Published in final edited form as:

Pediatr Emerg Care. 2015 July ; 31(7): 465–472. doi:10.1097/PEC.0000000000000483.

RNA Biosignatures in Adolescent Patients in a Pediatric Emergency Department with Pelvic Inflammatory Disease

Fran Balamuth, MD PhD^{1,2}, Zhe Zhang, PhD³, Eric Rappaport, PhD⁴, Katie Hayes, MS⁵, Cynthia Mollen, MD MSCE^{1,2}, and Kathleen E. Sullivan, MD, PhD^{2,6}

¹Division of Emergency Medicine, The Children's Hospital of Philadelphia

²University of Pennsylvania Perelman School of Medicine

³Bioinformatics Core Facility, The Children's Hospital of Philadelphia

⁴Microarray Core Facility, The Children's Hospital of Philadelphia

⁵Division of Emergency Medicine, Children's National Medical Center

⁶Division of Allergy and Immunology, The Children's Hospital of Philadelphia

Keywords

Biochemical marker; Adolescent health; Sexually transmitted infection

Introduction

Pelvic inflammatory disease (PID) consists of infection and inflammation of the upper female genital tract. PID typically begins in the lower genital tract in the vagina and cervix and then spreads to the upper genital tract including uterus, fallopian tubes, ovaries, and peritoneum¹⁻³. PID is thought to be largely a polymicrobial infection, but several sexually transmitted organisms have been implicated including *Neisseria gonorrhoeae* and *Chlamydia trachomatis*. It has been suggested that other organisms may also be involved in the development of PID including *Trichomonas vaginalis*⁴ and *Mycoplasma genitalium*^{5,6}. If untreated, PID can lead to long term complications in young women including ectopic pregnancy, chronic pelvic pain, and infertility¹.

Because the consequences of under-recognition and under-treatment of this condition are severe, the Centers for Disease Control and Prevention (CDC) have maintained broad diagnostic criteria for PID to ensure that cases are not missed⁷. The current minimum diagnostic criteria for PID in a sexually active female include the presence of lower abdominal pain and any of the following: cervical motion tenderness, adnexal tenderness, or uterine tenderness⁷.

Corresponding Author: Fran Balamuth MD, PhD Division of Emergency Medicine Children's Hospital of Philadelphia 3501 Civic Center Blvd, CTRB 9206 Philadelphia, PA 19014 Phone: 215 590 7295 balamuthf@email.chop.edu.

Financial Disclosure: The authors have no financial relationships relevant to this article to disclose.

Conflict of Interest: The authors have no conflicts of interest to disclose.

This clinical diagnosis can be particularly difficult in the adolescent population, where discomfort with pelvic examination is common and pain can be difficult to interpret⁸⁻¹⁰. Although PID may be challenging to diagnose in the adolescent female population, these patients represent a high risk group: PID has been shown to be more common in adolescent females, particularly those with early sexual intercourse and lower socioeconomic status^{9, 11}. The ability to diagnose PID is critical in the pediatric emergency department, which has been shown to be an important source of accessing health care for high risk adolescents, and where lower abdominal pain is a common chief complaint¹².

In addition to the long term complications of a missed PID diagnosis mentioned above, there are also potential complications with over-diagnosis such as inappropriate antibiotic use and unnecessary health care resource utilization. Because of the dangers involved with both over and under-diagnosis of PID in the adolescent emergency department population, the development of novel methods to improve the accuracy of diagnosis of PID is essential. Over the past decade, the development of RNA based microarray technology has allowed clinicians to contemplate new ways to approach such diagnostic dilemmas. There are preliminary data in several areas including acute infection, inflammatory bowel disease, and childhood cancer which suggest that broad evaluation of the human immune response to disease as measured by RNA expression patterns may correctly identify distinct patient sets¹³⁻¹⁷. The array technology collapses a set of signals into a metric for which accurate quantitation and analytic methods have been validated¹⁸. Signals that may have been modest individually can be amalgamated into a measurable signature and thus may have superior function diagnostically. It has been established that unique patterns of inflammatory gene expression can be identified in pathologic specimens from different abdominal processes including appendicitis and inflammatory bowel disease^{19, 20}. Furthermore, it has been suggested that proteomic based strategies might be effective ways of identifying surgical causes of abdominal pain in adolescent females²¹.

