

## ARTICLE OPEN

Comparative transcriptomic analysis of *Gardnerella vaginalis* biofilms vs. planktonic cultures using RNA-seqJoana Castro<sup>1,2,3</sup>, Angela França<sup>1</sup>, Katie R. Bradwell<sup>4</sup>, Myrna G. Serrano<sup>4</sup>, Kimberly K. Jefferson<sup>3</sup> and Nuno Cerca<sup>1</sup> 

Bacterial vaginosis is the most common gynecological disorder affecting women of reproductive age. Bacterial vaginosis is frequently associated with the development of a *Gardnerella vaginalis* biofilm. Recent data indicates that *G. vaginalis* biofilms are more tolerant to antibiotics and are able to incorporate other bacterial vaginosis-associated species, yielding a multi-species biofilm. However, despite its apparent role in bacterial vaginosis, little is known regarding the molecular determinants involved in biofilm formation by *G. vaginalis*. To gain insight into the role of *G. vaginalis* in the pathogenesis of bacterial vaginosis, we carried out comparative transcriptomic analysis between planktonic and biofilm phenotypes, using RNA-sequencing. Significant differences were found in the expression levels of 815 genes. A detailed analysis of the results obtained was performed based on direct and functional gene interactions. Similar to other bacterial species, expression of genes involved in antimicrobial resistance were elevated in biofilm cells. In addition, our data indicate that *G. vaginalis* biofilms assume a characteristic response to stress and starvation conditions. The abundance of transcripts encoding proteins involved in glucose and carbon metabolism was reduced in biofilms. Surprisingly, transcript levels of vaginolysin were reduced in biofilms relative to planktonic cultures. Overall, our data revealed that gene-regulated processes in *G. vaginalis* biofilms resulted in a protected form of bacterial growth, characterized by low metabolic activity. This phenotype may contribute towards the chronic and recurrent nature of bacterial vaginosis. This suggests that *G. vaginalis* is capable of drastically adjusting its phenotype through an extensive change of gene expression.

npj Biofilms and Microbiomes (2017)3:3; doi:10.1038/s41522-017-0012-7

## INTRODUCTION

Bacterial vaginosis (BV) is the most prevalent vaginal condition in women of reproductive age and can cause several problems, such as preterm birth, endometritis, and increased risk of acquisition and transmission of sexual transmitted diseases.<sup>1</sup> Examination of vaginal biopsy specimens has demonstrated that most cases of BV are characterized by the adherence of a bacterial biofilm to the vaginal epithelium, and that *Gardnerella vaginalis* is the predominant species of the biofilm mass.<sup>2</sup> However, *G. vaginalis* colonization does not always lead to BV.<sup>3</sup> Biofilm formation represents a protected mode of growth that allows cells to survive in the acidic vaginal environment.<sup>4</sup> *G. vaginalis* can also adopt a planktonic phenotype that differs greatly from biofilm lifestyle.<sup>5</sup> It is postulated that a biofilm provides an ecological advantage over planktonic bacteria.<sup>6</sup>

Importantly, biofilm infections are particularly problematic because sessile bacteria are generally much more tolerant to antibiotics than planktonic cells.<sup>6</sup> Evidence suggests that biofilm formation contributes significantly to BV treatment failure and high recurrence rates.<sup>7,8</sup> Targeting virulence factors represents a new paradigm in the development of new and effective treatments to prevent and treat biofilm-associated infections.<sup>9</sup> Therefore, a better understanding of BV-associated *G. vaginalis* biofilm physiology and virulence is needed to understand the high persistence and resistance of biofilm cells.

The purpose of our study was, therefore, to identify the major transcriptomic features of BV-associated *G. vaginalis* biofilms, as compared to their planktonic counterparts, using high-throughput RNA-sequencing (RNA-seq). Transcriptomic comparisons between biofilm and planktonic cultures that have been carried out for *Staphylococcus aureus*,<sup>10</sup> *Staphylococcus epidermidis*,<sup>11</sup> *Streptococcus mutans*,<sup>12</sup> and *Streptococcus pneumoniae*,<sup>13</sup> indicate that gene-regulated processes in the biofilm led to a protective mode of growth by developing an effective cellular response to stress and decreasing metabolic activity.

