Human Immunology 80 (2019) 335–338

Contents lists available at ScienceDirect

Human Immunology

journal homepage: www.elsevier.com/locate/humimm

Genotyping of Italian patients with Behçet syndrome identified two novel *ERAP1* polymorphisms using sequencing-based approach

Maria Carmela Padula[a,b], Pietro Leccese[a,*], Nancy Lascaro[a], Teresa Carbone[a], Michele Gilio[a], Angela Anna Padula[a], Giuseppe Martelli[b], Salvatore D'Angelo[a,c]

[a] Rheumatology Institute of Lucania (IReL) and Rheumatology Department of Lucania, San Carlo Hospital of Potenza, via Potito Petrone, Potenza 85100, Italy
[b] Department of Science, University of Basilicata, viale dell'Ateneo Lucano, Potenza 85100, Italy
[c] Basilicata Ricerca Biomedica (BRB) Foundation, Matera 75100, Italy

ARTICLE INFO

Keywords:
Behçet syndrome
ERAP1
Genotyping
Polymorphisms

ABSTRACT

The endoplasmic reticulum aminopeptidase protein 1 gene (*ERAP1*) is related to several human diseases, including Behçet syndrome (BS), a multisystemic disorder with unknown etiology. ERAP1 is involved in immune response and its role can be influenced by gene single nucleotide variations (SNVs).

We genotyped the *ERAP1* whole structure in 50 consecutive BS patients and 50 ethnically-matched healthy controls using both bioinformatics and molecular methodologies.

We identified two novel heterozygous missense SNVs of *ERAP1* exon3 responsible for the p.Glu183Val and p.Phe199Ser changes. The first variation was recognized in 7/50 (14%) BS patients and involved the substrate binding site (p.Glu183) required for the anchorage of the peptide N-terminal group. The SNV was predicted to be a damaging variation, as well as the p.Phe199Ser substitution (PolyPhen-2 and SIFT on line software). 3D protein structure prediction showed a change in energy score when the wild-type and the variant states were compared, probably influencing the substrate binding and the protein folding. The first variation was associated to a more stable protein chain, while the second polymorphism was related to a less stable protein chain.

Our data need to be tested in larger genetic studies.

1. Introduction

Behçet syndrome (BS) is a multifactorial disorder with several clinical manifestations, including recurrent oral and genital ulcers, ocular inflammation, gastrointestinal ulceration and skin lesions. Both genetic and environmental factors contribute to BS susceptibility and clinical phenotype [1–4]. Although the Human Leukocyte Antigen-B51 (HLA-B*51) was described as the major BS susceptibility locus [5], its role in disease pathogenesis remains elusive and only partially explains the genetics risk of BS [6,7]. Non-HLA polymorphisms have also been investigated and their association with BS has been demonstrated for several genes, with particular focus on the endoplasmic reticulum aminopeptidase protein 1 gene (*ERAP1*) [3,4]. *ERAP1* is formed by 20 exons and encodes a 941 residues amino-peptidase with a central role in peptide trimming, a step required for the generation of many HLA class I-binding peptides. The enzyme processes N-terminally extended antigenic precursors for optimal loading onto HLA molecules. The trimming specificities have been related to the presence of several single nucleotide variations (SNVs), leading to efficient, hypoactive and hyperactive phenotypes. The different SNVs role in the enzymatic function is not clear: some variations may interact with the active site, others polymorphic sites may influence the dynamics of the open-closed conformational change, others are outside the active site and their function is very difficult to define [6–12].

The most significant previously found *ERAP1* SNVs are: rs17482078 (p.Arg725Gln), rs2287987 (p.Met349Val), rs30187 (p.Lys528Arg), rs10050860 (p.Asp575Asn), and rs27044 (p.Gln730Glu). The polymorphisms genotyping in several populations, in particular in Turkish [11,13] and Spanish [14] cohorts were consistent with the association between *ERAP1* and BS, as well as with an epistatic interaction between

Abbreviations: BS, Behçet syndrome; ERAP1, Endoplasmic reticulum aminopeptidase protein 1; HC, Healthy Controls; HCA, hydrophobic cluster analysis; HGVS, Human Genome Variation Society; HLA, Human Leukocyte Antigen; ISG, International Study Group; PCR, Polymerase Chain Reaction; PolyPhen-2, Polymorphism Phenotyping v2; SIFT, Sorting Intolerant from Tolerant; SNVs, single nucleotide variations

