Unsuspected Associations of Variants within the Genes NOTCH4 and STEAP2-AS1 Uncovered by a GWAS in Endemic Pemphigus Foliaceus

Augusto, Danillo G.; de Almeida, Rodrigo C.; Farias, Ticiana D. J.; Magalhaes, Wagner C. S.; Malheiros, Danielle; Lima-Costa, Maria Fernanda; Barreto, Mauricio L.; Horta, Bernardo L.; Kumar, Vinod; Wittig, Michael

Published in:
Journal of Investigative Dermatology

DOI:
10.1016/j.jid.2021.04.017

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

Publication date:
2021

Unsuspected Associations of Variants within the Genes *NOTCH4* and *STEAP2-AS1* Uncovered by a GWAS in Endemic Pemphigus Foliaceus

Journal of Investigative Dermatology (2021) **141,** 2741–2744; doi:10.1016/j.jid.2021.04.017

TO THE EDITOR

Pemphigus foliaceus is a blistering autoimmune disease of the skin representing a public health issue in Brazil, where it is endemic and neglected. Sporadic cases are reported across the globe. Nevertheless, an astonishing prevalence of more than 3% was reported for endemic pemphigus foliaceus (EPF) in some Brazilian regions (Schmidt et al., 2019), the highest ever reported for an autoimmune disease worldwide. Although the reasons for its endemicity are not clear, it has been suggested that environmental factors, such as agricultural activities, insect bites, and others, may trigger the disease in genetically susceptible individuals (Aoki et al., 2015; Qian et al., 2016).

We previously described strong associations with alleles of *HLA* class II genes (Pavoni et al., 2003; Petzl-Erler and Santamaria, 1989). Furthermore, through candidate gene association studies, we revealed polymorphisms of immune-related genes altering susceptibility to EPF (Petzl-Erler, 2020). Although the results based on candidate gene association studies established a strong role of genetic factors in EPF pathogenesis, this approach cannot reveal unsuspected susceptibility loci.

Here, we present a GWAS in EPF. This study was approved by the Human Research Ethics Committee of the Federal University of Paraná and the Brazilian National Human Research Ethics Committee (CONEP), protocol number CAAE 02727412.4.0000.0096, under the Brazilian Federal laws. All individuals gave written informed consent following the Declaration of Helsinki. The discovery cohort was composed of 234 patients approached in the endemic areas of Mato Grosso do Sul State, in a reference hospital specialized in EPF (Pemphigus Adventist Hospital, Campo Grande), and 5,658 controls that are part of the EPIGEN-Brasil initiative, which is based on three well-defined population-based cohorts from Brazilian regions (Kehdy et al., 2015; Magalhães et al., 2018). All patients enrolled in this study were diagnosed by experienced dermatologists on the basis of clinical features, histopathological features, immunofluorescence, and evaluation of antidesmoglein autoantibodies.

Individuals were genotyped for SNPs with the Illumina platform (Illumina, San Diego, CA) using microarray chips CoreExome-24 v1.1 for patients and HumanOmni2.5 for controls. Quality control was performed as described previously (Anderson et al., 2010; Kehdy et al., 2015). In summary, we eliminated related individuals and those with large-scale differences in ancestry (Supplementary Figure S1). Further, we excluded markers whose genotypes deviated from Hardy-Weinberg equilibrium ($P < 0.001$) and those with strong linkage disequilibrium ($r^2 > 0.8$). We also excluded markers with minor allele frequency <0.10 and call rate <96%. Principal component analysis was used to control for outliers and to merge datasets further. After quality control, 204,967 markers remained for the logistic regression analysis, assuming an additive model using four principal components as covariates to correct for possible population stratification. We observed minimal overall inflation of the genome-wide statistical results ($\lambda_{GC} = 1.05$; Supplementary Figure S1).

