Investigation of transcriptionally regulated differential gene expression as a potential protection signature in the Pemphigus vulgaris disease model.

Tyler M. Chinsky, Ricki Chen, Sarah Holan, Harshitha B Suresha Kristina Seiffert-Sinha and Animesh A Sinha

Department of Dermatology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY

University at Buffalo
The State University of New York

ABSTRACT

- Pemphigus vulgaris (PV) is a rare IgG-mediated autoimmune disorder that serves as an excellent model to investigate organ specific autoimmunity in humans.
- Clinically, PV presents as painful blistering of the mucosa and skin. The pathogenesis of PV stems from the production of circulating autoantibodies directed primarily against desmoglein 3 and 1, key cadherins that serve a critical role in the adhesion of epithelial cells.
- This destruction frequently leads to patient outcomes that reflect heightened risks of infection and mortality. Pathogenicity in Pemphigus vulgaris (PV) has long been associated with certain HLA alleles, specifically HLA DRB1*0402 and DQB1*0503, that are a requirement for the initiation of disease.
- The majority of individuals that express the pemphigus associated HLA alleles do not develop disease and despite carrying genetic risk, they remain healthy. Therefore, we propose an entirely novel hypothesis that PV patients develop disease at least partially due to an inability to mitigate key immune pathways activated by the PV associated HLA alleles that drive autoimmunity, while healthy individuals who carry the PV associated HLA risk alleles (HLA-matched controls) remain healthy by down regulating a set of key genes (the "protection signature") that thwart the activation of potential autoimmune pathways due to the PV associated HLA risk alleles they carry.
- Our hypothesis is supported by genome-wide transcriptional data that reveal the existence of two distinct expression signatures:
  - (1) An "HLA associated disease signature" encompassing drivers of disease in PV patients with active disease
  - (2) A "protection signature" comprising genes exhibiting a strong down regulation found only in HLA matched controls, but not in HLA-unmatched controls who do not carry the PV associated HLA risk alleles, and thus have no need to block potential HLA linked autoimmune activation. This data support the notion of a highly regulated and counterbalanced network of genes and pathways that underpin the delicate balance between health and disease, knowledge that could be leveraged to envision wholly new strategies to prevent the development and/or progression of autoimmunity.

OBJECTIVES

- ❑ To evaluate the importance of HLA association in PV patients and controls and establish a PV-associated "HLA disease" signature and the "protection" signature.
- ❑ To identify differentially expressed genes (DEGs) associated with the "HLA disease" signature and the "protection" signature.
- ❑ To investigate functional pathways and genes in the "HLA disease" signature and the "protection" signature relevant to understanding underlying disease mechanisms and pathogenesis as compared with those linked to the protection signature.

METHODS

- PBMCs from the following samples were used for microarray analysis:[4]
  21 PV patients : 13 active (PV-A)
  8 remittent (PV-R)
  10 healthy controls : 4 matched for DRB1*0402^{+} and/or DQB1*0503^{+} (MCR)
  6 unmatched for DRB1*0402^{-} and/or DQBI*0503^{-} (UMCR)
- Microarray analyses were run using Affmetrix (HG)-U133 Plus 2.0 microarray chips (Waltham, MA) on the 31 patients and controls.
- Raw data was processed, .CEL files created, and normalized using RMA in Partek Genomic Suite v 7.19.1125 (St. Louis, MO).
- Normalized data was used to make Principal Component Analysis (PCA) graphs, along with an unsupervised clustering of all normalized samples.
- Of the 31 samples, all PV-R and two PV-A (PV198 and PV187) were removed to show differences between controls and active patients.
- Differentially expressed genes (DEGs) were detected in Partek Genomic Suite v 7.19.1125 by designing two analysis of variance (ANOVA) models showing the correlation between disease state and HLA status:
  1. "HLA Associated Disease Signature": PV-A samples vs. UMCR samples, then overlaying the corresponding MCR expression levels
  2. "Protection Signature": MCR samples vs. PV-A + UMCR samples
- DEGs were filtered in the disease and protection signatures using a log2 fold change cutoff of ≥ ±1.2 and 2.5, respectively, in addition to an FDR and p-value ≤ 0.05.
- Gene Ontology analysis was completed for biological processes using Partek Genomic Suite v 7.19.1125, filtered by fisher exact test and enrichment p-value ≤ 0.05, and forest plots created showing the number of up- vs down-regulated genes enriched in each process.
- KEGG pathway analysis was completed on the filtered biological processes and DEGs using Partek Pathway Tool v 7.19.1125 (St. Louis, MO) and filtered by both Fisher's Exact Test and enrichment p-value ≤ 0.05.