* Corresponding author.
E-mail addresses: mcpadula25@gmail.com (M.C. Padula), pietroleccese1979@gmail.com (P. Leccese), nancylascaro@libero.it (N. Lascaro), carbone_teresa@libero.it (T. Carbone), angela.padula@ospedalesancarlo.it (A.A. Padula), giuseppe.martelli@unibas.it (G. Martelli).

https://doi.org/10.1016/j.humimm.2019.02.003
Received 15 October 2018; Received in revised form 7 February 2019; Accepted 7 February 2019
Available online 08 February 2019

Table 1
Demographic features and clinical manifestations of Italian patients with Behçet's syndrome included in the study.

Demographics	BS patients, n = 50
Age, years	46.10 ± 12.19
Male/Female	29/21
Clinical manifestations	
Oral ulcers, n (%)	50 (100%)
Genital ulcers, n (%)	21 (42%)
Skin disease, n (%)	42 (84%)
Ocular involvement, n (%)	30 (60%)
Neurologic involvement, n (%)	18 (36%)
Vascular involvement, n (%)	10 (20%)
Joint involvement, n (%)	29 (58%)
HLA-B51 positivity	34 (68%)

Abbreviations: BS, Behçet's syndrome; n, number of subjects.

ERAP1 and HLA-B [14].

2. Materials and methods

2.1. Patient recruitment

A total of 50 consecutive BS patients (29 males and 21 females) with a mean age of 45.78 years (range: 26–67 years) were recruited at Rheumatology Institute of Lucania (IReL), a tertiary BS outpatients clinic. The patients met the International Study Group (ISG) diagnostic criteria of BS, consisting in the presence of recurrent oral ulcerations plus two of the following criteria: recurrent genital ulcerations, eye lesions, skin lesions, positive results from a pathergy test. Patient demographic and clinical features were reported in Table 1.

A group of 50 Italian healthy controls (HC), 28 males and 22 female with a mean age of 44.31 years (range: 27–65 years), was selected among unrelated hospital and university employees and recruited after ruling out any history of inflammatory and autoimmune rheumatic diseases. Prior to enrolling in the present study, all subjects provided their written, informed consent. The Regional Ethics Committee approved the study (Permit Number: 705/2017).

2.2. Genotyping

Specific primers for the coverage of all *ERAP1* gene regions (20 exons, intron-exon boundaries, and 5′ and 3′ UnTranslated Regions) were designed using the NCBI Primer-Blast tool (Table 2). Genomic DNA was extracted from whole blood using EuroGOLD Blood DNA Mini kit (Euroclone®, Italy). DNA purity and concentration were determined using the NanoDrop™ 1000 spectrophotometer (NanoDrop Technologies, Inc). PCR amplification was carried out using the Q5 Hot Start High-Fidelity DNA Polymerase (BioLabs Inc, New England) according to the manufacturer's recommendations. The following conditions were used: 1) initial denaturation: 98 °C/5 min; 2) thermocycling: 98 °C/1 min; 60 °C/1 min; 72 °C/2 min (30 cycles); 3) final extension: 72 °C/5 min. PCR products were analyzed by gel electrophoresis (1.5% agarose gel). Good-quality PCR amplicons were directly sequenced using the GATC Biotech Sanger sequencing service and bioinformatically analysed using Mutation Surveyor software (SoftGenetics, USA) and NCBI-Blast Nucleotide on line similarity search tool.

2.3. Downstream bioinformatics analysis

The functional impact of the *de novo* allelic variants were predicted using on line Polymorphism Phenotyping v2 (PolyPhen-2) software [15] and Sorting Intolerant from Tolerant (SIFT) algorithm [16]. The 3D protein modelling was obtained using Protean 3D program (LaserGene software) [17] based on 3QNF Reference sequence (Protein Data Bank, PDB).

3. Results

We identified two *de novo* gene variations. The first novel variation was a heterozygous SNV at 18169 nucleotide position (NG_027839.1:g.18169A > T; HGVS nomenclature) (Fig. 1a) that was found in 7/50 (14% of cases) BS patients. This SNV was a missense variation responsible for the substitution glutamate to valine at 183 amino acid site (NP_057526.3:p.Glu183Val; HGVS nomenclature) (Fig. 1b). The change was predicted to be damaging with maximum score when two prediction software were queried (PolyPhen-2 score: 1.00; SIFT score: 0.00).