The strongest association signals were within the major histocompatibility complex, specifically within intergenic regions in the *HLA* class II region (Figure 1). In addition, we found a suggestive association ($P < 5 \times 10^{-5}$) with the intronic *rs2854050* in *NOTCH4*, a non-*HLA* gene within the major histocompatibility complex (Table 1). According to the Genotype Tissue Expression portal (https://gtexportal.org), this SNP has an expression quantitative trait loci effect on 13 different genes in 15 tissues (including skin and whole blood). More specifically, it is associated with lower *HLA-C* gene expression in sun-exposed skin and higher *HLA-DQA2* expression in whole blood ($P < 10^{-6}$). Variation in *NOTCH4* has been strongly associated with alopecia, also an autoimmune skin disease (Petukhova et al., 2010). Another non-*HLA* suggestive association was with *rs6968049,* located at the long non-coding RNA gene *STEAP2-AS1* on chromosome 7. Detailed annotation of all variants with the most significant *P*-values is given in Supplementary Tables S1 and S2. Results for all 204,967 variants are in Supplementary Table S3.

For replication and validation, we analyzed an independent cohort of 95 EPF and 153 pemphigus vulgaris (PV) patients, recruited at the University Hospital of the Ribeirão Preto Medical School of the University of São Paulo, Brazil. These cohorts were genotyped with the iPLEX MassARRAY System (Agena Bioscience, Inc., San Diego, CA) and compared with an independent subset of 1,000 controls from EPIGEN-Brasil. We replicated six of nine associations in the independent EPF cohort (Table 1 and Supplementary Figure S2). These results point to the importance of closely looking at suggestive associations because we replicated the associations of *NOTCH4* and *STEAP2-AS1* variants in the independent cohort.

Considering the similarities of pemphigus foliaceus and PV regarding their pathogenesis (Hammers and

Abbreviations: EPF, endemic pemphigus foliaceus; PV, pemphigus vulgaris

Accepted manuscript published online 13 May 2021; corrected proof published online 15 September 2021