The peripheral immune response to organisms associated with PID is much less well studied. However, there are data in animal models which suggest that unique T cell populations are involved in the immune response to *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infection^{22, 23}. More recently, it has been shown that differential peripheral immune responses to *Chlamydia trachomatis* may predict upper vs. lower genital tract disease²⁴ and also that endometrial inflammation by T, and B cells is associated with *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, and *Trichomonas vaginalis* infection⁴. Further, genetic polymorphisms in Toll-like receptors have been suggested to play a role in *Chlamydia trachomatis* infection and PID²⁵. These data suggest that a detailed analysis of gene expression in peripheral leukocytes of adolescent girls with PID may yield important information about the inflammatory response in these patients, and may ultimately lead to the development of novel diagnostic tools. In this study, we compared RNA biosignatures in adolescent female patients with PID to those in healthy control patients.

Materials and Methods

Study Design/Site

This cross sectional study was conducted in the emergency department (ED) and operating room (OR) of a quaternary care children's hospital with more than 90,000 annual visits. The study was approved by the institutional review board at the study site. Study enrollment occurred between October 1 2009 and June 30 2012.

Subject Selection

Adolescent females ages 12-19 years (inclusive) who presented to the ED with a chief complaint of abdominal pain or genitourinary complaint (dysuria, vaginal discharge, flank pain) and who were diagnosed with PID by an attending physician were approached for enrollment. Diagnosis of PID was based on the CDC criteria, and included the presence of abdominal pain with cervical motion tenderness or adnexal tenderness or uterine tenderness as documented by an attending physician or fellow ⁷. Control samples were collected from adolescent females ages 12-19 (inclusive) in the operating room who were undergoing an elective surgical procedure for a non-abdominal problem. Patients with a history of autoimmune disease or immunodeficiency disorder were excluded. Patients were also excluded if they had been diagnosed with a sexually transmitted infection in the past 60 days, or if they had been treated with antibiotics in the past 30 days. Once written consent was obtained from eligible subjects or their parents/guardians, a 3 ml blood sample was collected and stored. Because the diagnosis of PID constitutes a sexually transmitted infection, PID patients were permitted to consent for the study despite their age. For control patients, consent was obtained from parents/guardians and assent was obtained from the patient.

Laboratory Testing and Diagnostic Imaging

Testing for sexually transmitted infection was performed on patients diagnosed with PID at the discretion of the treating physician. *Chlamydia Trachomatis* (CT) and *Neisseria gonorrhoeae* (GC) testing was via urine APTIMA Combo 2 Assay and *Trichomonas vaginalis* (TV) testing used the vaginal OSOM Trichomonas rapid test. Further laboratory testing such as complete blood count (CBC) and inflammatory marker testing including c-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) were performed at the discretion of the treating physician. Where available, results were abstracted from the medical record in retrospective fashion. Descriptive statistics were performed using STATA 11.0 (www.stata.com, College Station, TX).

Array Processing

Samples were collected in Tempus RNA collection tubes²⁶ and stored at -80°C pending analysis. RNA was isolated in two mixed batches containing PID patients and controls using Tempus Spin RNA Isolation kit at the Clinical Trial Resource Center at the Children's Hospital of Philadelphia using standard manufacturer's instructions. RNA quality was confirmed by Agilent Bioanalyzer assay. The arrays were processed in a dedicated core facility using the Affymetrix system (www.affymetrix.com).

Data Analysis

Data analysis was performed in collaboration with the authors' institutional bioinformatics core facility. RNA degradation was assessed by determining the average measurements of all genes at the 5' end-probe, 3' end-probe, and all probes in the middle. All samples met predefined quality standards.

To determine if given genes were present or absent in each sample, probes matched to a given gene were compared to all of the probes not mapped to any known genes. Those control probes were a mixture of mismatch probes, negative controls, spike-in controls, and probes targeted to unknown transcripts. We compared the probes matched to each known gene to all control probes via Student's t test and calculated a p value for each gene in the sample. A gene was called present if the p value is ≤ 5 , marginal if $0.05 > p > 0.1$, and absent if $p > 0.1$.

Affymetrix probes were grouped into unique Entrez gene IDs using custom library file downloaded from the BRAINARRAY database (<http://brainarray.mbni.med.umich.edu>). The raw data in .CEL files were normalized and summarized by the RMA (Robust Multichip Averaging) method to generate an N*M matrix, where N is the number of unique Entrez genes and M is the number of samples. The normalized data were log2-transformed for statistical analysis.