A second variation was found within the same exon in 3/50 patients (6%) carrying also the first SNV. The second substitution was a heterozygous thymine to cytosine at 18217 nucleotide position (NG_027839.1:g.18217 T > C; HGVS nomenclature) (Fig. 1a). This site was characterized by the rs146396644 known variation (NG_027839.1:g.18217 T > G; NP_057526.3:p.Phe199Cys; HGVS nomenclature). The novel change led to the amino acid phenylalanine to serine substitution (NP_057526.3:p.Phe199Ser; HGVS nomenclature) (Fig. 1b) and showed a predicted pathogenic effect (Poly-Phen2 score: 0.99; SIFT score: 0.00).

Table 2
Primer pairs designed for the coverage of all *ERAP1* regions.

Primers name	Primers sequence (Forward – Reverse)
ERAP 1a	5′TCGGTCCCCAACTTGAGC3′ – 5′GGGAAGCCCCTGGCTAA3′
ERAP 1b	5′TCCCTTCCTGGGAGTTGC3′ – 5′CCCAGGCCCTCAGCC3′
ERAP 2	5′TTTCTGTGAATGCTGGGTGG3′ – 5′ATAATGAATTTAGCGTGGGCAG3′
ERAP 3	5′CACCATGACCATGCCCAG3′ – 5′GGCTAGTGCAAGTCACAGA3′
ERAP 4	5′CCCCAAGTGCTGGGATTG3′ – 5′ACCAGAACAAGCTGGTCAC3′
ERAP 5	5′ATTGGCAACCATATCCCACA3′ – 5′CTGAGGCAACTACAGGGATG3′
ERAP 6-7	5′GCTGGTTCAAACACCTATCCA3′ – 5′ACTGTCAAGGAAGTTATCAGTG3′
ERAP 8	5′TAGGTTGCCCCCAAAATGTG3′ – 5′TAACCCCACAACCTTGCCTA3′
ERAP 9-10	5′CCTGCAGTTGCGTTTTCAAAG3′ – 5′TCAGAAAGTTTGGCACAGAGC3′
ERAP 11	5′GTAGCGGCTGTTTGAGCATA3′ – 5′CCATCCTACAAGGCAGATGT3′
ERAP 12	5′TGAGAGCTTGGCTGTAGCA3′ – 5′CATCCAAGATTGGGCAAGGA3′
ERAP 13	5′ATGGCCAGAGATGTAGTTGC3′ – 5′GATTCCTTCCTTCAAGTGCC3′
ERAP 14	5′TGAAGCAAAAAGTTGACTTCCC3′ – 5′GTACCATGAAGAGTCCCAGA3′
ERAP 15	5′GTCCTGAAGTCTTGTTGCAT3′ – 5′GCAGGGGAGACACTTAAACT3′
ERAP 16	5′ATGCTTCTCTGATCATGGGG3′ – 5′TGTGCTGACTTGCAGTATGT3′
ERAP 17	5′ACCTTTGTTACTGGCAGAC3′ – 5′CAGCAAGGCTAAAACACAGT3′
ERAP 18	5′CACACCGATCAGTTTTTAGC3′ – 5′AGGTTTTTACTGGCAGAGC3′
ERAP 19	5′GACAGCTGGGACTGCCTT3′ – 5′TTGGAGCCAAAACAGCCA3′
ERAP 20	5′GTCTGGGCTTTGTCATTTGT3′ – 5′CTGTGAGGCAAGGCTGTG3′

```
a)
Query  1      GATACTAGCATCAACACAATTTGTACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGA  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  18146  GATACTAGCATCAACACAATTTGAACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGA  18205

Query  61     TGAACCTGCCTCCAAAGCAAGTTTCTCAATCAAAATTAGAAGAGAGCCAAGGCACCTAGC  120
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  18206  TGAACCTGCCTTCAAAGCAAGTTTCTCAATCAAAATTAGAAGAGAGCCAAGGCACCTAGC  18265

b)
Query    2 ILASTQFVPTAARMAFPCFDEPASKASFSIKIRREPRHL 118
Sbjct  176 ILASTQFEPTAARMAFPCFDEPAFKASFSIKIRREPRHL 214
                 *                 *
```

Fig. 1. a) Nucleotide alignment between the queries (*de novo* SNVs-containing sequence) and the subject (*ERAP1* RefSeq) using BlastN. The polymorphic sites (g.18169A > T and g.18217 T > G) are highlighted in the box. b) Amino acid alignment between the queries (*de novo* SNVs-containing sequence) and the subject (ERAP1 protein RefSeq) using BlastX. The polymorphic sites (p.E183V and p.F199S) are highlighted in bold font and by the asterisks.

