ID No.	K2102
Project Title	Multiple sclerosis: analysis of the T and B cell receptor repertoire in Mexican patients
Principal Investigator	Martha Espinosa-Cantellano (Prof., Center for Research and Advanced Studies)
Project Members IMSUT Host Researcher	Seiya Imoto (Prof., IMSUT)
Members	Hugo David Gonzalez Conchillos (PhD Student, Center for Research and Advanced Studies) Miriam Patricia Perez Saldivar (PhD Student, Center for Research and Advanced Studies) Satoru Miyano (Project Prof., Tokyo Medical & Dental Univ.) Rui Yamaguchi (Team Leader, Aichi Cancer Center) Kazuma Kiyotani (Senior Researcher, Japanese Foundation for Cancer Research) Yusuke Nakamura (Director, Japanese Foundation for Cancer Research)

Report

The aim of this project is to analyze the repertoire of T and B cell receptors (TCR and BCR) by next generation sequencing (NGS) in patients with multiple sclerosis (MS) during relapse and remission phases.

Preparation of libraries and NGS were performed in 2018 at the Japanese Foundation for Cancer Research (JFCR) from October 1 to October 31. Blood samples from 15 Mexican MS patients were collected during relapse, 10 of which were also obtained during remission. In addition, 6 patients with other inflammatory neurological diseases (OIND) were included as controls. For the analysis of BCR repertoire, RNA from peripheral blood mononuclear cells (PBMC) of each patient was used for the synthesis of cDNA and the subsequent generation of heavy- and light-chain (kappa and lambda) libraries. For TCR analysis, RNA was obtained from the PBMCs of each patient and, in addition, PBMC samples of 4 of the patients during the relapse and remission phases were isolated into subpopulations of T lymphocytes ($CD4^+$, $CD8^+$ and $CD25^+$) using magnetic beads coupled to specific antibodies. For PBMC samples, the TCR beta chain was amplified, while for the lymphocyte subpopulations, both chains (Alpha and Beta) were amplified. An Illumina MiSeq system with a MiSeq reagent v2 kit (300 cycles) was used for NGS.

To learn how to analyze BCR and TCR sequencing data, the two Mexican PhD students (Hugo David Gonzalez Conchillos and Miriam Patricia Perez Saldivar) visited IMSUT from January 25 to February 17, 2020. This stay allowed them to learn how to apply Tcrip and Bcrip pipelines developed by the Japanese colleagues at IMSUT to the Mexican data. The Mexican PhD students applied these algorithms to assign each sequencing reading to the V, J, and C

reference sequences of TCR and BCR obtained from IMGT (www.imgt.org). Thus, the BCR and TCR clonotypes of each patient were obtained.

In addition, the students learned to calculate the diversity index (DI) of unique combinations of V (D) J with CDR3 sequences from the inverse of Simpson diversity index, within Tcrip and Bcrip pipelines.

Preliminary results of BCR repertoire analysis

- Diversity Index

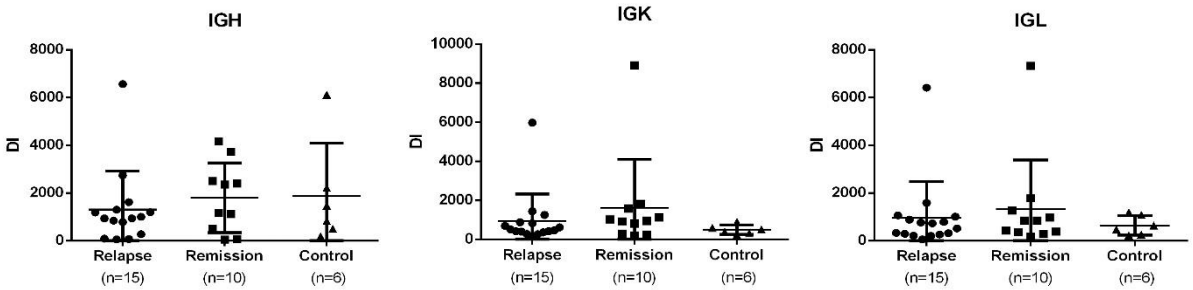


Fig. 1. Diversity index (DI) of IGH, IGK and IGL. Although less diversity seemed to be present in the IGH repertoire of patients during relapse phase compared to remission and to controls, no significant difference between the three groups was found after performing Mann-Whitney test.

- Shared clonotypes

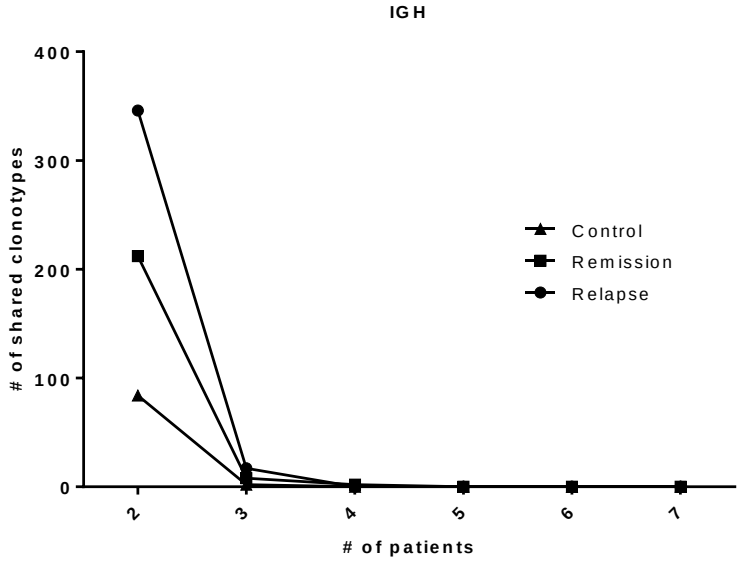


Fig. 2. Shared IGH clonotypes of relapse, remission and control patients. More shared clonotypes were found among relapse patients.