Herein, we sequenced the transcriptome of BV-associated *G. vaginalis* biofilms and planktonic cultures and used a data analysis approach based on direct and functional gene interactions, namely gene set enrichment and cluster analysis.

## RESULTS

## Transcriptome analysis

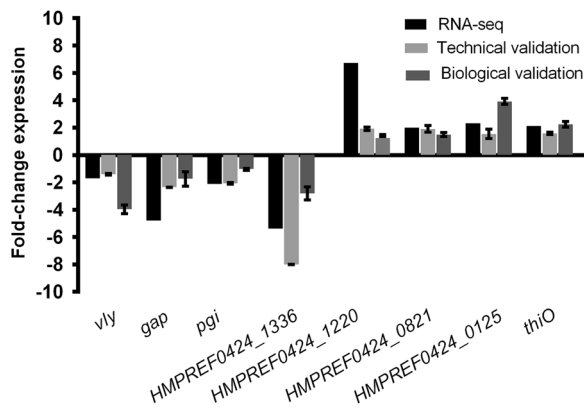
A total of 561,302 (planktonic phenotype) and 311,643 (biofilm phenotype) sequencing reads were obtained for the complementary DNA (cDNA) libraries. Before trimming the raw data, we identified the genes, with the reads per kilobase per million (RPKM) above 1.00, expressed in each condition. We only detected three genes uniquely expressed in biofilm cells, whereas 11 genes were found uniquely in planktonic cells. However, the majority of

<sup>1</sup>Centre of Biological Engineering (CEB), Laboratory of Research in Biofilms Rosário Oliveira (LIBRO), University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal; <sup>2</sup>Instituto de Ciências Biomédicas Abel Salazar (ICBAS), University of Porto, Rua de Jorge Viterbo Ferreira 228, 4050-313 Porto, Portugal; <sup>3</sup>Department of Microbiology and Immunology, Virginia Commonwealth University, Richmond, VA 23298-0678c, USA and <sup>4</sup>Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA 23284, USA

Correspondence: Nuno Cerca ([nunocerca@ceb.uminho.pt](mailto:nunocerca@ceb.uminho.pt))

Received: 20 July 2016 Revised: 17 November 2016 Accepted: 22 November 2016

Published online: 02 February 2017



**Fig. 1** qPCR validation of the transcription of differentially expressed genes randomly selected. Technical validation means that we used the same total RNA utilized for libraries construction. Biological validation means that we used new total RNA obtained from independent experiments performed under same biological conditions. The data indicate the fold-change expression of genes in *G. vaginalis* biofilms cells compared to planktonic cells. For qPCR experiments, the bars represent the mean and the error bars the standard error of the mean (mean  $\pm$  SEM)

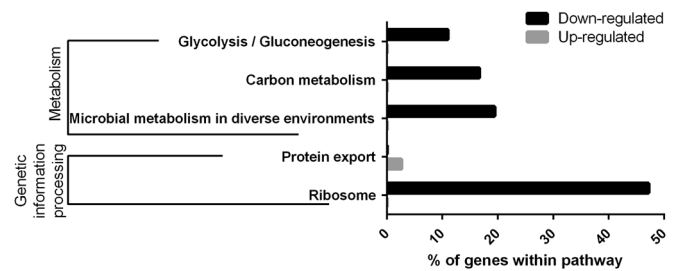
gene transcripts that were only detected in planktonic or biofilm cells, encoded uncharacterized proteins or transfer RNA, as shown in the Supplementary Material (Table S1).

Our data indicated that within the 1045 genes that were transcribed in both conditions, 815 (78%) were differentially expressed between planktonic and biofilm cells. For downstream analysis, only genes with fold-changes above two were considered. Transcript levels of 309 (30%) genes were elevated, whereas 36 (3%) were reduced in biofilms. Among the transcripts that were more abundant in biofilms, 78 encoded hypothetical proteins. In an effort to find homology with known proteins, we performed a BLAST analysis, a search in the Pfam database (version 29.0) for Pfam domains<sup>14</sup> and used the PSORTb program (v.3.0)<sup>15</sup> to predict their subcellular localization. The results are shown in Table S2. Interestingly, 53% of these proteins might have cytoplasmic membrane localization, suggesting that part of these proteins could have a transporter function.