Data from 9 patients and 9 controls was used to identify genes that were differentially expressed between the two groups. Unpaired comparisons were performed by the SAM (Significance Analysis of Microarrays) method on all measured genes. The differential expression of each gene between groups was represented by two indices: the magnitude and the statistical significance of the difference. The magnitude was represented by fold change and the difference of group averages (the same as log2-ratio of un-logged data). The significance was represented by p value and FDR (false discovery rate).

The selection of significant genes included two steps. First, genes with p values more than 0.05 were excluded; second, the remaining genes were sorted by their fold changes and the genes with the highest and lowest changes were selected into two gene lists. Both lists were inputted into DAVID (<http://david.abcc.ncifcrf.gov>) for functional classification. Nearest neighbor classification was performed on the training set, and 17 additional patients were classified using this method. Ingenuity software was used to perform network analysis.

All statistical procedures, including SAM analysis, hierarchical clustering, and principal components analysis, were performed within the R statistical environment (<http://www.r-project.org>)

Results

Patient characteristics

Fifty patients with PID and 9 control OR patients were enrolled during the study period. PID patients were slightly older than control patients (16.3 vs 15.2 years old, $p=0.006$). Of the PID patients, 90% were African American, 6% were Caucasian, and 4% were Hispanic. Of

the control group, 45% were African American, 35% were Caucasian patients, and 9% Hispanic.

Identification of sexually transmitted and urinary pathogens and laboratory testing

Of the patients with clinical PID, 9 had positive urine testing for *Chlamydia trachomatis* or *Neisseria gonorrhoeae*. Of these, 2 had *Chlamydia trachomatis* only, 2 had *Neisseria gonorrhoeae* only, and 5 had infection with more than one organism. 5 patients had positive testing for *Trichomonas vaginalis*, and 36 patients did not have positive testing for any of these organisms. 47 patients had urine cultures sent, 2 had urinary tract infections identified. One of these patients had co-infection with *Trichomonas vaginalis* and the other had co-infection with *Neisseria gonorrhoeae*.

We compared white blood cell count, absolute neutrophil count, C-reactive protein, and erythrocyte sedimentation rate values among patients undergoing RNA expression analysis between patients with positive and negative testing for STI. There was no significant difference in mean values between the two patient groups (Table 1). Three patients were not evaluable as they did not have laboratory testing sent.

RNA expression analysis of training set

We defined a training set of PID patients as those with either *Chlamydia trachomatis* or *Neisseria gonorrhoeae* infection and compared them to control patients. A two group comparison using the SAM method identified 170 differentially expressed genes (DEGs) with p values less than 0.001, corresponding to a false discovery rate of 0.04. A heatmap including these DEGs in all 18 training subjects is shown in figure 1. Noticeably, almost 80% (135) of the DEGs were down-regulated in PID patients, where the average expression level was reduced by 14-63%. The 35 up-regulated DEGs had increases of 12-251% in PID patients compared to controls. Seven of the up-regulated DEGs are small nucleolar RNAs, suggesting a potential role for non-coding RNA in gene regulation in PID.

Network analysis of differentially expressed genes

We went on to perform networking analyses using Ingenuity (www.ingenuity.com) to identify dysregulated pathways in the training set. This analysis is shown in Figure 2. Many of the genes identified in this analysis are involved with the immune response. For example, several genes involved in cytokine signaling such as SOCS3 (suppressor of cytokines 3), CSF2RB (Common subunit to Type I cytokine receptors), and JAK3 (Janus associated kinase 3) were identified. Other genes of interest include PI3K (Phosphatidylinositol 3 kinase) and PAG1 (Phosphoprotein associated with glycolipid enriched microdomains), which are both involved with immune cell signaling. Several transcription factors including RELB, BAZ2A, and MLL4 were also identified. Interestingly, SLC26A3 was also identified in our analysis as a differentially expressed gene. This gene is involved in chloride anion exchange at mucosal surfaces.

Characterization of unknown patients

We then performed principal component analysis using the 170 DEGs identified in the training set on 15 additional patients with clinical PID and 2 controls. Five PID patients in

this set had positive testing for *Trichomonas vaginalis*, and all of these patients clustered with the PID patients in the training set. Both controls (501,502) clustered with the control patients. The remainder of the 10 patients in this set had clinical findings consistent with PID but did not test positive for any of the identified STIs. Of these, 5 clustered with the PID patients and 5 clustered with the controls. The heatmap from this analysis is shown in Figure 3.