- IGHVJ segment use

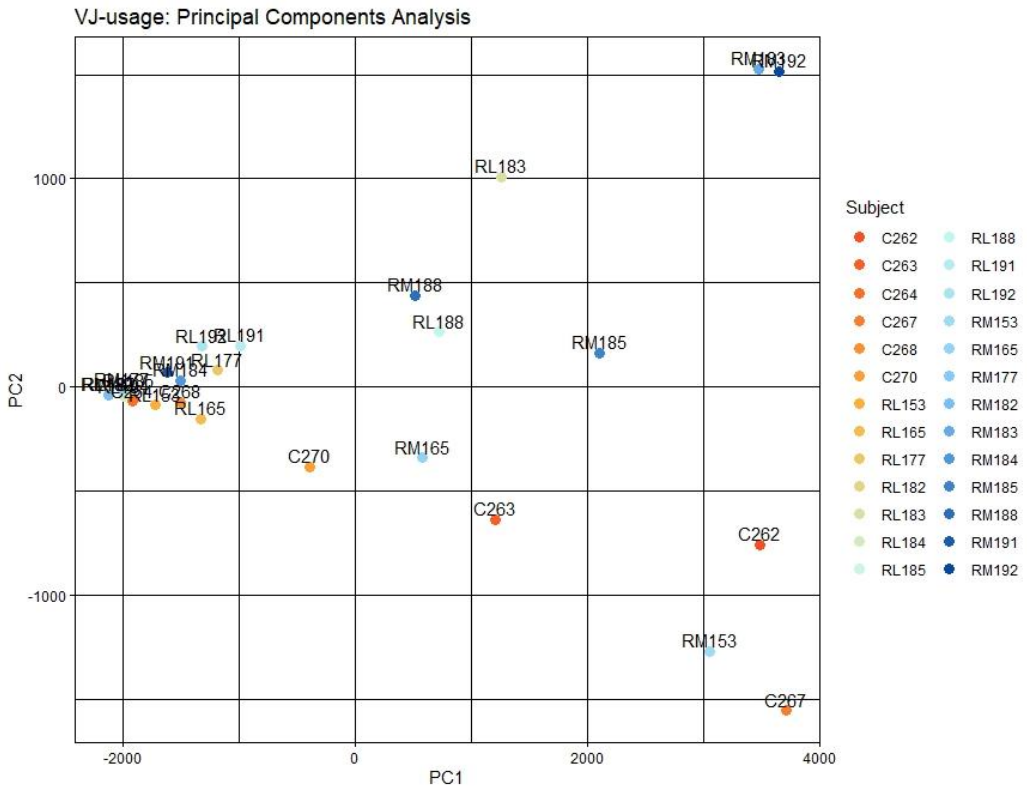


Fig. 3. IGH-VJ segment use. Analysis of the VJ-segment use profile of IGH revealed that relapse patient´s sequences seemed to cluster together while those from remission and control patients were scattered all over the graph.

- IGHV gen use

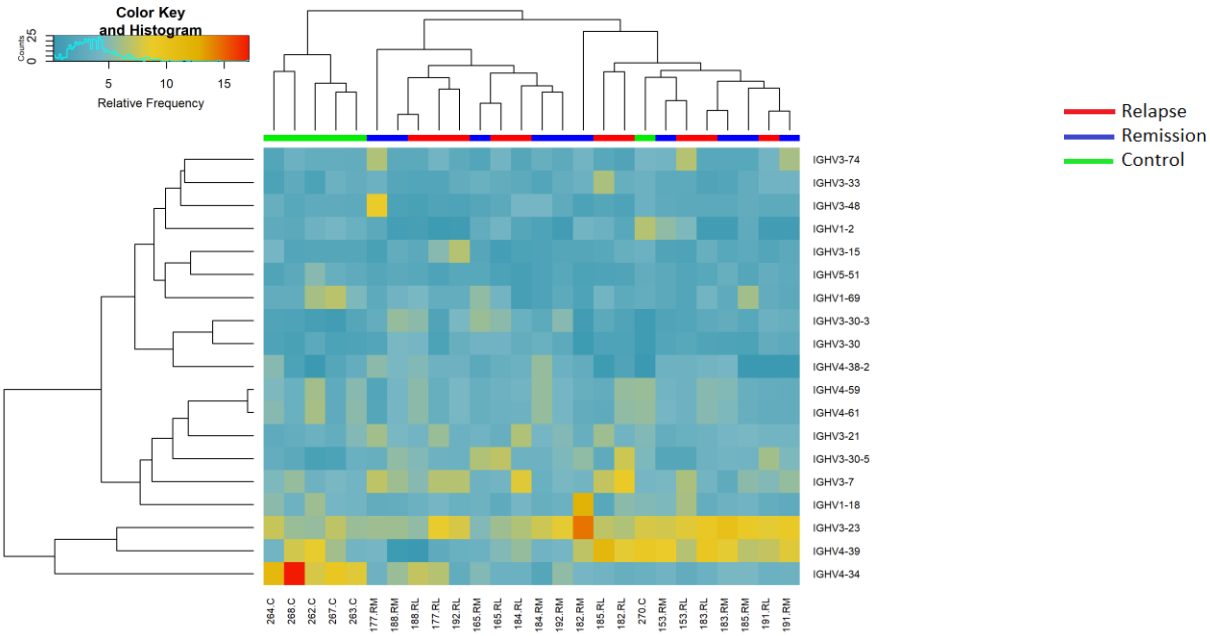


Fig. 4. IGHV gen use. Analysis of the V-segment use profile revealed that controls were separated from MS patients. In addition, IGHV4-34 gene seem to predominate in controls, while the IGHV3-23 gene showed a higher frequency in MS patients.

- IGH SHM

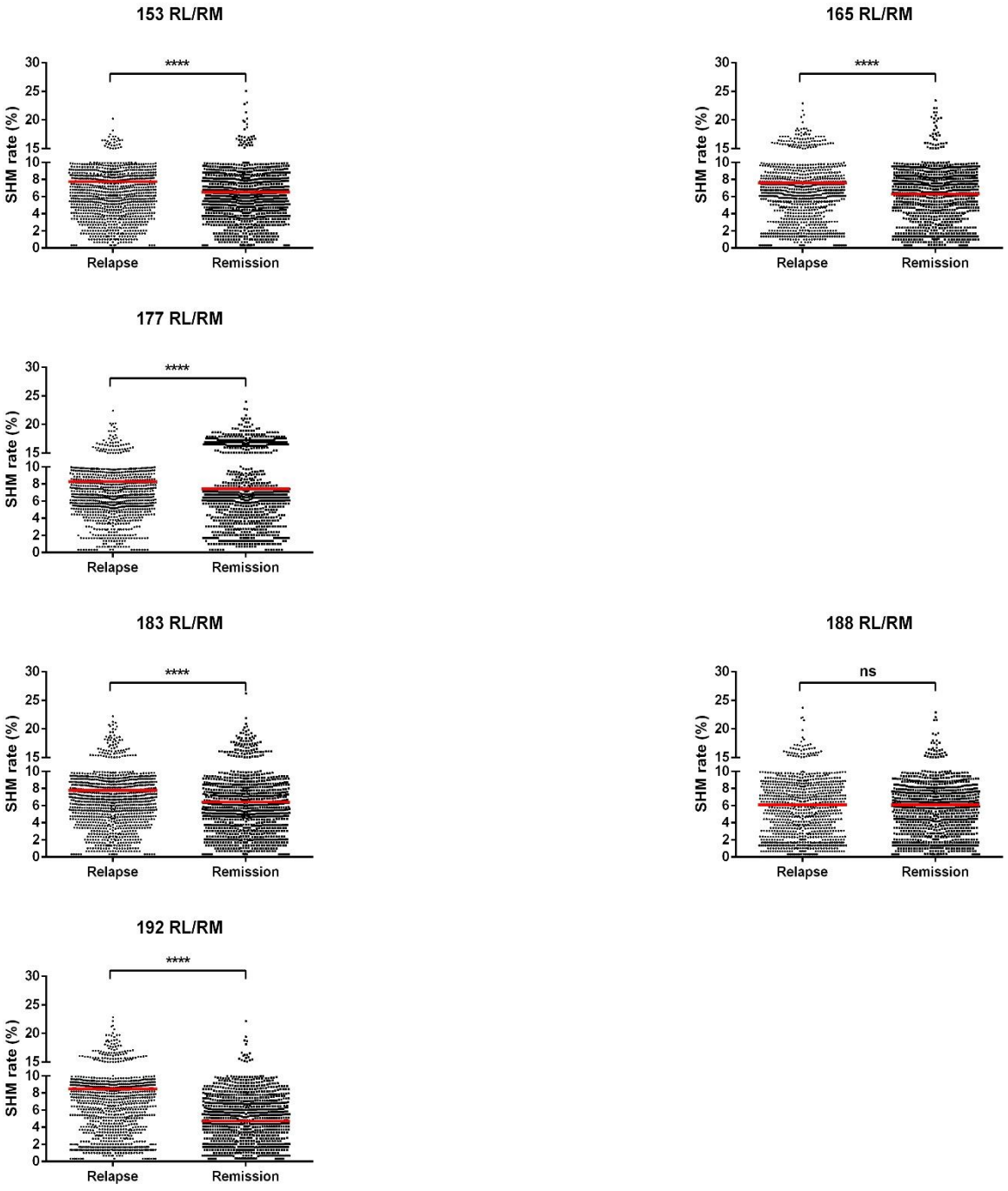


Fig. 5. IGH SHM of 6 Relapse/Remission patients. Each dot represents the SHM rate (%) of one clonotype, and red lines correspond to the median values. A higher SHM rate was observed in MS patients during relapse. ****P< 0.001