In order to confirm the results obtained by RNA-seq, transcripts detected in greater or lesser abundance in biofilms were randomly selected and their relative levels quantified by quantitative PCR (qPCR). Both RNA used for cDNA libraries construction (technical validation) and RNA obtained by performing new experiments (biological validation) were used for validation. As can be seen in Fig. 1, the same trend was observed in all measurements (qPCR and RNA-seq).

#### Enrichment analysis of genes with increased and decreased transcription

GO annotation, placement of genes on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and enrichment analysis of the genes with down and upregulated transcription was accomplished using STRING.<sup>16</sup> Significant enrichment was only found ( $p < 0.05$ , false discovery rate (FDR)-corrected) in KEGG pathways (Fig. 2). As could be expected, classes associated with metabolism were found significantly enriched among the genes with decreased transcription, suggesting that biofilm cells were less metabolically active than planktonic cells. Conversely, protein export was found to be enriched among transcripts that were elevated in biofilm cells.



**Fig. 2** KEGG pathways found significantly enriched ( $p < 0.05$ ) within the genes with increased and decreased transcription in biofilm cells

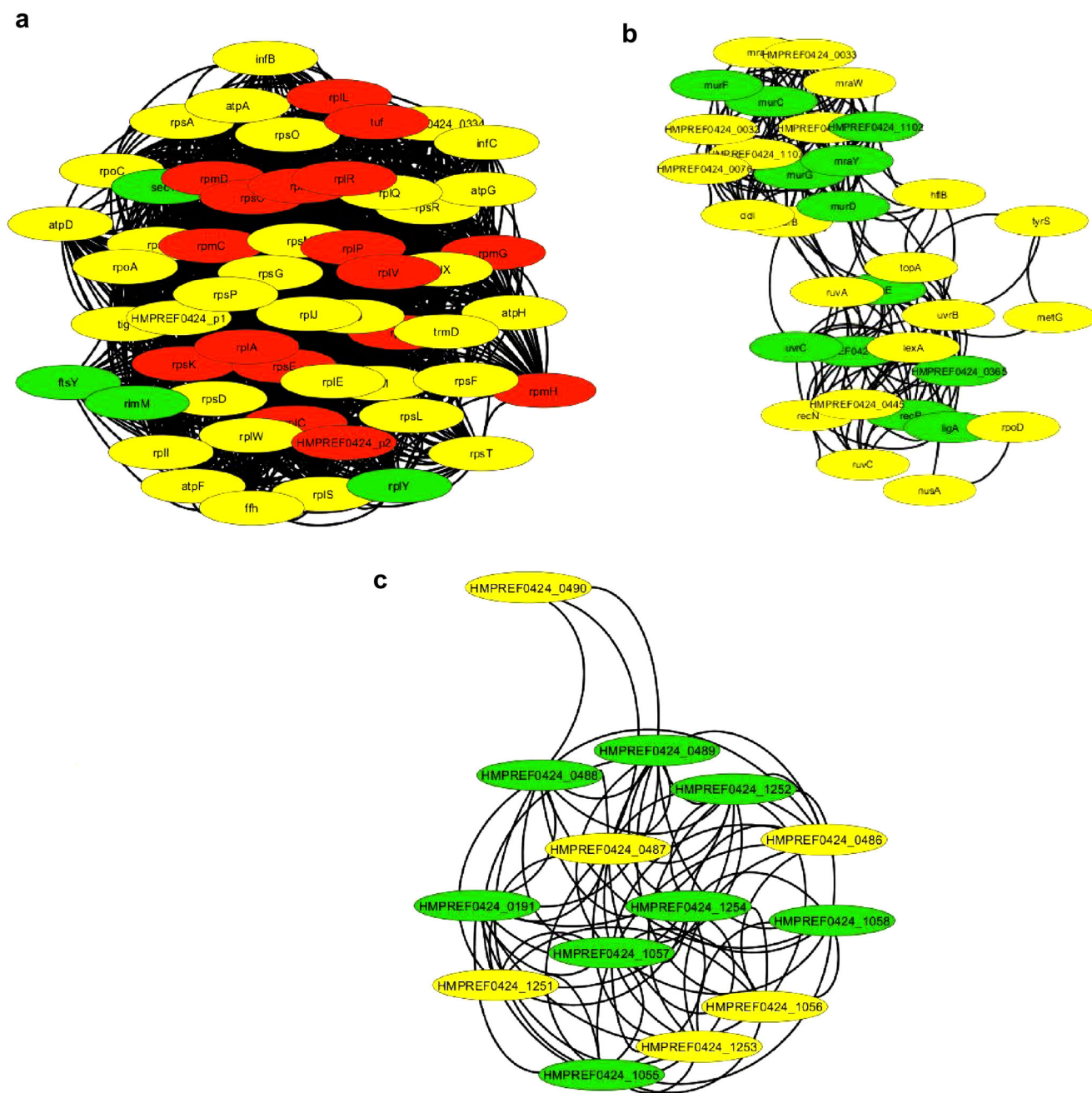
#### Cluster analysis

Gene clustering analysis was based on direct and functional gene interactions using Cytoscape.<sup>17</sup> Cytoscape was used to create a gene interaction network including all differentially expressed genes and neighbors, yielding a total of 764 nodes and 7685 edges (complete gene network of differentially expressed genes is shown in the Supplementary Material Fig. S1). Among the differentially expressed genes, we found 22 clusters. Significant enrichment was found ( $p < 0.05$ , FDR-corrected) in biological processes or KEGG pathways associated with translation and metabolic process (Fig. 3a), cell-wall biogenesis and mismatch repair (Fig. 3b), and antimicrobial resistance (Fig. 3c).