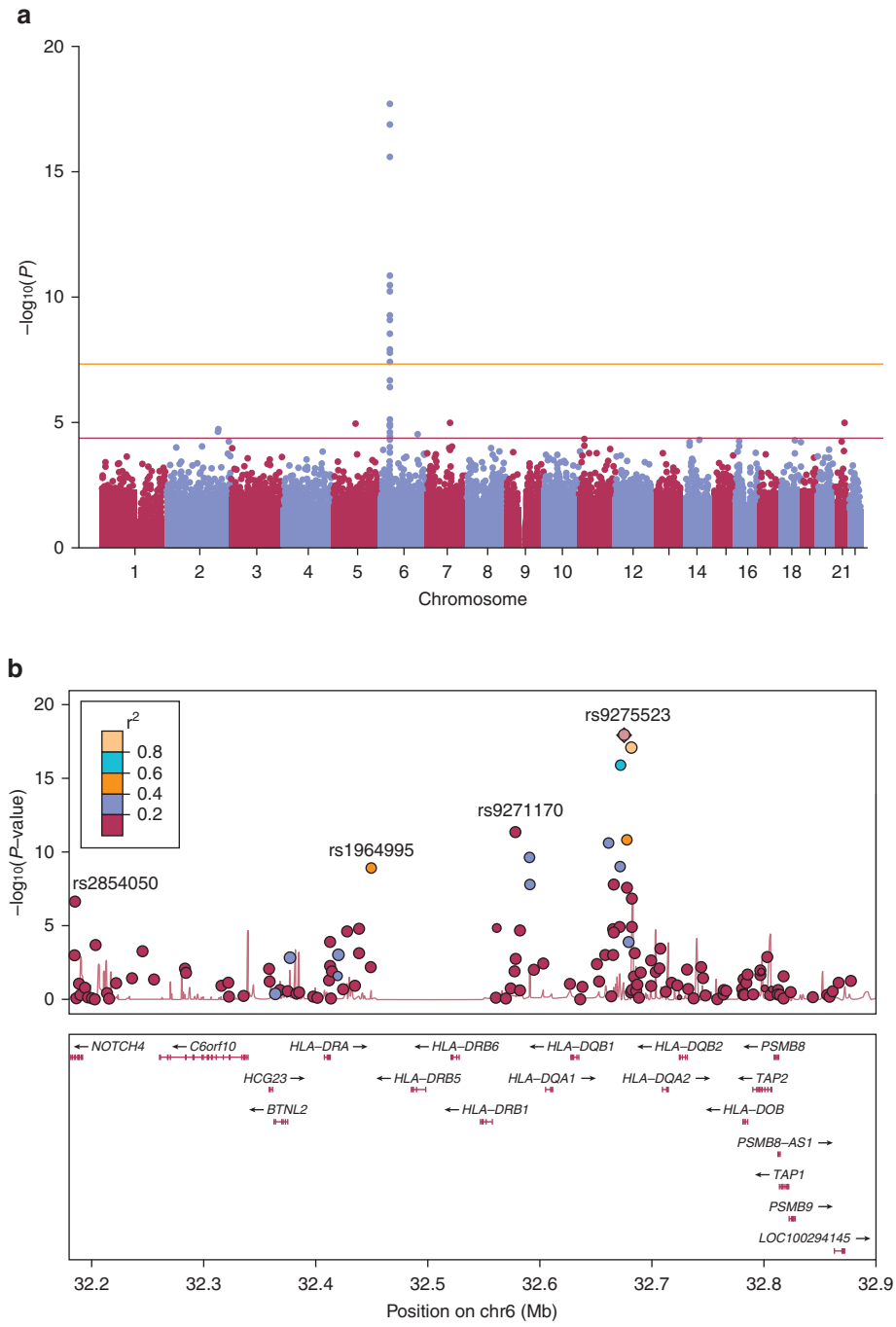

Figure 1. GWAS identifies variants associated with endemic pemphigus foliaceus. (**a**) Manhattan plot of the GWAS on endemic pemphigus foliaceus. The genome-wide significance level is set at 5×10^{-8}, and a threshold of 5×10^{-5} was set to indicate variants suggestively associated. Each dot represents an SNP. Y-axis shows the negative logarithm of the association *P*-value, and numbers in the X-axis indicate chromosomes. (**b**) Detailed visualization of variants within the HLA region. Each dot represents an SNP, color-coded according to linkage disequilibrium patterns with other variants. Y-axis shows the negative logarithm of the association *P*-value for the genes that are shown in the X-axis. Adapted from LocusZoom online tool (http://http://locuszoom.org/). chr, chromosome.

Stanley, 2020; Vodo et al., 2018), we checked if variants associated with EPF were also associated with PV (Table 1). We found three variants associated with EPF and PV. The variant *rs10947332*A* was associated with an increased risk of EPF and decreased risk of PV. We provide further insights into this result by analyzing all publicly available *HLA* and SNP genotyping data of 2,214 individuals of the 1000 Genomes Project populations (Abi-Rached et al., 2018). We found that *DRB1*01-DQB1*05:01* is the only HLA haplotype in strong linkage disequilibrium with *rs10947332*A* (D′ = 0.96), whereas *rs10947332*G* occurs in all other *DRB1-DQB1*

Table 1. Associations of Genetic Variants with Three Independent Pemphigus Cohorts