In addition, we performed unsupervised clustering of the entire dataset and displayed the results as two principal components. These data are shown in figure 4. As would be expected, for this clinical population, the results were diverse with some subjects partitioning with the known PID patients and some partitioning with the controls.

Contribution to gene expression variability from known clinical and laboratory values

Figure 4 shows that principal component 1 accounts for over 60% of the total variance of gene expression data. We then investigated the correlation between this composite signature variable with known clinical and laboratory values such as duration of symptoms, white blood cell count, absolute neutrophil count (ANC), C-reactive protein, and erythrocyte sedimentation rate. A significant correlation would indicate a possible relationship of the gene expression pattern with known biomarkers. Figure 5 shows that among the eight measured biomarkers and clinical features, only the ANC correlated with principal component I with marginal significance, suggesting that gene expression data provides additional predictive value beyond known indicators.

Discussion

This study demonstrated the feasibility of collecting and analyzing RNA samples from adolescent girls with PID in a pediatric emergency department setting, and offers preliminary data for the development of a novel diagnostic strategy in a high risk population. We were able to identify 170 genes that were differentially expressed in PID patients compared to controls, many of which are involved in the inflammatory process. This adds to the growing medical literature indicating that RNA expression patterns may be of diagnostic use in pediatric infections. As we hypothesized, the patients with clinical PID and negative STI testing partitioned equally into the PID and control sets when analyzed by PCA analysis. It is known that not all adolescents with PID symptoms have true upper genital tract infection²⁷. Intriguingly, all patients with *Trichomonas vaginalis* infection partitioned with the PID training set suggesting that this method is sensitive for genital tract infection. Further supporting the validity of the signature is the finding that there was a correlation with elevation of the C-reactive protein.

There are several limitations to the current study, the largest of which is the limited sample size. Although both the training and validation sets contained patients with clinical PID, there were differences in identified infections between the two groups. Future studies would need to test the validity of our identified signature in patients with GC/CT infection in addition to the patients studied here. Patients with clinical PID but negative STI testing are a particularly interesting group. It is possible that these patients truly do have PID, and either that they have classic PID caused by organisms that are not identified by urine PCR/antigen

testing, or that they have PID caused by a novel organism. Alternatively, although these patients were diagnosed with PID in the ED, it is possible that these are false positive diagnoses based on these patients' experience of pain on pelvic exam, and not true upper genital tract inflammation. In addition, we cannot rule out the possibility that some control subjects may have had asymptomatic infection with STI, however as none of the subjects had abdominal pain at the time of enrollment, they did not meet criteria for PID. Development of a larger data set including patients with known infection as well as STI negative patients will help to address this question. Future studies designed to confirm the training set findings on a larger sample should be performed. Identification of a critical gene set which could serve as a rapid biomarker is a longer term goal.

Conclusion

We have demonstrated that RNA sample collection from adolescents in the ED setting is feasible. We identified a set of genes that were differentially expressed in PID patients vs. controls, many of which are involved in inflammation. It is our ultimate hope that these findings will lead to improved accuracy of diagnosis of PID in this high risk population. The ability to diagnose PID more accurately could lead to more rapid treatment and fewer long term complications from pelvic inflammatory disease. Furthermore, a relatively non-invasive test for adolescents would be advantageous. Patients with a non-inflammatory signature may not require further clinical evaluation, thus saving unnecessary antibiotics, unnecessary testing, resulting in reduction of health care cost and hospital admissions.

Acknowledgments

Funding Source:

The project described was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant UL1TR000003, as well as by the Foerderer Young Investigator Award at The Children's Hospital of Philadelphia. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The sponsor had no role in study design, collection, analysis, interpretation of data, writing of the manuscript, or the decision to submit the manuscript for publication. Dr. Balamuth received career development support from NIH NHLBI K12-HL109009, though the funders were not involved in design and conduct of the study; collection, management, analysis, interpretation of the data; preparation, review, or approval of the manuscript.

