATDNQGGKLI
	SS04	CMV	TRAV21	TRAJ37	APSGNTGKLI
	SS04	CMV	TRAV8-3	TRAJ20	AVGGNDYKLS
	MF01	Influenza A	TRAV27	TRAJ49	AGRTGNQFY
	MF01	Self-antigen	TRAV12-2	TRAJ45	AVNGGGADGLT
TRB	SS02	Self-antigen	TRBV5-1	TRBJ2-5	ASSLGETQY
	SS09	CMV	TRBV19	TRBJ2-1	ASRRTGGSYNEQF

Note: Only cases sharing the same V gene (public clonotypes), with V-score ≥200 and ≥3 sequence reads are shown.

4 | DISCUSSION

Prior investigations described T cell restrictions in CTCL.[9,11] These findings are supported by our study, in which we found that the predominant TRA and TRB CDR3 regions of SS patients were highly expanded (median frequency >93%, range: 11.7%–96.8%) and closely correlated, suggesting that the identified predominant TRA and TRB CDR3 regions assemble together in the same TR found in the malignant clone. We employed psoriasis patients as negative controls for clonality and, as expected, no remarkable CDR3 expansions were observed, with no CDR3 representing more than 1.8% of the overall T cell population. These results are similar to those obtained in another RNA sequencing study in peripheral T cell lymphoma, where patients with dominant T cell clones (65/76) exhibited predominant clone frequencies between 11%–99%, whereas in normal T cell populations these frequencies were below 2%.[7] Nonetheless, to optimally assess these values, a much larger sample size must be evaluated. All these findings pave the way for the implementation of RNA sequencing as an option for the analysis of clonal TR rearrangements in T cell malignancies.

We also found CDR3 regions almost identical to the predominant in all SS patients, whose frequencies were exponentially correlated with those of the predominant CDR3 regions. Even though the possibility of sequencing errors cannot be completely discarded (and therefore, some caution is required), different findings point to the existence of almost identical CDR3 regions naturally occurring in SS patients: (i) several almost identical sequencies were detected multiple times (multiple reads); and (ii) several of those detected only once (one read) involved two base changes for one amino acid change. Recent studies have shown that TRs that recognize the same epitopes often exhibit similar sequences.[17,19–21] The identification of these clusters of almost identical CDR3 regions in the same SS patient suggests that increased and specific antigenic pressure may produce groups of T cells equipped with TRs responding to the same antigen. Eventually, the malignant transformation of one of these T cell clones could account for our results. Alternatively, as previously described in other lymphoid neoplasms,[22,23] CDR3 regions almost identical to the predominant could represent intratumor diversification, whose impact on the clinical course of SS patients should be characterized.

Additionally, we identified HLA similarities mainly for HLA class II DR and DQ genes, which could be associated with the classical idea that $CD4^+$ T cells, as those malignant cells found in CTCL, recognize peptides on class II HLA complexes found in antigen-presenting cells.[24] Previous investigations addressed to find HLA associations in patients with CTCL identified similarities mainly for HLA class II genes, whereas class I associations were less consistent.[12,13] In detail, the DRB1*01, DQA1*01:01, DQB1*05 and B*35 alleles were related to poor prognosis in 46 MF Caucasian patients.[14,15] Compared to healthy subjects from the same population, our SS cohort exhibited a trend towards increased DRB1*01 and DQA1*01:01 frequencies, and a specific group of 5 SS patients presented the DQA1*01:01 allele in combination with some of the other poor prognosis alleles disclosed by Brazzelli and colleagues. Other investigations pointed to DRB1*11 and DQB1*03 as susceptibility alleles for CTCL.[12,13] In our study, a second group of four SS patients exhibited both DRB1*11 and DQB1*03 together with the DQA1*05:05 allele, which further supports previous results and draws attention to DRB1*11-DQA1*05:05-DQB1*03 as a susceptibility haplotype for SS. Finally, the Western European HLA haplotype was found in two SS and the one erythrodermic MF subjects. This could potentially point to common immunogenetic susceptibilities in the development of both entities. However, since the studied cohort belongs to the Spanish population, where A*29:02-B*44:03-C*16:01-DRB1*07:01-DQA1*02:01- DQB1*02:02 is one of the most frequent haplotypes,[18,25] its potential association with SS/MF should be better characterized. For both cases (specific alleles and haplotypes), the evaluation of larger Spanish SS cohorts should be performed to corroborate the HLA associations suggested by this study.