				Discovery - EPF 234 Patients and 5,658 Controls				Replication - EPF 95 Patients and 1,000 Controls				Validation – PV 153 Patients and 1,000 Controls			
Chr	Annotation	SNP	Allele	OR	L95	U95	*P*-Value	OR	L95	U95	*P*-Value	OR	L95	U95	*P*-Value
6	*HLA* class II[1]	rs9275523	*A*	5.26	3.63	7.62	1.82E−18	6.91	4.78	10.00	1.37E−26	4.20	3.28	5.39	1.14E−29
6	*HLA* class II[2]	rs9271170	*T*	3.22	2.29	4.52	1.42E−11								
6	*HLA* class II[1]	rs10947332	*A*	4.06	2.68	6.16	3.74E−11	3.35	2.39	4.69	3.27E−11	0.36	0.22	0.61	2.03E−05
6	*HLA* class II[2]	rs17533090	*T*	3.46	2.34	5.12	5.58E−10	2.42	1.69	3.46	4.32E−06	1.62	1.22	2.16	0.002
6	*HLA* class II[3]	rs1964995	*G*	2.72	1.97	3.75	9.18E−10								
6	*NOTCH4*[4]	rs2854050	*T*	4.21	2.41	7.35	4.08E−07	2.02	1.16	3.51	0.0183	0.92	0.54	1.57	0.894
5	Intergenic[5]	rs1032757	*T*	0.31	0.19	0.52	5.83E−06								
7	*STEAP2-AS1*[6]	rs6968049	*A*	2.02	1.48	2.77	1.13E−05	1.65	1.23	2.23	0.0010	0.98	0.77	1.26	0.899
21	*DSCAM* [4]	rs2837819	*G*	0.40	0.27	0.60	1.19E−05	0.66	0.42	1.03	0.072	0.70	0.50	1.00	0.052
2	Intergenic[7]	rs2357149	*T*	2.09	1.49	2.93	2.07E−05	1.47	1.03	2.10	0.041	1.02	0.77	1.36	0.886
2	Intergenic[7]	rs1601324	*G*	1.98	1.44	2.73	2.57E−05								
6	*IL20RA*[4]	rs1744061	*T*	0.47	0.33	0.67	3.28E−05								
11	*SERGEF*[4]	rs1548528	*A*	1.81	1.36	2.41	5.24E−05	1.18	0.85	1.64	0.316	0.82	0.64	1.04	0.110
18	Intergenic[8]	rs12962837	*A*	2.79	1.69	4.61	6.05E−05	1.08	0.70	1.66	0.736	0.98	0.71	1.35	0.935

Abbreviations: Chr, chromosome, EPF, endemic pemphigus foliaceus; L95, lower endpoint of the 95% confidence interval; PV, pemphigus vulgaris; U95, upper endpoint of the 95% confidence interval.
[1]Intergenic between *HLA-DQB1* and *HLA-DQA2*.
[2]Intergenic between *HLA-DRB1* and *HLA-DQA1*.
[3]Intergenic between *HLA-DRA* and *HLA-DRB5*.
[4]Intronic variant.
[5]Between the genes *ATP6AP1L* and *LOC105379050*.
[6]Long non-coding RNA gene.
[7]Between the genes *TMEFF2* and *PCGEM1*.
[8]Between the genes *DCC* and *MBD2*.

haplotypes. Haplotype *DRB1*01-DQB1*05:01* is strongly associated with increased risk of EPF and decreased risk of PV (Brochado et al., 2016; Gil et al., 2017; Petzl-Erler, 2020), and the linkage disequilibrium pattern explains the differential association of rs10947332 with these two diseases. None of the non-*HLA* variants associated with EPF were associated with PV, suggesting that their effect is restricted to EPF.

In summary, our results confirmed that *HLA* class II variants are the strongest genetic factors involved in EPF etiology. We also showed that EPF and PV share some intergenic susceptibility variants in the *HLA* class II region. The associations with intergenic variants may contribute to identifying causal variants; understanding the mechanism underpinning the associations between *HLA* class II genotypes and EPF; and exploring the differences and similarities of EPF, sporadic pemphigus foliaceus, and PV in the future. In addition, we identified at least two associations of EPF with possibly regulatory variants in non-*HLA* genes. Hopefully, this study will add a step ahead in comprehending this complex and unique disease that affects thousands of individuals worldwide.

Data availability statement

All data are available either in Supplementary Material or on request. Data from EPIGEN-Brasil are at https://www.ebi.ac.uk/ega/datasets/EGAD00010000787.