Preliminary results of TCR repertoire analysis

- Diversity Index

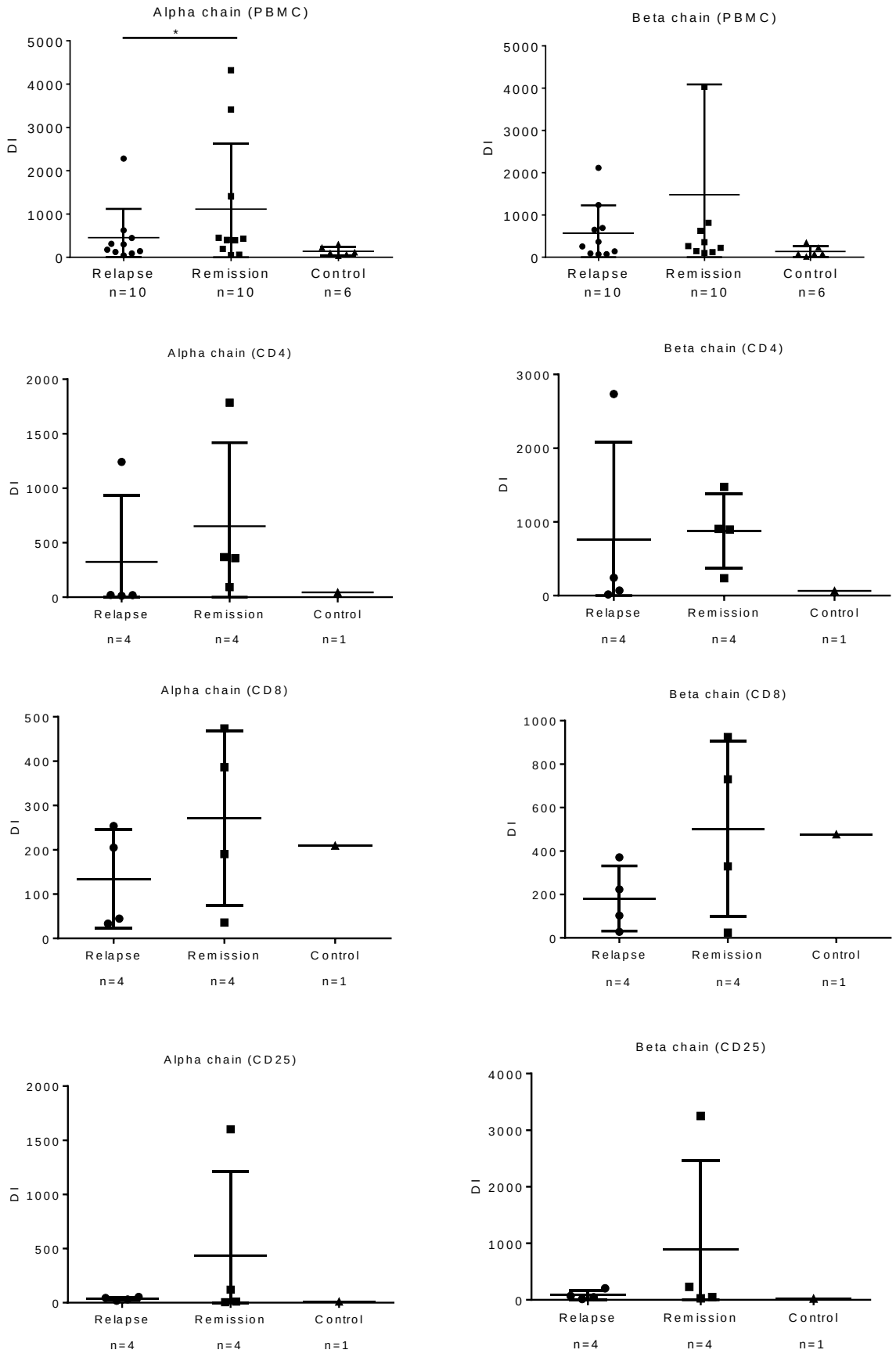
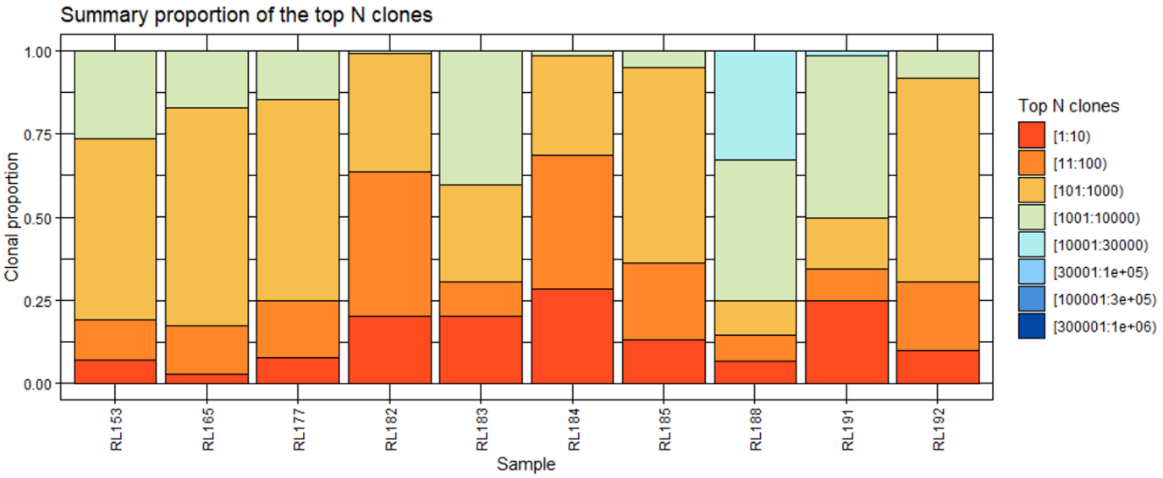