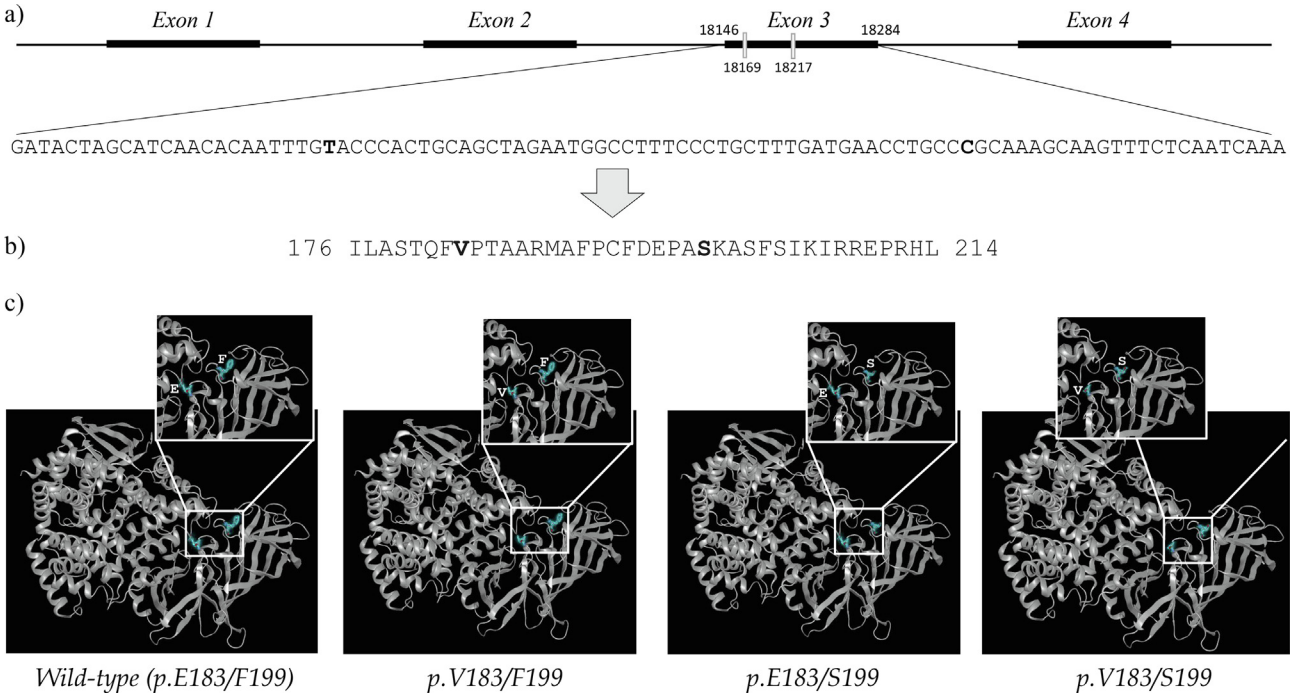

Fig. 2. a) Localization of our *de novo* SNVs at exon 3 *ERAP1* and nucleotide sequence showing the polymorphic sites in bold font. b) amino acid sequence showing both polymorphic sites (p.E183V and p.F199S) in bold font. c) 3D protein structure output by 3D Protean software for wild-type, p.E183Val, p.F199S, p.E183V plus p.F199S polymorphic sequences. The polymorphic residues are in blue colour.

The novel polymorphisms have localized within the *ERAP1* exon 3 (Fig. 2a) and have been submitted and released in GenBank database with MK252970 accession number. Interestingly, both SNVs were not found in none of 50 matched healthy controls.

The 3D protein structure prediction was obtained by protein sequence (Fig. 2b and c). The two amino acids were localized very close to each other in the 3D protein structure (Distance [Å]: 13.61). The computational analysis showed a change in energy score when the wild-type and the variant states were compared. The first variation was associated to a more stable protein chain (ΔE: −2.022), while the second polymorphism led to a less stable protein chain (ΔE: 2.967).

