Hierarchical Clustering of B-Cell Receptor Structures in Splenic Marginal Zone Lymphoma

Marcatili, Paolo; Zibellini, Silvia; Rattotti, Sara; Chailyan, Anna; Varettoni, Marzia; Morello, Lucia; Boveri, Emanuela; Lucioni, Marco; Bonfichi, Maurizio; Gotti, Manuel; Fiaccadori, Valeria; Paulli, Marco; Tramontano, Anna; Arcaini, Luca

Publication date: 2012

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Splenic marginal zone lymphoma (SMZL) is frequently associated with HCV infection and autoimmune disorders. Previous studies demonstrated a biased usage of immunoglobulin heavy variable genes (IGHV) and, in some cases, stereotyped B-cell receptors (BCRs). This characterization, however, is mainly based on the heavy chain alone, even if strong evidences are emerging on the role of light chain (Bikas et al. Leukemia 2012).

The aim of this study was to analyze IG light variable genes (IGLV) of SMZL BCRs, VL-VH pairing and structural information and to investigate the sequence-structure-antigen (AG) relationship. To this end, we analyzed the VL-VH paired sequences of BCR from 52 SMZL pts (38 BM and 14 PB) diagnosed according to Matutes criteria (Leukemia, 2008). Sequences were analyzed using the IMGT/DBs and the IMGT/V-QUEST tool. The PIWS web server was used to build 3-D models of all antibodies (Abs). The Abs structures were compared using LGA and clustered together according to a score accounting for structure and sequence similarity. Using the DIGIT DB and tools, all the clusters were analyzed and compared to other IGs.

Based on the IGHV nucleotide sequence identity to the germline, 7 sequences (13%) were considered ‘truly unmutated’ (100% sequence identity). 20 (39%) were ‘minimally or borderline mutated’ (97-99%) whereas 25 (48%) were ‘significantly mutated’ (<97%). IGHV families were used as follows: IGHV3 (58%), IGHV1 (27%) and IGHV4 (15%). The majority of pts carried kappa light chain (69%). The most frequently used IGLV families were IGLV3 (58%) and IGLV1 (28%), the most frequent IGLV family was IGLV1 (56%). The VL-VH paired sequences, the two pairings IGHV3-23/IGKV3-20 (n=6) and IGHV1-02/IGLV1-47 (n=3) were significantly over-represented when compared to CLL and DIGIT DB sequences, indicating that the pairing between VL-VH chains was non-random. The IGHV1-02/IGLV1-47 paired sequences showed a high number of somatic mutations (>3%), whereas samples using the IGHV1-02 gene (n=10) but a VL gene other than from IGLV1-47 displayed a low number of mutations, suggesting a significant role for the light chain.

In order to analyze the possible functional role of light chain, we analyzed the structural similarity of AG binding sites (ABSs), performing hierarchical clustering on the similarity obtained by an all-against-all structural superposition of each ABS. Twenty structural clusters were identified (8 with ≥ 3 samples) (Fig. 1). Considering IGs in the same major groups, they showed a similar mutation rate, pointing out a likely common AG selection at least in a fraction of pts (Fig. 1). In most cases, IGs in the same clusters display ABSs with similar physicochemical characteristics: positively charged binding sites (2 clusters), hydrophobic patches (3 clusters) or small pockets in the middle of the ABS (3 clusters) might be clue for different AGs specific for each cluster. HCV infection was found in 1 major and 2 minor clusters (Fig. 1), mainly associated with unmutated clones, indicating a likely common antigenic stimulation. In the other major clusters, the role for an AG-driven selection different from HCV in SMZL lymphomagenesis can be postulated. In particular, 3 clusters, containing both mutated and unmutated samples, displayed a statistically significant similarity to CLL clones (p<0.05), and 1 cluster was structurally similar to autoimmune clones (Kawasaki disease) (p=0.05). Of note, other clusters showed a degree of similarity with samples connected to diseases that involve an AG independent or superantigenic stimulation (EBV, Rabies virus, Rotavirus).

In conclusion, the multi-layered characterization of the sequence and structure properties of paired VL-VH in SMZL identified a non-random pairing between heavy and light chains. Structural cluster analysis identified Abs with similar physicochemical properties, similar mutation rate and similar HCV status in a fraction of our dataset. Comparing Abs of our cases to a large dataset of human annotated Abs derived from the DIGIT DB, a subset resulted similar to CLL or autoimmune clones, whereas other Abs appeared more similar to polyreactive Abs and to Abs possibly targeted by superantigens. These findings could explain the large diversity observed in the IGs expressed in SMZL and provide new insights in SMZL pathogenesis.

*The first two authors equally contributed to this paper