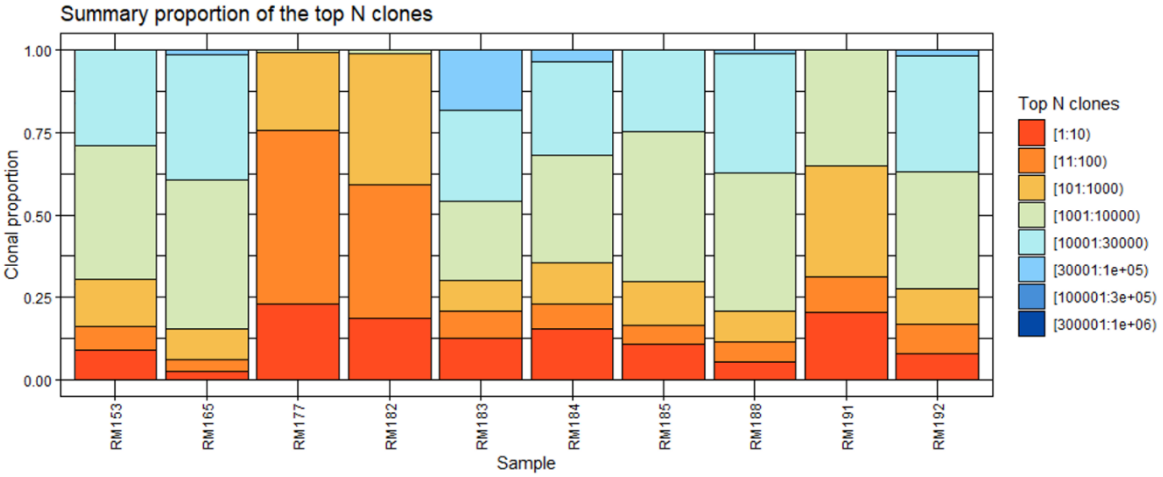
Fig. 4. Inverse of Simpson's diversity index of PBMC, $CD4^+$, $CD8^+$ and $CD4^+/CD25^+$ populations for alpha and beta chains. A lower diversity index was observed in the relapse phase of MS patients compared to remission and controls; however, no statistical difference was obtained with the Mann-Whitney test.

- Proportion of clonotypes

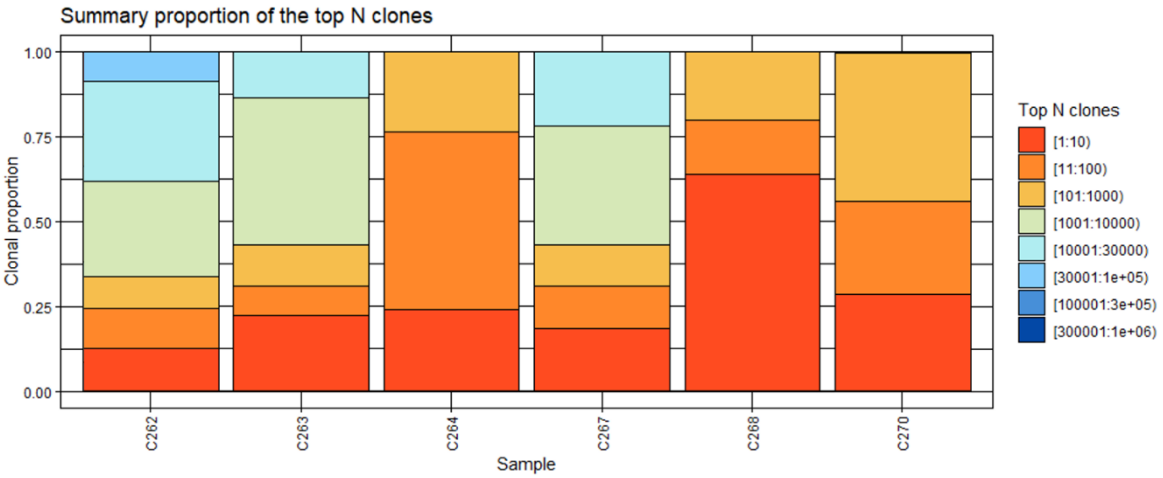
A) Relapse TRB.



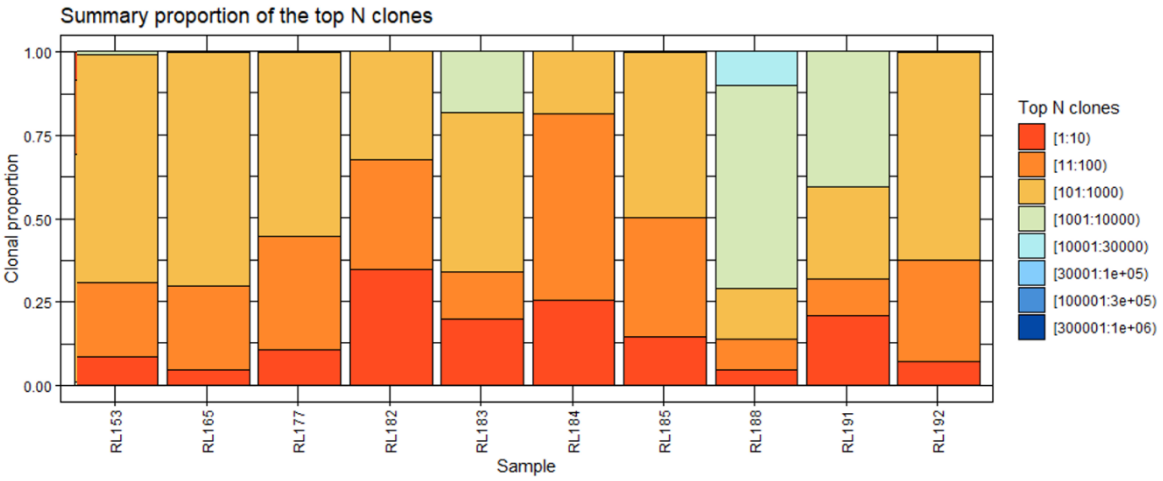
B) Remission TRB.



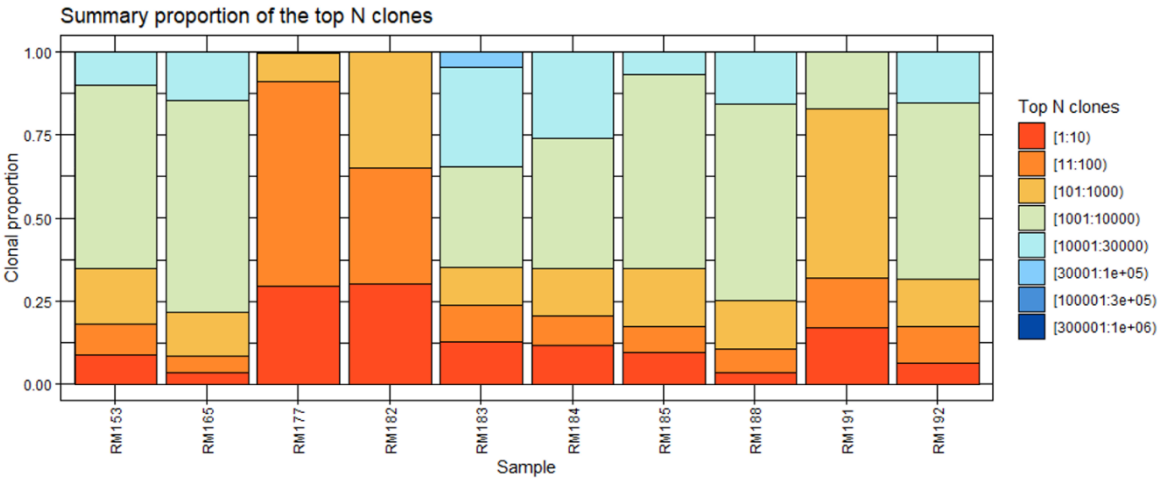
C) Control TRB.



D) Relapse TRA.



E) Remission TRA.



F) Control TRA.