Besides, we identified CDR3 regions shared between distinct patients, most of them displaying different nucleotide sequences, which excludes the possibility of cross- contamination. These shared CDR3 regions could be the result of several independent recombination events converging in the same CDR3 (convergent recombination) or due to the presence of the same antigens acting in different individuals and expanding the same specific T cell clone (convergent antigen selection).[26] Prior T cell repertoire

studies performed with limited sampling depth (e.g., using cloning strategies) hardly allowed the identification of shared T cell clones among patients, which led to considering them as disease-specific (convergent antigen selection).[27,28] Nonetheless, other investigations carried out with high-throughput methods in healthy unrelated individuals demonstrated a huge overlap in TR repertoires.[29] We also detected a significant increased frequency of shared CDR3 regions for TRA compared to TRB, which is in accordance with a recent high-throughput study performed in 20 healthy subjects which demonstrated that TRA repertoires are more similar between individuals than TRB repertoires.[30] These previous high-throughput investigations in healthy donors suggest that most of the shared CDR3 regions found among our patients could have been generated by convergent recombination rather than convergent antigenic selection (the latter being potentially associated with the pathogenesis of the disease). However, the predominant TRA and TRB CDR3 regions of patient SS07 (which could be considered the CDR3 regions of the malignant T cell clone) were also found in three other patients displaying different nucleotide sequences, pointing to common antigens potentially related to SS pathogenesis.

Among the different antigen matches provided by VDJdb, the most restrictive analysis (same CDR3 and V gene) highlighted specific T cell clones associated with CMV. Previous studies reported a significant association between CMV infection or reactivation and SS or MF development or progression.[31,32] A match with a T cell clonotype able to recognize a self-epitope found in melanocytes was also identified. The previous finding, which was observed in the erythrodermic MF patient from our series, is compatible with the particular skin tropism that characterizes the malignant T cells of this entity, which express markers of skin resident T cells[2] and cutaneous T cell homing receptors (CLA antigen). When less restrictive criteria were applied (same CDR3 but different V gene), the predominant CDR3 regions of two distinct SS patients matched CDR3 regions found in the context of Epstein–Barr virus and Influenza A infections. As these antigen matches concerned the predominant CDR3 regions (probably corresponding to the malignant clone), an association between these specific viruses and malignant T cell proliferation in these two patients could be hypothesized. Nonetheless, since the V genes that accompanied these common CDR3 regions were discordant, these results should be interpreted with caution. Additionally, a match between a CDR3 region of a CMV-specific T cell clone was identified for the predominant TRB CDR3 region of the erythrodermic MF patient, also exhibiting discordant V genes. Some studies pointed to Epstein–Barr virus as a likely etiologic agent or promoter of SS and MF pathogenesis,[33,34] and CMV infection was associated with CTCL, as previously discussed.[31,32] However, functional studies should be performed to confirm the stimulation of the T cell clones of these patients by the aforementioned pathogens.

In all, we have demonstrated that $CD4^+$ T cells of SS patients express highly expanded and correlated predominant TRA and TRB CDR3 regions, CDR3 almost identical to the predominant in the same patient, and shared CDR3 regions enriched in antigen matches between different patients (including CMV, Epstein–Barr virus, Influenza virus and self-antigens). No remarkable CDR3 expansions were found in psoriasis patients. It is worth noting that, contrarily to exome data, RNA sequencing provides information regarding all transcribed genes found in the sample, as previously published for the SS cohort of the present study.[16] All these findings, together with the rapid decrease of sequencing costs and the increasing importance of tumor profiling in the context of personalized medicine, point to high-throughput RNA sequencing as a useful tool for parallel clonality and biological assessment in SS.