Abbreviations

BAZ2a	Bromodomain adjacent to zinc finger domain protein 2A
CBC	complete blood count
CDC	Centers for Disease Control and Prevention
CRP	C-reactive protein
CSF2RB	common subunit to Type I cytokine receptors
CT	Chlamydia Trachomatis
CT Scan	computed tomography scan
ED	Emergency Department

ESR	Erythrocyte sedimentation rate
FDR	false discovery rate
GC	Neisseria Gonorrhoeae
JAK3	Janus associated kinase 3
MLL4	mixed-lineage leukemia 4
OR	operating room
PAG1	phosphoprotein associated with glycolipid enriched microdomains
PCA	principal components analysis
PI3K	phosphatidyl inositol 3 kinase
PID	Pelvic Inflammatory Disease
RMA	robust multichip averaging
SAM	significance analysis of microarrays
SLC26A3	solute carrier family 26 member 3
SOCS3	suppressor of cytokines 3
STI	sexually transmitted infection
TV	Trichomonas Vaginalis

References

1. Gray-Swain MR, Peipert JF. Pelvic inflammatory disease in adolescents. *Curr Opin Obstet Gynecol.* 2006; 18(5):503–510. [PubMed: 16932044]
2. Simms I, Stephenson JM. Pelvic inflammatory disease epidemiology: what do we know and what do we need to know? *Sex Transm Infect.* 2000; 76(2):80–87. [PubMed: 10858707]
3. Sweet RL. Pelvic Inflammatory Disease: Current Concepts of Diagnosis and Management. *Curr Infect Dis Rep.* 2012
4. Reighard SD, Sweet RL, Vicetti Miguel C, Vicetti Miguel RD, Chivukula M, Krishnamurti U, et al. Endometrial leukocyte subpopulations associated with Chlamydia trachomatis, Neisseria gonorrhoeae, and Trichomonas vaginalis genital tract infection. *Am J Obstet Gynecol.* 2011; 205(4):324, e321–327. [PubMed: 21777898]
5. Short VL, Totten PA, Ness RB, Astete SG, Kelsey SF, Haggerty CL. Clinical presentation of Mycoplasma genitalium Infection versus Neisseria gonorrhoeae infection among women with pelvic inflammatory disease. *Clin Infect Dis.* 2009; 48(1):41–47. [PubMed: 19025498]
6. Bjartling C, Osseir S, Persson K. Mycoplasma genitalium in cervicitis and pelvic inflammatory disease among women at a gynecologic outpatient service. *Am J Obstet Gynecol.* 2012; 206(6):476, e471–478. [PubMed: 22483084]
7. Sexually Transmitted Diseases. [May 11, 2010] Treatment Guidelines. 2006. <http://www.cdc.gov/std/treatment/>.
8. Simms I, Warburton F, Westrom L. Diagnosis of pelvic inflammatory disease: time for a rethink. *Sex Transm Infect.* 2003; 79(6):491–494. [PubMed: 14663128]
9. Simms I, Stephenson JM, Mallinson H, Peeling RW, Thomas K, Gokhale R, et al. Risk factors associated with pelvic inflammatory disease. *Sex Transm Infect.* 2006; 82(6):452–457. [PubMed: 16901918]