ORCIDs

Danillo G. Augusto: http://orcid.org/0000-0001-5665-8547
Rodrigo C. de Almeida: http://orcid.org/0000-0001-7966-056X
Ticiana D. J. Farias: http://orcid.org/0000-0002-3606-2428
Wagner C. S. Magalhães: http://orcid.org/0000-0003-3575-8068
Danielle Malheiros: http://orcid.org/0000-0002-2528-9707
Maria Fernanda Lima-Costa: http://orcid.org/0000-0002-3474-2980
Maurício L. Barreto: http://orcid.org/0000-0002-0215-4930
Bernardo L. Horta: http://orcid.org/0000-0001-9843-412X
Vinod Kumar: http://orcid.org/0000-0002-2775-6049
Michael Wittig: http://orcid.org/0000-0003-4240-4356
Andre Franke: http://orcid.org/0000-0003-1530-5811
Hauke Busch: http://orcid.org/0000-0003-4763-4521
Enno Schmidt: http://orcid.org/0000-0002-1206-8913
Ana Maria Roselino: http://orcid.org/0000-0002-2709-1825
Eduardo Tarazona-Santos: http://orcid.org/0000-0003-3508-3160
Angelica B. W. Boldt: http://orcid.org/0000-0002-0902-9622
Maria Luiza Petzl-Erler: http://orcid.org/0000-0002-0345-5276

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

We warmly thank all the individuals who voluntarily enrolled in this study. Special thanks to Hospital Adventista do Pênfigo for kindly opening their doors for our lab members and for treating patients with pemphigus with so much care and respect. We thank the staff of the Laboratório de Genética Molecular Humana, Universidade Federal do Paraná for their support. We also thank the statistician Chao Zhao for revision and advice with metanalysis. This work was supported by grants from the following funding agencies: Fundação Araucária (PRONEX FA/CNPq protocolo 50530 convenio 116/2018 and 9894.413.43926.1904/2013), Conselho Nacional

de Desenvolvimento Científico e Tecnológico - CNPq (470483/2014-8), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (400648/2014-8 and finance Code 001). DGA acknowledges funding from Young Talent Attraction, Science without Borders Program (CAPES 88881.067970/2014-01). ABWB acknowledges the Conselho Nacional de Desenvolvimento Científico e Tecnológico fellowship (314288/2018-0). TDJF received a scholarship under the International Sandwich Doctorate Program (Capes - PDSE 88881.132221/2016-01) and housing assistance from the German Academic Exchange Service, Deutscher Akademischer Austauschdienst - DAAD. The EPIGEN-Brasil project was funded by Departamento de Ciência, Tecnologia e Inovacão of the Brazilian Ministry of Health. HB, AF, and ES acknowledge funding by the Deutsche Forschungsgemeinschaft, German Research Foundation under Germany's Excellence Strategy — EXC 22167-390884018.

AUTHOR CONTRIBUTIONS

Conceptualization: DGA, RCDA, DM, ABWB, MLPE; Data Curation: RCDA, WCSM; Formal Analysis: RCDA, TDJF, WCSM, DGA; Investigation: DGA; Methodology: VK; Resources: DGA, MFLC, MLB, BLH, MW, AF, HB, ES, AMR, ETS, ABWB, MLPE; Supervision: MLPE; Writing - Original Draft Preparation: DGA, MLPE, TDJF; Writing - Review and Editing: DGA, RCDA, TDJF, WCSM, DM, MFLC, MLB, BLH, VK, MW, AF, HB, ES, AMR, ETS, ABWB, MLPE

***Danillo G. Augusto*[1,2], *Rodrigo C. de Almeida*[1,3], *Ticiana D.J. Farias*[1], *Wagner C.S. Magalhães*[4,5], *Danielle Malheiros*[1], *Maria Fernanda Lima-Costa*[6], *Maurício L. Barreto*[7,16], *Bernardo L. Horta*[8], *Vinod Kumar*[9,10,11,15], *Michael Wittig*[12], *Andre Franke*[12], *Hauke Busch*[13], *Enno Schmidt*[13], *Ana Maria Roselino*[14], *Eduardo Tarazona-Santos*[5], *Angelica B.W. Boldt*[1] and *Maria Luiza Petzl-Erler*[1,*]**