No correlation was found between the presence of novel SNVs and the most significant known polymorphisms (rs30187, rs17482078, rs27044) in our patients group. We found significantly higher frequencies of mutant rs30187GG (40.0% vs 24.0%), rs17482078AA (18.0% vs 0.0%) and rs27044GG (50.0% vs 30.0%) homozygous genotypes in patients compared with controls ($p < 0.05$).

4. Discussion

ERAP1 role is recently studied for its role in the immune response and several SNVs are reporting as BS risk loci, both known and novel variations [6,7,11–14,18]. *ERAP1* SNVs have been found to affect the protein trimming function and predispose to several human diseases, including autoimmune and rheumatic disorders, with partially unknown pattern. The variations led to altered antigen presentation and cytotoxic response influencing the antigenic peptide repertoire and the immunodominance hierarchy [8,14,19]. These effects are related to the critical role of the conformational dynamic of the enzyme changing from a “closed” conformation to an “open” conformational state. The first conformation could be triggered by the binding of a selective substrate with a structural change in the active site loop containing the conserved *exo*-peptidase GAMEN motif (317-321 amino acids) [8].

Our study was performed recruiting an explorative cohort of BS patients from Southern Italy, a population characterized by low disease prevalence. In fact, BS is a rare disease in Italy. We investigated the entire *ERAP1* sequence and identified two novel exon 3 polymorphisms (p.Glu183Val and p.Phe199Ser) that could contribute to increase the number of known SNVs. These SNVs were not previously reported in literature and in specific databases probably due to the significant variance in gene allelic frequencies among populations of different ethnic groups and different genetic ancestry.

Of note was the first variation, the substitution glutamate (hydrophilic amino acid) to valine (hydrophobic amino acid) in correspondence of a critical site (183 position). In fact, the glutamate at 183 position is a conserved site involved in the substrate binding, due to its significant role in the anchorage of the N-terminal amine group of the peptides.

For the second variation we also recognized the substitution of a phenylalanine (hydrophobic amino acid) to serine (hydrophilic amino

acid).

The different amino acid chemical properties could affect the protein three-dimensional shape and the protein interacting pattern. In order to address this point, the computational prediction of the 3D protein structure of wild type and SNVs-containing proteins was predicted recognizing a change in the energy and stability of the protein. The energy gap is a measure of how such gene variations could affect the protein stability and influence the folded/unfolded states [20]. Both the novel SNVs were predicted to be damaging and were characterized by the change of energy score: p.Glu183Val causes stabilization, while p.Phe199Ser induces destabilization. At the time we assume a link between our SNVs and the protein stability, that is critical for the enzyme conformational state change and activity, as well as for the substrate binding, to be validated in future proteomic studies.

In conclusion, we found two novel SNVs, probably two rare mutations, not previously reported in literature, but our study has some limitations. In fact, we genotyped a relatively small group from Southern Italy probably characterized by increased sensitivity to genetic drift and the hypothesis of ethnic-specific founder variations could be advanced. However, this hypothesis and the possible functional and pathogenic role of the *de novo* SNVs remains to be validated in a larger case-control study including a higher number of both patients and healthy controls.

Declaration of interest

None.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Maria Carmela Padula: Conceptualization, Data curation, Formal analysis, Investigation, Software, Writing - original draft. **Pietro Leccese:** Data curation, Investigation, Supervision, Validation, Writing - review & editing. **Nancy Lascaro:** Data curation, Investigation, Writing - review & editing. **Teresa Carbone:** Data curation, Visualization. **Michele Gilio:** Data curation, Visualization. **Angela Anna Padula:** Data curation, Investigation, Supervision, Validation, Writing - review & editing. **Giuseppe Martelli:** Conceptualization, Data curation, Investigation, Supervision, Validation, Writing - review & editing. **Salvatore D'Angelo:** Conceptualization, Supervision, Validation, Writing - review & editing.

Acknowledgment

Many thanks to Professor Ignazio Olivieri to have conveyed us the importance of honesty, humility and enthusiasm for the research and to have supported all the steps of the present investigation. Thanks to Dr. Vincenzo Trotta for his support in statistical analyses.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humimm.2019.02.003.