#### The top 10 most significantly down or upregulated genes in biofilms

Table 1 lists the 10 transcripts with the greatest increase and the 10 with the greatest decrease in biofilm cells. Among the transcript decreased, we found ribosomal proteins suggesting that biofilms had decreased level of translation. Furthermore, BV-associated *G. vaginalis* biofilm cells showed decreased transcript levels of genes encoding several factors involved in energy production, such as *HMPREF0424\_1336*, a gene encoding primary receptors for chemotaxis and transport of many sugar based solutes. In addition, the expression levels of genes associated with glucose metabolic pathways were also lower in biofilms cells, namely *gap*, that also has a role in oxidoreductase activity, and *gpmA*, that displays an important role in a subpathway of the glycolysis pathway (glycolysis/gluconeogenesis pathway of *G. vaginalis* 409-05 is shown in the Supplementary Material Fig. S2), which itself is part of carbohydrate degradation. Taken together, these results imply that *G. vaginalis* biofilm cells are characterized by the reduction of basic cell processes (translation) and metabolism (glycolysis and carbon metabolism).

Among the transcripts elevated in biofilm cells, we found *HMPREF0424\_0563*, a gene with a molecular function related to hydrolase activity. Furthermore, in biofilm cells we found an overexpression of the *HMPREF0424\_1220* gene encoding an aminotransferase involved in amino acid biosynthesis. A similar trend was reported for *Neisseria meningitidis*.<sup>18</sup> Interestingly, we found *HMPREF0424\_0420*, a gene that encodes the LPXTG-motif cell anchor domain-containing protein, which can be involved in biofilm formation, as described in Gram-positive bacteria.<sup>19</sup> Moreover, in BV-associated *G. vaginalis* biofilm cells, transcript levels of the gene *HMPREF0424\_0573*, which encodes a LysM domain-containing protein possibly associated with autoaggregation of *G. vaginalis*, were also increased, similar to what was observed for *Lactobacillus reuteri* biofilms.<sup>20</sup> Transcripts encoding the ComEA protein (*HMPREF0424\_0943*), which is involved in DNA repair and NLPA lipoprotein (*HMPREF0424\_0888* gene), which is involved in ABC transporters were also found in greater abundance in biofilms cells.