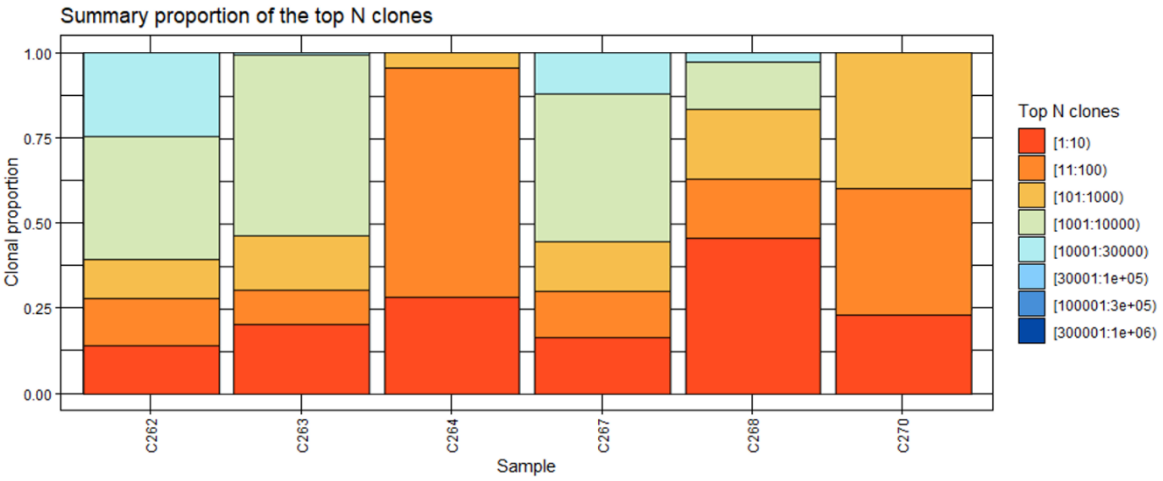
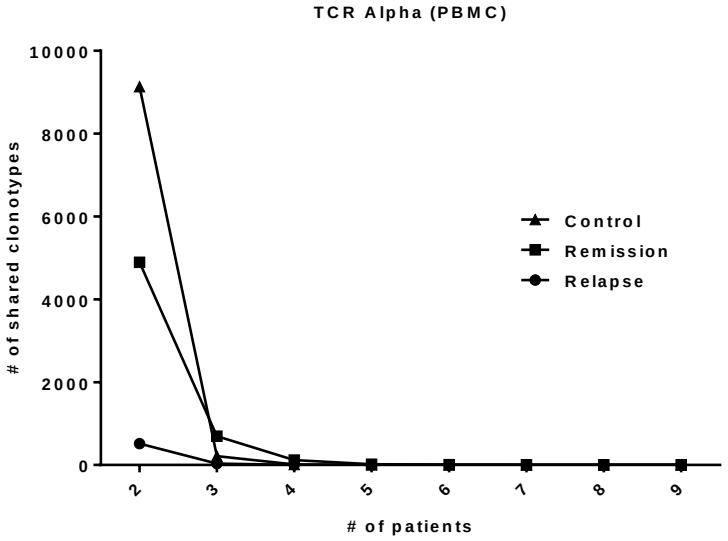


Fig. 5. Top proportion of clonotypes. The number of clonotypes that represents 50% of the entire repertoire for both chains is between 100 and 1,000 different clonotypes for relapse

phase, but in the case of remission phase, the same percentage of repertoire is represented for 1,000 to 10,000 clonotypes.

- Shared clonotypes



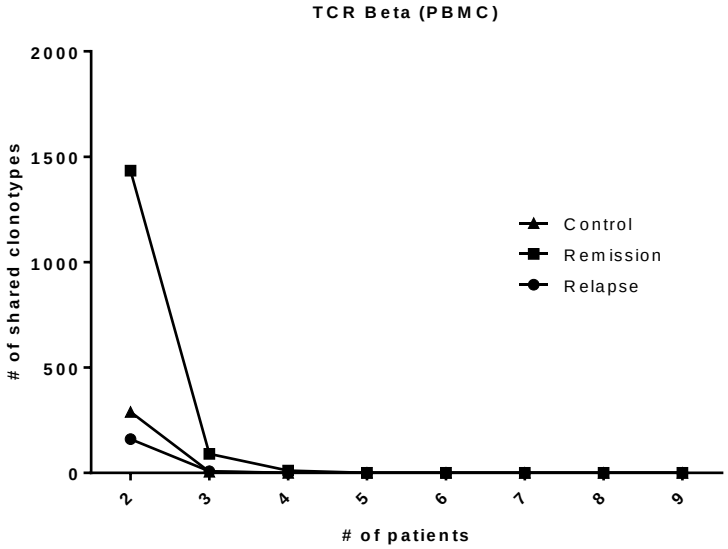
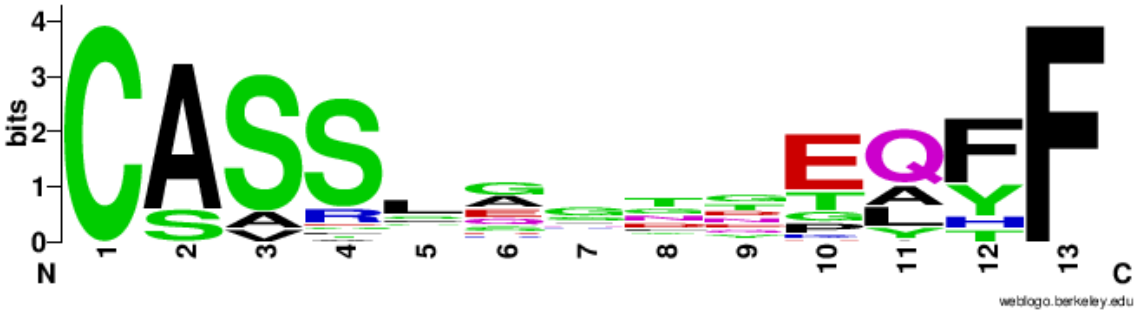


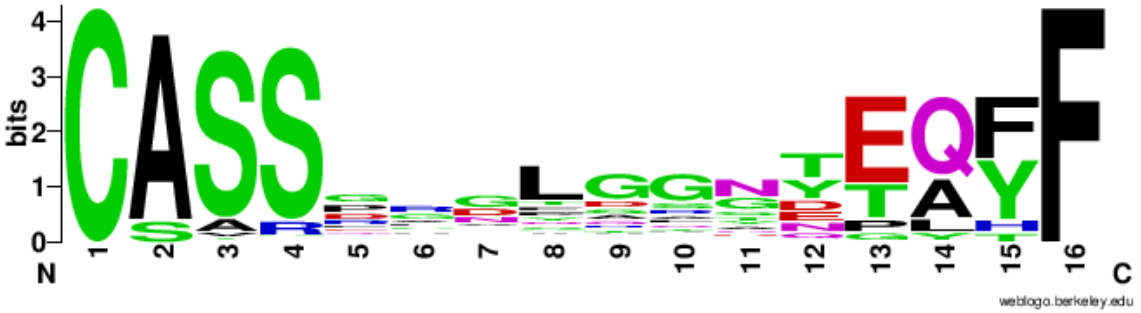
Fig. 6. Shared clonotypes between phases. Related to the diversity index, which is higher in the remission phase and the top clonotype representing the repertoire, it is in this remission phase where a greater number of clonotypes are shared among patients for both chains.

- Sequence logos.