10. Blake DR, Fletcher K, Joshi N, Emans SJ. Identification of symptoms that indicate a pelvic examination is necessary to exclude PID in adolescent women. *J Pediatr Adolesc Gynecol.* 2003; 16(1):25–30. [PubMed: 12604142]
11. Abu Raya B, Bamberger E, Kerem NC, Kessel A, Srugo I. Beyond “safe sex”—can we fight adolescent pelvic inflammatory disease? *Eur J Pediatr.* 2012
12. Goyal M, McCutcheon M, Hayes K, Mollen C. Sexual history documentation in adolescent emergency department patients. *Pediatrics.* 2011; 128(1):86–91. [PubMed: 21646260]
13. Garcia I, Mayol G, Rios J, Domenech G, Cheung NK, Oberthuer A, et al. A threegene expression signature model for risk stratification of patients with neuroblastoma. *Clin Cancer Res.* 2012; 18(7):2012–2023. [PubMed: 22328561]
14. Ramilo O, Allman W, Chung W, Mejias A, Ardura M, Glaser C, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood.* 2007; 109(5):2066–2077. [PubMed: 17105821]
15. Kentsis A, Lin YY, Kurek K, Calicchio M, Wang YY, Monigatti F, et al. Discovery and validation of urine markers of acute pediatric appendicitis using highaccuracy mass spectrometry. *Ann Emerg Med.* 55(1):62–70. e64. [PubMed: 19556024]
16. Mannick EE, Bonomolo JC, Horswell R, Lentz JJ, Serrano MS, Zapata-Velandia A, et al. Gene expression in mononuclear cells from patients with inflammatory bowel disease. *Clin Immunol.* 2004; 112(3):247–257. [PubMed: 15308118]
17. Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, et al. An interferoninducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature.* 2010; 466(7309): 973–977. [PubMed: 20725040]
18. Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet.* 2006; 7(1):55–65. [PubMed: 16369572]
19. Murphy CG, Glickman JN, Tomczak K, Wang YY, Beggs AH, Shannon MW, et al. Acute appendicitis is characterized by a uniform and highly selective pattern of inflammatory gene expression. *Mucosal Immunol.* 2008; 1(4):297–308. [PubMed: 19079191]
20. Fuss IJ, Neurath M, Boirivant M, Klein JS, de la Motte C, Strong SA, et al. Disparate CD4+ lamina propria (LP) lymphokine secretion profiles in inflammatory bowel disease. Crohn's disease LP cells manifest increased secretion of IFN-gamma, whereas ulcerative colitis LP cells manifest increased secretion of IL-5. *J Immunol.* 1996; 157(3):1261–1270. [PubMed: 8757634]
21. Reed JL, Strait RT, Kachelmeyer AM, Byczkowski TL, Ho ML, Huppert JS. Biomarkers to distinguish surgical etiologies in females with lower quadrant abdominal pain. *Acad Emerg Med.* 2011; 18(7):686–691. [PubMed: 21762231]
22. Gondek DC, Roan NR, Starnbach MN. T cell responses in the absence of IFNgamma exacerbate uterine infection with *Chlamydia trachomatis*. *J Immunol.* 2009; 183(2):1313–1319. [PubMed: 19561106]
23. Imarai M, Candia E, Rodriguez-Tirado C, Tognarelli J, Pardo M, Perez T, et al. Regulatory T cells are locally induced during intravaginal infection of mice with *Neisseria gonorrhoeae*. *Infect Immun.* 2008; 76(12):5456–5465. [PubMed: 18824531]
24. Collett BA, Newhall WJ, Jersild RA Jr. Jones RB. Detection of surface-exposed epitopes on *Chlamydia trachomatis* by immune electron microscopy. *J Gen Microbiol.* 1989; 135(1):85–94. [PubMed: 2476526]
25. Taylor BD, Darville T, Ferrell RE, Kammerer CM, Ness RB, Haggerty CL. Variants in toll-like receptor 1 and 4 genes are associated with *Chlamydia trachomatis* among women with pelvic inflammatory disease. *J Infect Dis.* 2012; 205(4):603–609. [PubMed: 22238472]
26. Asare AL, Kolchinsky SA, Gao Z, Wang R, Raddassi K, Bourcier K, et al. Differential gene expression profiles are dependent upon method of peripheral blood collection and RNA isolation. *BMC Genomics.* 2008; 9:474. [PubMed: 18847473]
27. Erb T, Beigi RH. Update on infectious diseases in adolescent gynecology. *J Pediatr Adolesc Gynecol.* 2008; 21(3):135–143. [PubMed: 18549965]

Figure 1.

Hierarchical clustering of 18 training samples including 9 PID patients (red) vs. 9 controls (green) and the 170 genes differentially expressed between the two groups ($p < 0.001$, $FDR=0.04$). Each column represents an individual patient, and each row represents a gene product.

Figure 2.
Network analysis of differentially expressed genes using Ingenuity software. Genes in red are overexpressed in PID patients as compared to controls.

Figure 3.

Heatmap demonstrated supervised clustering of entire data set. PID patients are represented in pink, control patients in green. Unknown patients are in blue. Each column represents an individual patient and each row represents a single gene. In the gene expression data, red indicates increased expression, while yellow indicates decreased expression.

Figure 4.

Principal component analysis using 170 DEGs and 35 samples, including 9 training PID patients (red), 9 training controls (green), 5 testing PID patients with *Trichomonas vaginalis* infection (orange), 2 testing controls (cyan), and 10 patients with clinical PID but no identified sexually transmitted infection (grey). Dashed lines arbitrarily split all samples into three clusters.

Figure 5.

Correlation between gene expression pattern and known biomarkers. Top panels corresponds to markers measured as numeric values and the p values were calculated by Spearman correlation. The bottom panels corresponds to markers with binomial values and the p values were calculated via applying Student's t test on two groups of principal component 1 values.

Table 1

Comparison of mean laboratory values between STI+ and STI- patients. STI+ patients tested positive for either CT, GC, or TV.

	STI+ (n=11)	STI- (n=10)	p
WBC	9.2	7.0	0.25
ANC	6186	4276	0.34
CRP	2.5	0.8	0.21
ESR	14.1	10.9	0.51