REFERENCES

1. Sinha A.A. et al., (1988) *Science*, 239 (4843):1026-9.
2. Lee E., et al., (2006) *Human Immunology*, 67 (1-2): 125-39
3. Tong et al., (2006) BMC Bioinformatics 18 (7): Suppl 5:S7
4. Dey-Rao R., Seiffert-Sinha K., Sinha A.A., (2013) *Genes and Immunity*, 14(8): 487-99.

RESULTS AND DISCUSSION

1. Unsupervised Hierarchical Clustering

HLA-unmatched controls (UMCR; do not carry the PV associated HLA risk alleles) were added to the mixture of all PV patients (PV) and HLA matched control samples (MCR) and all samples were sorted by unsupervised cluster analysis.

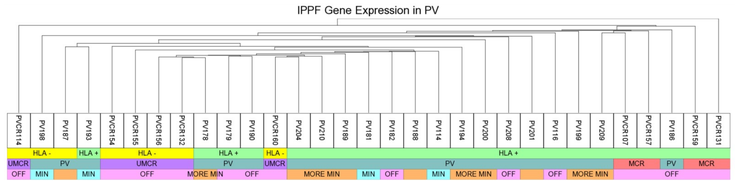


Figure 1: Unsupervised hierarchical clustering of matched control, unmatched control, and PV samples

- Thus far, there is no trend visible based on therapy or disease status, however there are potential trends based off HLA status separating HLA- patients [PVCR114, PV198, PV187, PVCR154, PVCR156, PVCR132, PVCR160] and most HLA+ samples [exceptions: PV193, PV178, PV179, and PV190]

➢ Overall, we note that there is an incomplete separation of patients and controls by disease. We observe a further significant parsing of samples based on HLA status; a theme repeated by PCA.

❑ These data highlight HLA as a potential key driver of gene expression, overriding even disease status, suggesting the existence of a HLA associated disease signature.

2. Principal Components Analysis

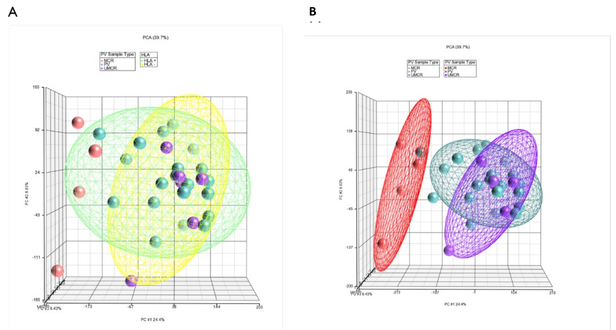


Figure 2. Principal component analyses of matched controls, unmatched controls, and PV active samples, and HLA positive and negative samples

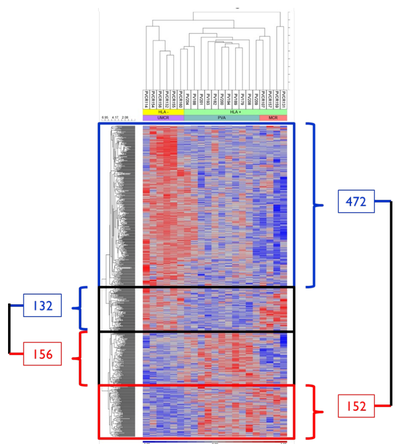
➢ HLA status: HLA-matched MCR group with HLA-matched PV samples (green ellipsoid), while the HLA-unmatched PV samples group with the HLA-unmatched controls (UMCR)(yellow ellipsoid) (Figure 2A).

➢ Disease status: HLA-matched MCR clusters away from all PV and HLA-unmatched control samples (UMCR), a separation that is most likely driven by the previously described subset of 155 DEGs demonstrating opposite regulation in MCR from PV and UMCR, forming the "protection" signature [2] (Figure 2B).

3. HLA Associated Disease Signature

Unsupervised clustering of the 21 patient and control samples revealed complete separation based on HLA status further showing the association of HLA alleles DRB1*0402 and DQB1*0503 to the progression and inhibition of disease.

Figure 3: HLA Associated Disease Signature



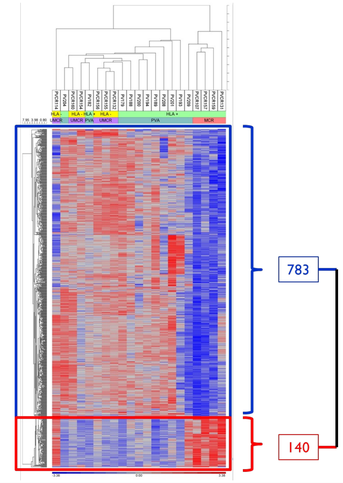
➢ Revealed 912 DEGs which separated into two categories:

1. **HLA driven genes:**
   - Sets of genes regulated due to HLA
   - Observed where PV-A and MCR have similar expression levels
   - Contained 472 downregulating and 152 upregulating genes
2. **HLA independent genes:**
   - Sets of genes regulated due to disease
   - Showed inverse expression levels in PV-A compared to MCR & UMCR, regardless of HLA status
   - Contained 132 downregulating and 156 upregulating

4. Protection Signature

We have previously hypothesized that the protection signature to block progression to disease in genetically susceptible, but healthy individuals may harbor targets that could potentially be manipulated to thwart, or even reverse, disease.[4]

Figure 4. Protection Signature



➢ Rendered 923 DEGs showing near complete separation based on HLA status, with the exceptions of PV204 and PV182

➢ A clear inverse relation can be seen based on MCR expression levels compared with PV-A and UMCR

➢ This suppression of critical pathways in genetically susceptible individuals is required to halt the expression of autoimmune phenotypes

➢ These protecting genes in MCR patients include 783 downregulating and 140 upregulating genes

5. Gene Ontology and KEGG Pathway Analyses

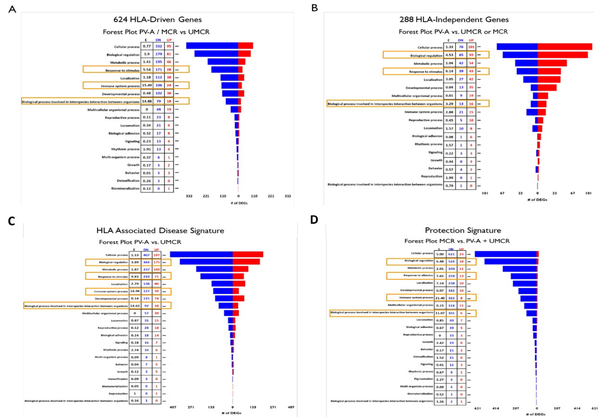


Figure 5: GO analysis forest plots of both HLA associated disease and protection signatures

➢ GO analysis of both HLA independent and driven gene sets showed a large difference in processes enriched (boxed in orange): immune system processes for HLA driven and interaction and regulation in HLA independent (Figures 5A, 5B)

➢ While when comparing the HLA associated disease signature overall with the protection signature reveals the same processes enriched (boxed in orange): Immune function, cellular regulation and stimulus, organismal interactions (Figures 5C, 5D)

KEGG enriched pathways of both HLA associated disease and protection signatures

➢ KEGG pathway analysis revealed 21 pathways significantly mapped to 17 to genes within the HLA associated disease signature, and 20 pathways significantly mapped to 28 genes within the protection signature

❑ There are unique pathways between HLA driven and independent genes, indicating that these pathways are needed for disease expression in genetically susceptible individuals.

❑ We also observe that healthy individuals carrying PV-associated HLA alleles (MCRs) may be "protected" from progression to disease by the down-/up-regulation of genes/pathways (in blue) that are found to be activated in the "disease" signature.

HLA Associated Disease Signature		Protection Signature
HLA Independent	HLA Driven	
Endocytosis	Kaposi Sarcoma Associated Herpesvirus Infection	Chemokine signaling
Yersinia infection	Glycerophospholipid metabolism	Sphingolipid metabolism
Shigellosis	Cytokine-Cytokine receptor interaction	FoxO signaling
IL-17 signaling	Epstein Barr	Shigellosis
Salmonella infection	Hepatitis C	Yersinia infection
Hepatitis C	Osteoclast Differentiation	Pathogenic Escherichia coli infection
Tight Junction	Influenza A	Regulation of actin cytoskeleton
	TLR Signaling	Endocytosis
	Hepatitis B	C-type lectin receptor signaling
	Natural Killer Cell Mediated Cytotoxicity	Tuberculosis
	Measles	Renin secretion
	MAPK Signaling pathway	Neutrophil extracellular trap formation
	Th17 Cell-Mediated Differentiation	Bacterial invasion of epithelial cells
	HLA Driven	Salmonella infection
	Kaposi Sarcoma Associated Herpesvirus Infection	Malaria
	Glycerophospholipid metabolism	Phagosome
	Cytokine-Cytokine receptor interaction	Rheumatoid arthritis

CONCLUSIONS + NOVEL HYPOTHESIS

- ❑ We conclude that the HLA matched healthy individuals (MCRs) and PV patients that have the same immunogenic background, share many of the same dysregulated disease-specific pathways and processes.
- ❑ Further investigation, via functional validation and NGS analyses, into these genes is necessary to validate these signatures.
- ❑ We hypothesize that *HLA status is central to both autoimmune activation in PV patients and protection in HLA-matched controls.*