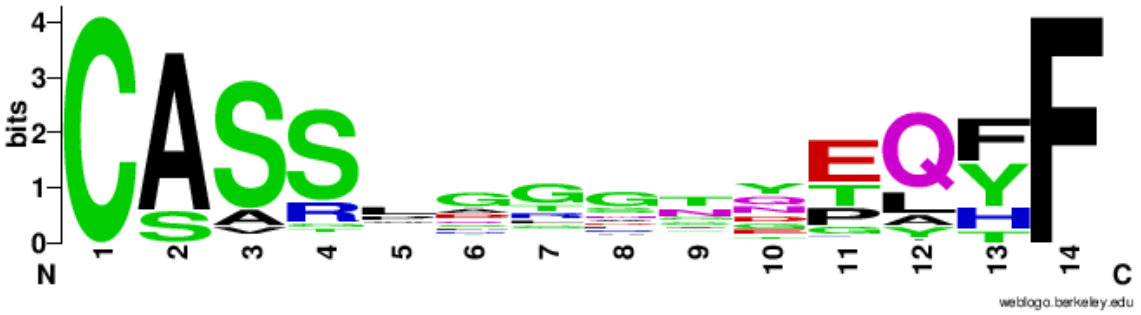
A) Relapse TRB.



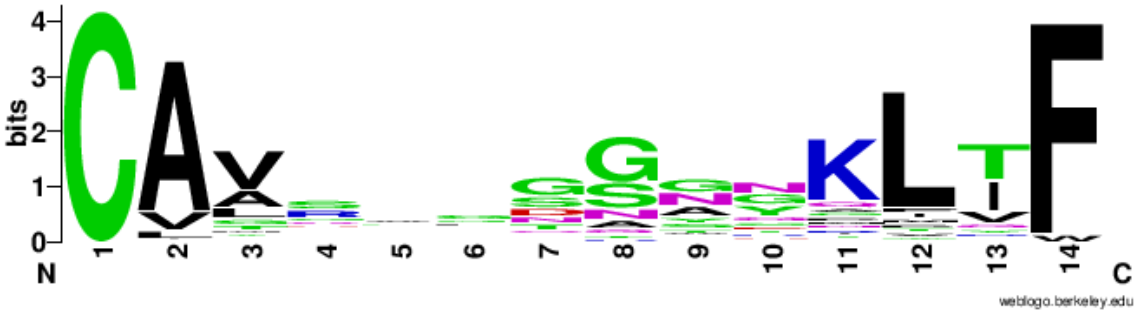
B) Remission TRB.



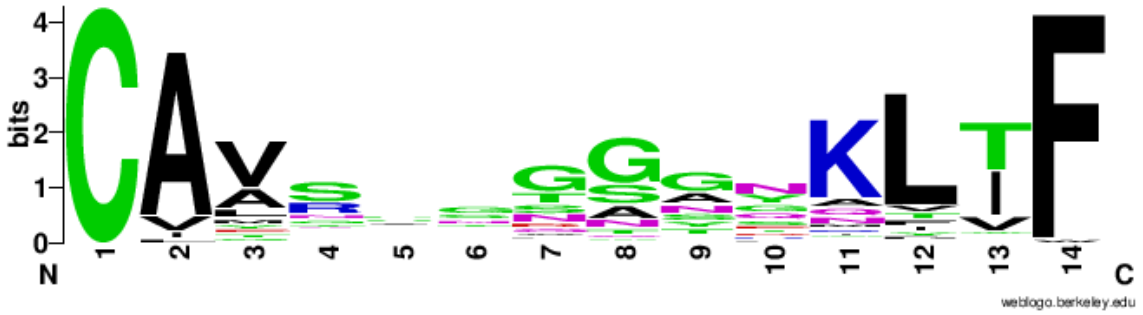
C) Control TRB.



D) Relapse TRA.



E) Remission TRA.



F) Control TRA.

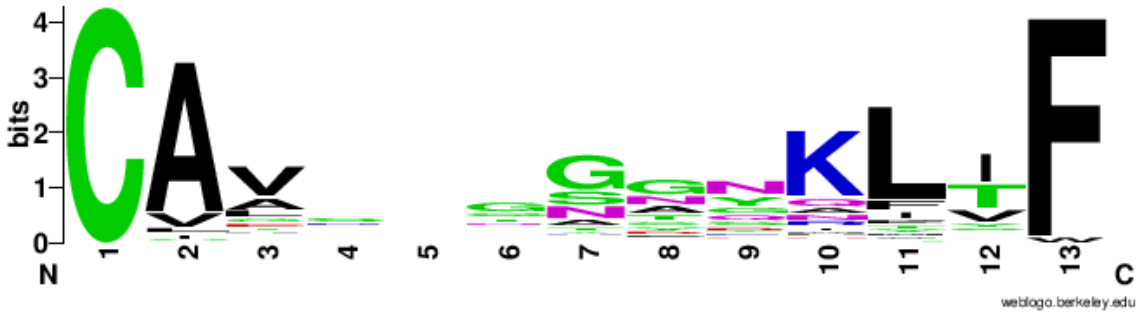
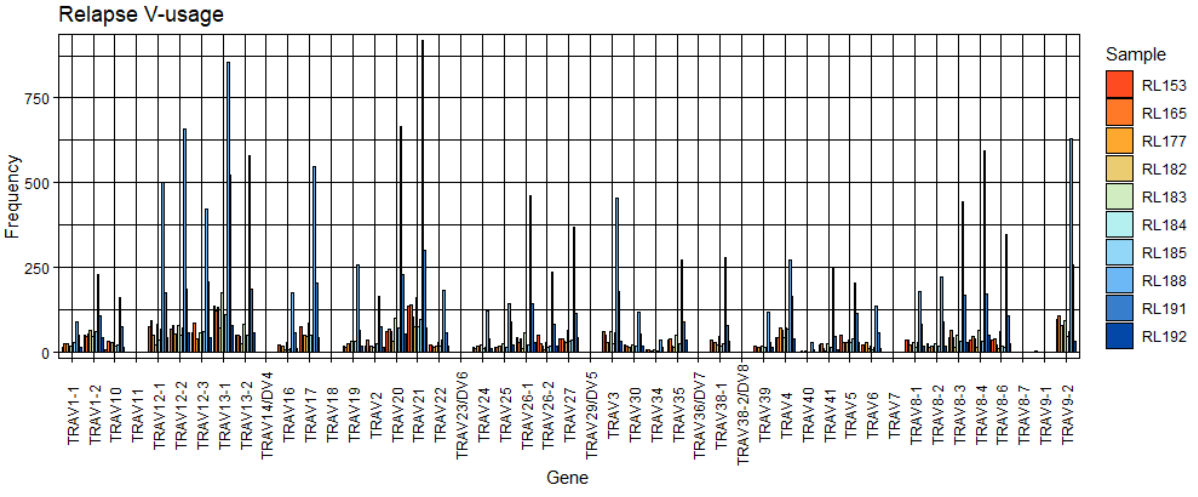
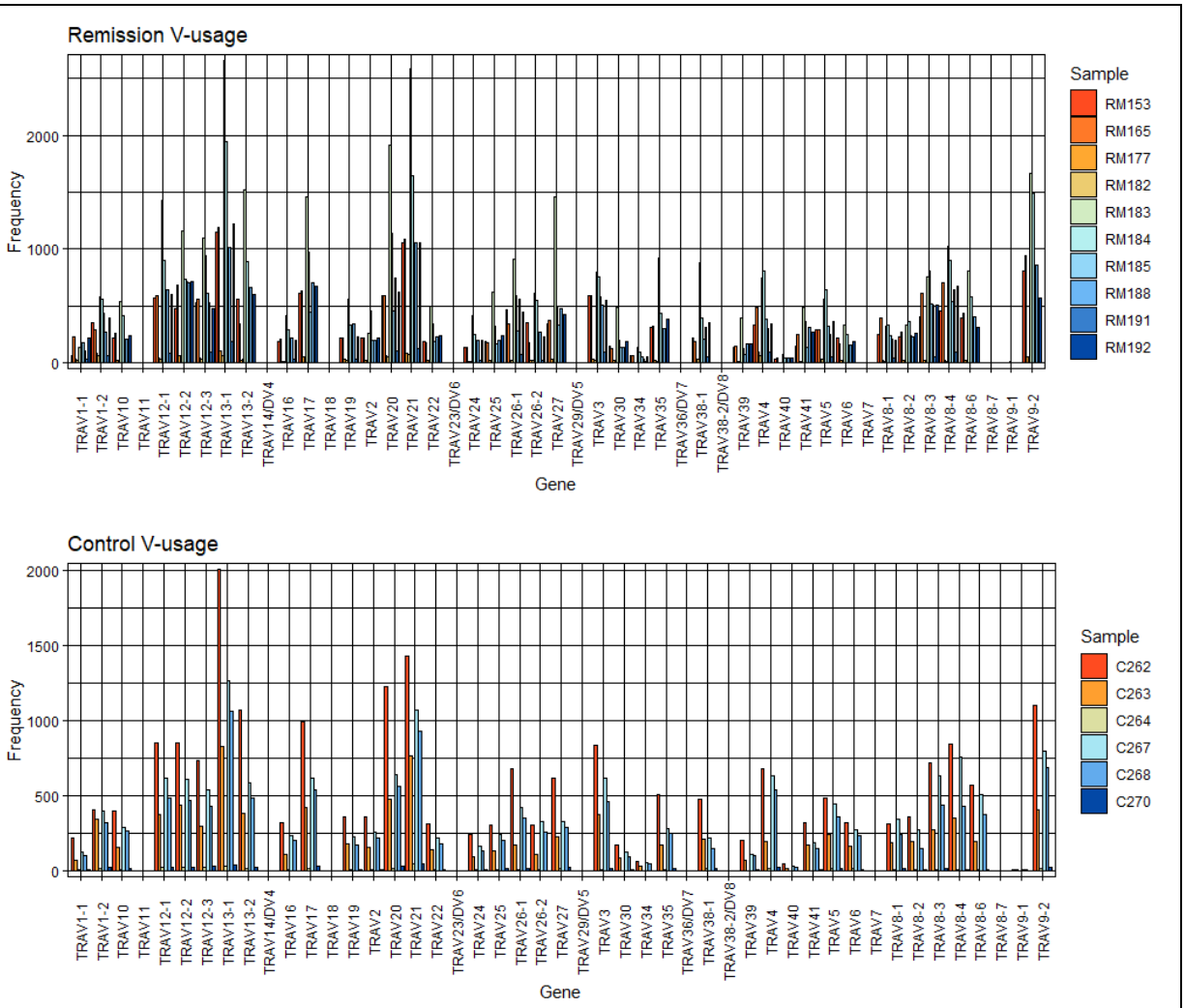