AUTHOR CONTRIBUTIONS

Conceptualization: GB, DLA, RMP, BE, RR, FG; Data Curation: GB, DLA, RR; Formal Analysis: GB, RR; Funding Acquisition: BE, RMP, FG; Methodology: AGL, AP, MLS; Resources: RMP, TE, MPGM, ILL, OS, BB, MM, RR, FG; Supervision: BE, RMP, FG; Writing—Original Draft Preparation: GB, DLA; Writing—Review and Editing: All authors.

AFFILIATIONS

[1]Grup de Recerca Translacional en Neoplàsies Hematològiques, Cancer Research Programme, IMIM-Hospital del Mar, Barcelona, Spain
[2]Laboratori de Citogenètica Molecular, Servei de Patologia, Hospital del Mar, Barcelona, Spain
[3]Karches Center for Oncology Research, The Feinstein Institutes for Medical Research, Northwell Health, Manhasset, New York, USA
[4]Servei de Dermatologia, Hospital del Mar, Barcelona, Spain
[5]Programa de Doctorat en Medicina, Departament de Medicina, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain
[6]Grup de Recerca en Malalties Inflamatòries i Neoplàsiques Dermatològiques, Inflammation and Cardiovascular Disorders Research Programme, IMIM-Hospital del Mar, Barcelona, Spain
[7]Laboratori de Recerca Translacional en Microambient Tumoral, Cancer Research Programme, Barcelona, Spain
[8]Servicio de Inmunología, Hospital Clínico Universitario Virgen de la Arrixaca-IMIB, Murcia, Spain
[9]Servei de Dermatologia, Hospital Clínic, IDIBAPS, Universitat de Barcelona, Barcelona, Spain
[10]Servei de Dermatologia, Hospital de Sant Pau, Barcelona, Spain
[11]Servei de Dermatologia, Hospital Vall d'Hebron, Barcelona, Spain
[12]Servei de Dermatologia, Hospital de Bellvitge, Barcelona, Spain
[13]Grup de Recerca Clínica Aplicada en Neoplàsies Hematològiques, Cancer Research Programme, IMIM-Hospital del Mar, Barcelona, Spain
[14]Laboratori de Biologia Molecular, Servei de Patologia, Hospital del Mar, Barcelona, Spain
[15]Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Barcelona, Spain
[16]Institut de Recerca Sant Joan de Déu, Hospital Sant Joan de Déu, Barcelona, Spain
[17]Institut de Biomedicina (IBUB), Universitat de Barcelona, Barcelona, Spain
[18]Centro de Investigaciones Biomédicas en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III (ISCIII), Madrid, Spain

ACKNOWLEDGMENTS

The authors want to thank Xavier Duran for statistical support. G. Blanco was funded by a grant from Fundación Española de Hematología y Hemoterapia (FEHH-Janssen) and Fundación Alfonso Martín Escudero. This work has been supported by the following grants: PI18/00021 and PI21/00390, Instituto de Salud Carlos III, Spanish Ministry of Economy and Competitiveness and 2017/SGR437,

2017/SGR738 and 2021/SGR025 from Generalitat de Catalunya and co-funded by the European Union.

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

ORCID

Gonzalo Blanco https://orcid.org/0000-0001-6557-3764
Daniel López-Aventín https://orcid.org/0000-0001-5097-5621
Ramon M. Pujol https://orcid.org/0000-0002-5622-6055
Andrea Gómez-Llonín https://orcid.org/0000-0002-7470-6730
Anna Puiggros https://orcid.org/0000-0001-9627-4978
Manuela López-Sánchez https://orcid.org/0000-0002-5094-0266
Teresa Estrach https://orcid.org/0000-0002-4601-2811
Mª Pilar García-Muret https://orcid.org/0000-0003-1714-5815
Ingrid López-Lerma https://orcid.org/0000-0002-7690-8926
Octavio Servitje https://orcid.org/0000-0002-9910-9617
Beatriz Bellosillo https://orcid.org/0000-0002-5335-2726
Manuel Muro https://orcid.org/0000-0001-9987-0994
Blanca Espinet https://orcid.org/0000-0002-4294-8145
Raquel Rabionet https://orcid.org/0000-0001-5006-8140
Fernando Gallardo https://orcid.org/0000-0002-8996-3828

