The adaptive immune response is characterized by an extreme diversity of the antigen receptors, and this is at the origin of immunoinformatics, a science at the interface between immunogenetics and bioinformatics. Immunoglobulins (IG) or antibodies [2] and T cell receptors (TR) [4] are managed and described in the IMGT® databases and tools at the level of receptor, chain and domain. The analysis of the IG and TR variable (V) domain rearranged nucleotide sequences is performed by IMGT/V-QUEST (since 1997), 50 sequences per batch [5] and, for next generation sequencing (NGS), by IMGT/HighV-QUEST, the high throughput version of IMGT/V-QUEST (portal begun in 2010; 500,000 sequences per batch) [6, 7]. The analysis of NGS scFv represents a challenge by their length (~850 bp) as they contain two V domains connected by a linker and there is no tool for the analysis of two V domains in a single chain. The functionality “Analysis of single chain Fragment variable (scFv)” has been implemented in IMGT/V-QUEST and, for NGS, in IMGT/High-V-QUEST for the analysis of the two V domains of IG and TR scFv [8]. For each sequence or NGS read, positions of the V-DOMAIN, linker and 3’-DOMAIN in the scFv are provided in the “V-orientated” sense. Each V-DOMAIN is fully characterized (gene identification, sequence description, junction analysis, characterization of mutations and amino changes). The functionality is generic and can analyse any IG or TR single chain nucleotide sequence containing two V domains, provided that the corresponding species IG reference directory is available. Nowadays, advances in NGS technology allow for longer reads (1000 bp and more), therefore full-length scFv. The positions of the 5’V-DOMAIN, linker and 3’V-DOMAIN in the scFv are provided in the ‘V-orientated’ sense. Each V-DOMAIN is fully characterized (gene identification, sequence description, junction analysis, characterization of mutations and amino changes). The functionality is generic and can analyse any IG or TR single chain nucleotide sequence containing two V domains, provided that the corresponding species IG reference directory is available. Nowadays, advances in NGS technology allow for longer reads (1000 bp and more), therefore full-length scFv.

The new functionality “Analysis of single chain Fragment variable (scFv)” provides the identification and full characterization of the two V-DOMAIN of full-length scFv by IMGT/V-QUEST online or, for NGS, by IMGT/HighV-QUEST. This functionality for scFv sequence analysis is generic for IG and TR, and to our knowledge, it is proposed for the first time by IMGT online tools, only. This functionality was used to analyse more than 450,000 scFv sequences from a combinatorial phage library, the sequencing reads of about 1000 bp were obtained with the PacBio Biosciences (PacBio) RS platform using single-molecule real-time (SMRT) circulating sequencing sequencing (CCS). The two V domains were identified and fully characterized in 89% of the ~349,000 reads for their sequencing quality and length. The “Analysis of single chain Fragment variable (scFv)” will facilitate and improve the description of the scFv content of combinatorial libraries, a key information in therapeutic antibody discovery, selection and development.

**Perspectives**

The need for the analysis of NGS sequences containing two V domains from IG or TR expressed repertoires is also rapidly rising, with novel methodological advances, as illustrated by single-cell sequencing of paired chains, paired recovery of transcripts and concatenation per single cell, or capture strategies. As IMGT/High-V-QUEST is generic for IG and TR, the functionality for the Analysis of single chain Fragment variable (scFv) can be used, without any change, for the characterization of the two V domains of various NGS single chains (IG or TR) which mimic the V domain pairing of the natural antigen receptor binding sites. It is expected that this will facilitate the identification of novel paratopes in infections, cancers, autoimmune diseases or neurodegenerative diseases.