[1]Programa de Pós-Graduação em Genética, Universidade Federal do Paraná, Curitiba, Brazil; [2]Department of Neurology, University of California, San Francisco, San Francisco, California, USA; [3]Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; [4]Núcleo de Ensino e Pesquisa, Instituto Mário Penna, Belo Horizonte, Brazil; [5]Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; [6]Instituto de Pesquisa Rene Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Brazil; [7]Instituto de Saúde Coletiva, Universidade Federal da Bahia, Salvador, Brazil; [8]Programa de Pós-Graduação em Epidemiologia, Universidade Federal de Pelotas, Pelotas, Brazil; [9]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; [10]Department of Internal Medicine, Radboud University Medical Center, Nijmegen, The Netherlands; [11]Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, The Netherlands; [12]Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany; [13]Lübeck Institute of Experimental Dermatology, University of Lübeck, Lübeck, Germany; [14]Division of Dermatology, Department of Clinical Medicine, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil; [15]Nitte (Deemed to be University), Division of Infectious Diseases, Nitte University Centre for Science Education and Research (NUCSER), Paneer Campus, Deralakatte, Mangaluru, India; and [16]Fundação Instituto Oswaldo Cruz, Centro de Integração de Dados e Conhecimentos para Saúde, Salvador, Brazil

[*]Corresponding author e-mail: perler@ufpr.br

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2021.04.017

REFERENCES

Abi-Rached L, Gouret P, Yeh JH, Di Cristofaro J, Pontarotti P, Picard C, et al. Immune diversity sheds light on missing variation in worldwide genetic diversity panels. PLoS One 2018;13:e0206512.

Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc 2010;5:1564–73.

Aoki V, Rivitti EA, Diaz LA. Cooperative Group on Fogo Selvagem Research. Update on fogo selvagem, an endemic form of pemphigus foliaceus. J Dermatol 2015;42:18–26.

Brochado MJ, Nascimento DF, Campos W, Deghaide NH, Donadi EA, Roselino AM. Differential HLA class I and class II associations in pemphigus foliaceus and pemphigus vulgaris patients from a prevalent Southeastern Brazilian region. J Autoimmun 2016;72:19–24.

Gil JM, Weber R, Rosales CB, Rodrigues H, Sennes LU, Kalil J, et al. Study of the association between human leukocyte antigens (HLA) and pemphigus vulgaris in Brazilian patients. Int J Dermatol 2017;56:557–62.

Hammers CM, Stanley JR. Recent advances in understanding pemphigus and bullous pemphigoid. J Invest Dermatol 2020;140:733–41.

Kehdy FSG, Gouveia MH, Machado M, Magalhães WCS, Horimoto AR, Horta BL, et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. Proc Natl Acad Sci USA 2015;112:8696–701.

Magalhães WCS, Araujo NM, Leal TP, Araujo GS, Viriato PJS, Kehdy FS, et al. EPIGEN-Brazil initiative resources: a Latin American imputation panel and the scientific workflow. Genome Res 2018;28:1090–5.

Pavoni DP, Roxo VMMS, Marquart Filho A, Petzl-Erler ML. Dissecting the associations of endemic pemphigus foliaceus (fogo selvagem) with HLA-DRB1 alleles and genotypes. Genes Immun 2003;4:110–6.

Petukhova L, Duvic M, Hordinsky M, Norris D, Price V, Shimomura Y, et al. Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. Nature 2010;466:113–7.

Petzl-Erler ML. Beyond the HLA polymorphism: a complex pattern of genetic susceptibility to pemphigus. Genet Mol Biol 2020;43:e20190369.

Petzl-Erler ML, Santamaria J. Are HLA class II genes controlling susceptibility and resistance to Brazilian pemphigus foliaceus (fogo selvagem)? Tissue Antigens 1989;33:408–14.

Qian Y, Culton DA, Jeong JS, Trupiano N, Valenzuela JG, Diaz LA. Non-infectious environmental antigens as a trigger for the initiation of an autoimmune skin disease. Autoimmun Rev 2016;15:923–30.

Schmidt E, Kasperkiewicz M, Joly P. Pemphigus. Lancet 2019;394:882–94.

Vodo D, Sarig O, Sprecher E. The genetics of pemphigus vulgaris. Front Med (Lausanne) 2018;5:226.