Fig 7. CDR3 sequence alignment. Logos from shared clonotypes between phases were observed. The logos from alpha chain CDR3 sequences were not different between the three groups; however, logos for beta chain show differences of length and amino acid position.

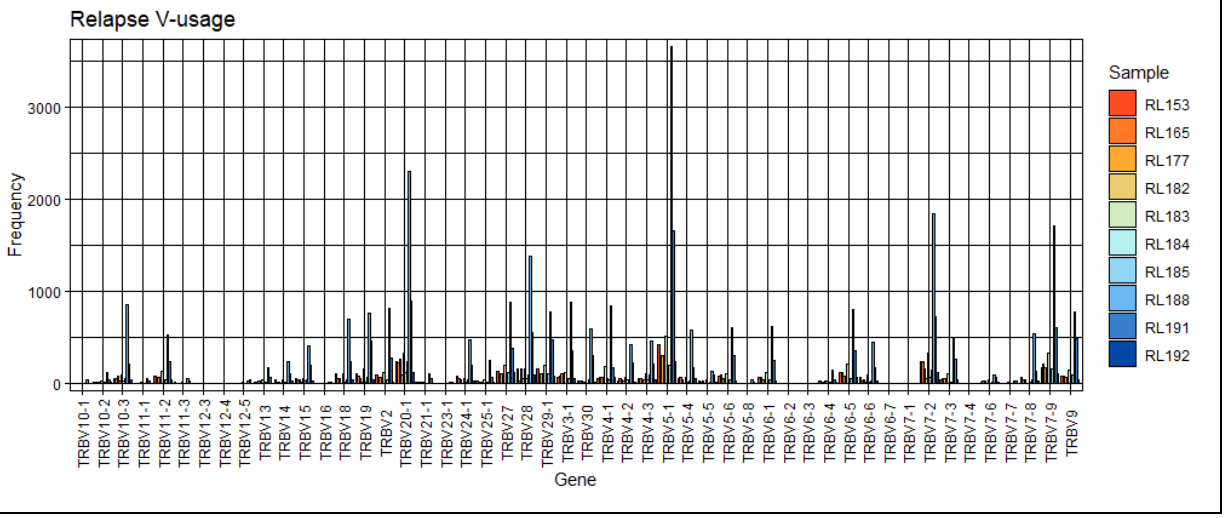
- TRBV gen use

A) Alpha chain




Remission V-usage
Frequency
Gene
Sample
RM153
RM165
RM177
RM182
RM183
RM184
RM185
RM188
RM191
RM192
Control V-usage
Frequency
Gene
Sample
C262
C263
C264
C267
C268
C270

B) Beta chain


Relapse V-usage
Frequency
Gene
Sample
RL153
RL165
RL177
RL182
RL183
RL184
RL185
RL188
RL191
RL192