REFERENCES

1. Whittaker SJ, Cerroni L, Willemze R, Siebert R. Sezary syndrome. In: Swerdlow SH, Campo E, Harris NL, et al., eds. *WHO classification of tumours of haematopoietic and lymphoid tissues*. 4th ed. International Agency for Research on Cancer; 2017:390-391.
2. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood*. 2010;116:767-771.
3. van Dongen JJ, Langerak AW, Brüggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 concerted action BMH4-CT98-3936. *Leukemia*. 2003;17:2257-2317.
4. Gallardo F, Bellosillo B, Serrano S, Pujol RM. Genotypic analysis in primary cutaneous lymphomas using the standardized BIOMED-2 polymerase chain reaction protocols. *Actas Dermosifiliogr*. 2008;99:608-620.
5. López-Aventín D, Gallardo F, Colomo L, et al. Diagnostic value of genotypic analysis in primary cutaneous lymphomas using standardized BIOMED-2 polymerase chain reaction protocols: experience in daily clinical practice. *Acta Derm Venereol*. 2021;101:adv00460.
6. Matos TR, de Rie MA, Teunissen MBM. Research techniques made simple: high- throughput sequencing of the T cell receptor. *J Invest Dermatol*. 2017;137:e131-e138.
7. Gong Q, Wang C, Zhang W, et al. Assessment of T cell receptor repertoire and clonal expansion in peripheral T cell lymphoma using RNA-seq data. *Sci Rep*. 2017;7:11301.
8. Janeway CA, Travers P, Walport M, Shlomchik M. *Immunobiology: the immune system in health and disease*. 5th ed. Garland Science; 2001.
9. Yawalkar N, Ferenczi K, Jones DA, et al. Profound loss of T cell receptor repertoire complexity in cutaneous T cell lymphoma. *Blood*. 2003;102:4059-4066.
10. Linnemann T, Gellrich S, Lukowsky A, et al. Polyclonal expansion of T cells with the TCR V beta type of the tumour cell in lesions of cutaneous T cell lymphoma: evidence for possible superantigen involvement. *Br J Dermatol*. 2004;150:1013-1017.
11. van der Fits L, Sandberg Y, Darzentas N, et al. A restricted clonal T cell receptor αβ repertoire in Sezary syndrome is indicative of superantigenic stimulation. *Br J Dermatol*. 2011;165:78-84.
12. Safai B, Myskowski PL, Dupont B, Pollack MS. Association of HLA-DR5 with mycosis fungoides. *J Invest Dermatol*. 1983;80:395-397.
13. Jackow CM, McHam JB, Friss A, Alvear J, Reveille JR, Duvic M. HLA-DR5 and DQB1*03 class II alleles are associated with cutaneous T cell lymphoma. *J Invest Dermatol*. 1996;107:373-376.
14. Brazzelli V, Rivetti N, Badulli C, et al. Immunogenetic factors in mycosis fungoides: can the HLA system influence the susceptibility and prognosis of the disease? Long-term follow-up study of 46 patients. *J Eur Acad Dermatol Venereol*. 2014;28:1732-1737.
15. Brazzelli V, Rivetti N, Badulli C, et al. Mycosis fungoides: association of KIR ligands and HLA-DQB1*05 with bad prognosis of the disease. *J Eur Acad Dermatol Venereol*. 2016;30:266-269.
16. Prasad A, Rabionet R, Espinet B, et al. Identification of gene mutations and fusion genes in patients with Sezary syndrome. *J Invest Dermatol*. 2016;136:1490-1499.
17. Shugay M, Bagaev DV, Zvyagin IV, et al. VDJdb: a curated database of T cell receptor sequences with known antigen specificity. *Nucleic Acids Res*. 2018;46:D419-D427.