References

[1] G. Hatemi, E. Seyahi, I. Fresko, R. Talarico, V. Hamuryudan, One year in review 2016: Behçet's syndrome, Clin. Exp. Rheumatol. 34 (2016) 10.
[2] P. Leccese, Y. Yazici, I. Olivieri, Behcet's syndrome in non endemic regions, Curr. Opin. Rheumatol. 29 (2017) 12.
[3] M. Takeuchi, D.L. Kastner, E.F. Remmers, The immunogenetics of Behçet's disease: a comprehensive review, J Autoimmun. 64 (2015) 137.
[4] A. Gul, Genetics of Behçet's disease: lessons learned from genomewide association studies, Curr. Opin. Rheumatol. 26 (2014) 56.
[5] S. Ohno, T. Asanuma, S. Sugiura, A. Wakisaka, M. Aizawa, K. Itakura, HLA-Bw51 and Behçet's disease, JAMA 240 (1978) 529.
[6] A.L. Hanson, T. Cuddihy, K. Haynes, D. Loo, C.J. Morton, U. Oppermann, et al., Genetic variants in ERAP1 and ERAP2 associated with immune-mediated diseases influence protein expression and the isoform profile, Arthritis Rheumatol. 70 (2018) 255.
[7] E. Reeves, C.J. Edwards, T. Elliott, E. James, Naturally occurring ERAP1 haplotypes encode functionally distinct alleles with fine substrate specificity, J. Immunol. 191 (2013) 35.
[8] A. Papakyriakou, E. Stratikos, The role of conformational dynamics in antigen trimming by intracellular aminopeptidases, Front. Immunol. 8 (2017) 946.
[9] G. Kochan, T. Krojer, D. Harvey, R. Fischer, L. Chen, M. Vollmar, et al., Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming, Proc. Natl. Acad. Sci. U.S.A. 108 (2011) 7745.
[10] J.M. Xavier, F. Davatchi, O. Abade, F. Shahram, V. Francisco, B.S. Abdollahi, et al., Characterization of the major histocompatibility complex locus association with Behçet's disease in Iran, Arthritis Res. Ther. 17 (2015) 81.
[11] M. Takeuchi, M.J. Ombrello, Y. Kirino, B. Erer, I. Tugal-Tutkun, E. Seyahi, et al., A single endoplasmic reticulum aminopeptidase-1 protein allotype is a strong risk factor for Behçet's disease in HLA-B*51 carriers, Ann. Rheum. Dis. 75 (2016) 2208.
[12] M.J. Ombrello, D.L. Kastner, E.F. Remmers, Endoplasmic reticulum-associated amino-peptidase 1 and rheumatic disease: genetics, Curr. Opin. Rheumatol. 27 (2015) 349.
[13] Y. Kirino, G. Bertsias, Y. Ishigatsubo, N. Mizuki, I. Tugal-Tutkun, E. Seyahi, et al., Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1, Nat. Genet. 45 (2013) 202.
[14] M. Conde-Jaldon, M.A. Montes-Cano, J.R. Garcia-Lozano, L. Ortiz-Fernandez, N. Ortego-Centeno, R. Gonzalez-Leon, et al., Epistatic interaction of ERAP1 and HLA-B in behçet disease: a replication study in the spanish population, PLoS One 9 (2014) 1.
[15] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations, Nat. Methods 7 (2010) 248.
[16] N.L. Sim, P. Kumar, J. Hu, S. Henikoff, G. Schneider, P.C. Ng, SIFT web server: predicting effects of amino acid substitutions on proteins, Nucleic Acids Res. 40 (2012) W452.
[17] T.N. Plasterer, PROTEAN. Protein sequence analysis and prediction, Mol. Biotechnol. 16 (2000) 117.
[18] M.C. Padula, P. Leccese, A.A. Padula, S. D'Angelo, G. Martelli, ERAP1 molecular characterization: identification of a de novo allelic variant, HLA 92 (2018) 44.
[19] X. Wang, J. Ma, J. Ma, Y. Wen, L. Meng, H. Yang, et al., Bioinformatics analysis of genetic variants of endoplasmic reticulum aminopeptidase 1 in ankylosing spondylitis, Mol. Med. Rep. 16 (2017) 6532.
[20] L. Quan, Q. Lv, Y. Zhang, STRUM: structure-based prediction of protein stability changes upon single-point mutation, Bioinformatics 32 (2016) 2936.