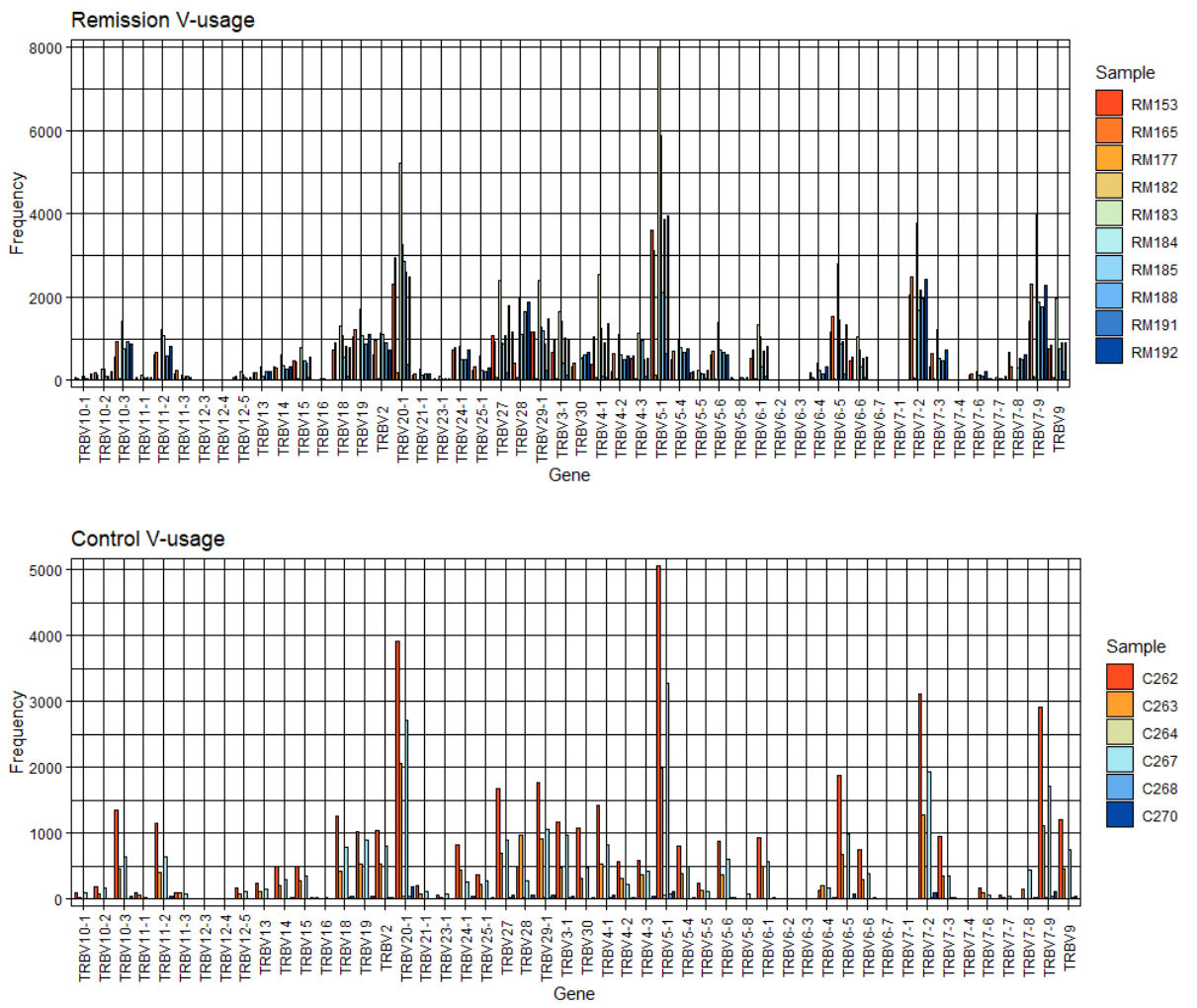


Fig 8. V gene usage. For beta chain, the most V gense used are 20-1, 28, 5-1, 7-2 and 7-9; for alpha chain, 21, 13-1, 20, 12-2 and 9-2. Most of the genes, are not reported previously as common gen among MS patients.

- Main components analysis.

A) PCA TRA

B) PCA TRB

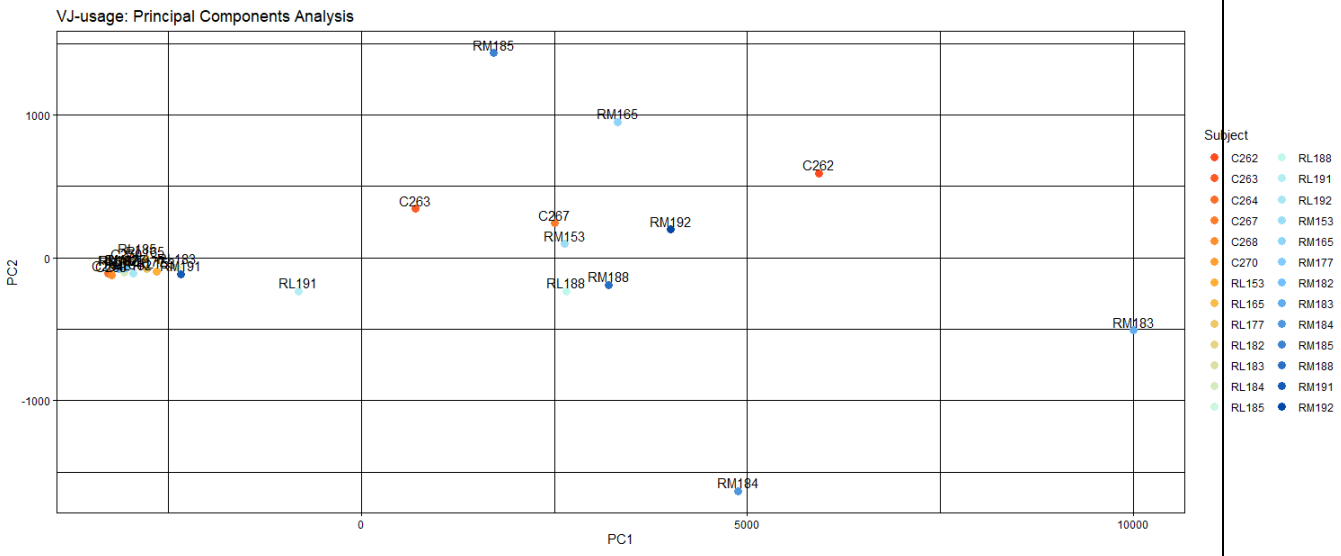


Fig 9. Principal components analysis. For both chains, most relapse samples are grouped into a single cluster, while remission and control samples are dispersed, indicating a correlation between the use of VJ segments among relapse patients.

Preliminary conclusions

Results suggest a clonal expansion in the BCR repertoire of relapse MS patients, although there is still analysis to be performed. The IGHV3-23 gene was identified as predominant in Mexican patients with MS, which could eventually serve as a molecular marker, provided more samples confirm this finding.

A higher diversity index was observed in patients in the remission phase compared to patients in relapse phases from PBMCs and T cell populations, which correlates with the SHM, since we found higher SHM rate in MS patients during relapse (due to clonal expansion).

A trend of lower diversity of TCR repertoire among patients in the relapse phase is observed compared to the same patients in the remission phase. In addition, in the relapse phase, fewer clones represent the repertoire of T-cell receptors and fewer clones are shared among these patients.

The logos demonstrate that the clonotypes in the remission phase are similar to those of control patients, while those observed during relapse are different in length and composition, suggesting the participation of these clonotypes in the acute phase of the disease.

Although gene usage is similar in all three groups, only relapse samples are grouped into a single cluster, indicating a correlation between them.

From the preliminary results, we can conclude that there is no common clonotype among MS patients, neither in relapse nor in remission. Therefore, the next step in the analysis of BCR and TCR repertoires is to identify predominant clonotypes per patient and predict the peptide that they recognize.

Considering the world sanitary emergency, no travel was allowed last year. Therefore, following IMSUT´s suggestion, we asked to have the reagents shipped from Japan to Mexico, to conclude this analysis. A group of healthy controls will be included this time, in order to compare the BCR and TCR repertoires of healthy people with MS patients in both relapse and

remission.

Collaboration between Cinvestav in Mexico and IMSUT and JFCR in Tokyo has been very fruitful. This collaboration has been fundamental in teaching Mexican students how to apply the bioinformatic tools developed in IMSUT to answer the questions of how the immune response is modulated in MS, through the analysis of BCR and TCR in MS Mexican patients. Once the remaining samples are sequenced in Mexico with the reagents sent from IMSUT, we will prepare the papers that will allow the students to obtain their PhD degree.