18. Montero-Martín G, Mallempati KC, Gangavarapu S, et al. High-resolution characterization of allelic and haplotypic HLA frequency distribution in a Spanish population using high-throughput next-generation sequencing. *Hum Immunol*. 2019;80:429-436.
19. Qi Q, Cavanagh MM, le Saux S, et al. Diversification of the antigen-specific T cell receptor repertoire after varicella zoster vaccination. *Sci Transl Med*. 2016;8:332ra46.
20. Dash P, Fiore-Gartland AJ, Hertz T, et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature*. 2017;547:89-93.
21. Glanville J, Huang H, Nau A, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature*. 2017;547:94-98.
22. Kostareli E, Sutton LA, Hadzidimitriou A, et al. Intraclonal diversification of immunoglobulin light chains in a subset of chronic lymphocytic leukemia alludes to antigen-driven clonal evolution. *Leukemia*. 2010;24:1317-1324.
23. Bagnara D, Tang C, Brown JR, et al. Post- transformation IGHV-IGHD-IGHJ mutations in chronic lymphocytic leukemia B cells: implications for mutational mechanisms and impact on clinical course. *Front Oncol*. 2021;11:640731.
24. Brodsky FM, Lem L, Bresnahan PA. Antigen processing and presentation. *Tissue Antigens*. 1996;47:464-471.
25. Muro M, Marín L, Torío A, et al. HLA polymorphism in the Murcia population (Spain): in the cradle of the archaeologic Iberians. *Hum Immunol*. 2001;62:910-921.
26. Pogorelyy MV, Minervina AA, Chudakov DM, et al. Method for identification of condition-associated public antigen receptor sequences. *Elife*. 2018;7:e33050.
27. Vardi A, Agathangelidis A, Stalika E, et al. Antigen selection shapes the T-cell repertoire in chronic lymphocytic leukemia. *Clin Cancer Res*. 2016;22:167-174.
28. Blanco G, Vardi A, Puiggros A, et al. Restricted T cell receptor repertoire in CLL-like monoclonal B cell lymphocytosis and early stage CLL. *Oncoimmunology*. 2018;7:e1432328.
29. Shugay M, Bolotin DA, Putintseva EV, Pogorelyy MV, Mamedov IZ, Chudakov DM. Huge overlap of individual TCR beta repertoires. *Front Immunol*. 2013;4:466.
30. Kitaura K, Shini T, Matsutani T, Suzuki R. A new high-throughput sequencing method for determining diversity and similarity of T cell receptor (TCR) α and β repertoires and identifying potential new invariant TCR α chains. *BMC Immunol*. 2016;17:38.

31. Herne KL, Talpur R, Breuer-McHam J, Champlin R, Duvic M. Cytomegalovirus seropositivity is significantly associated with mycosis fungoides and Sezary syndrome. *Blood*. 2003;101:2132-2136.
32. Ballanger F, Bressollette C, Volteau C, Planche L, Dreno B. Cytomegalovirus: its potential role in the development of cutaneous T cell lymphoma. *Exp Dermatol*. 2009;18:574-576.
33. Dreno B, Celerier P, Fleischmann M, Bureau B, Litoux P. Presence of Epstein-Barr virus in cutaneous lesions of mycosis fungoides and Sezary syndrome. *Acta Derm Venereol*. 1994;74:355-357.
34. Nahidi Y, Meibodi NT, Ghazvini K, Esmaily H, Hesamifard M. Evaluation of the association between Epstein-Barr virus and mycosis fungoides. *Indian J Dermatol*. 2015;60:321.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Blanco G, López-Aventín D, Pujol RM, et al. High-throughput RNA sequencing of the T cell receptor alpha and beta chains for simultaneous clonality and biological analyses in Sezary syndrome. *J Clin Lab Anal*. 2023;37:e24982. doi:10.1002/jcla.24982