**Fig. 3** Clusters generated by the MCODE plugin in Cytoscape. Red, green, and yellow circles represent fold-change values under  $-2$ , above  $2$ , and between  $-2$  and  $2$ , respectively. Biological process indicates enrichment ( $p < 0.05$ ) in translation and metabolic process (a); cell-wall biosynthesis biogenesis and mismatch repair (b); and KEGG indicates  $\beta$ -Lactamase resistance (c)

Upregulation of the transcription of potential virulence genes in *G. vaginalis* biofilms

Biofilm formation by pathogenic bacteria is often associated with altered virulence. Bacterial biofilms may suppress certain virulence factors while others are activated in order to evade immune defenses, and survive challenging conditions.<sup>21</sup> It was, therefore, of interest to determine the expression levels of previously annotated potential virulence genes.<sup>22</sup> As can be seen in Fig. 4, we found a slight increase in the *HMPREF0424\_0125* transcript, which encodes *TadE*-like protein. This might play an important role in adhesion to vaginal epithelial cells, similarly to what has been described in *Actinobacillus actinomycetemcomitans*.<sup>23</sup> Furthermore,

our data supported the previous hypothesis that *G. vaginalis* biofilm development is likely associated with type II glycosyltransferase.<sup>22</sup> Of note, glycosyltransferases are likely to be important for the biosynthesis of exopolysaccharide which in turn is important for biofilm formation. The ability to grow as a biofilm would likely confer an increase in antibiotic tolerance and resistance to mucosal immune defenses.<sup>24</sup> Herein, levels of transcripts encoding antimicrobial-specific resistance proteins belonging to efflux pump families were increased. In addition, the *HMPREF0424\_1196* transcript, which encodes a Rib-protein, was elevated in biofilm cells. Rib proteins belong to the  $\alpha$ -like protein (Alp)-family of highly repetitive surface antigens and are commonly found in Gram-positive pathogens.<sup>25</sup> These proteins

**Table 1.** List of the 10 genes with lowest and highest fold-change values among the differentially expressed genes in *G. vaginalis* cultured under biofilm vs. planktonic conditions

Gene	Definition	Fold-change (Biofilm vs. planktonic cells)
<b>Rank Downregulated</b>		
1	<i>HMPREF0424_0046</i> 50S ribosomal protein L34	-21.93
2	<i>HMPREF0424_0269</i> 50S ribosomal protein L30	-8.99
3	<i>HMPREF0424_0429</i> Exodeoxyribonuclease VII large subunit ( <i>xseA</i> )	-6.73
4	<i>HMPREF0424_0260</i> 30S ribosomal protein S3	-5.43
5	<i>HMPREF0424_1336</i> Periplasmic-binding protein and sugar-binding domain of the LacI family protein	-5.39
6	<i>HMPREF0424_0259</i> 50S ribosomal protein L22	-5.36
7	<i>HMPREF0424_0258</i> 50S ribosomal protein L2	-4.93
8	<i>HMPREF0424_0471</i> Glyceraldehyde 3-phosphate dehydrogenase domain-containing protein ( <i>gap</i> )	-4.81
9	<i>HMPREF0424_0276</i> 30S ribosomal protein S11	-3.86
10	<i>HMPREF0424_0394</i> Phosphoglyceratemetase ( <i>gpmA</i> )	-3.62
<b>Rank Upregulated</b>		
1	<i>HMPREF0424_0510</i> Uncharacterized protein	14.41
2	<i>HMPREF0424_0563</i> Pyroglutamyl-peptidase I	9.57
3	<i>HMPREF0424_1220</i> Aminotransferase, class I/II	6.74
4	<i>HMPREF0424_0420</i> LPXTG-motif cell wall anchor domain-containing protein	6.30
5	<i>HMPREF0424_0397</i> Uncharacterized protein	5.76
6	<i>HMPREF0424_0573</i> LysM domain-containing protein	5.51
7	<i>HMPREF0424_0943</i> ComEA protein	5.15
8	<i>HMPREF0424_0166</i> Uncharacterized protein	4.28
9	<i>HMPREF0424_0888</i> NLPA lipoprotein	4.24
10	<i>HMPREF0424_0797</i> Uncharacterized protein	4.12

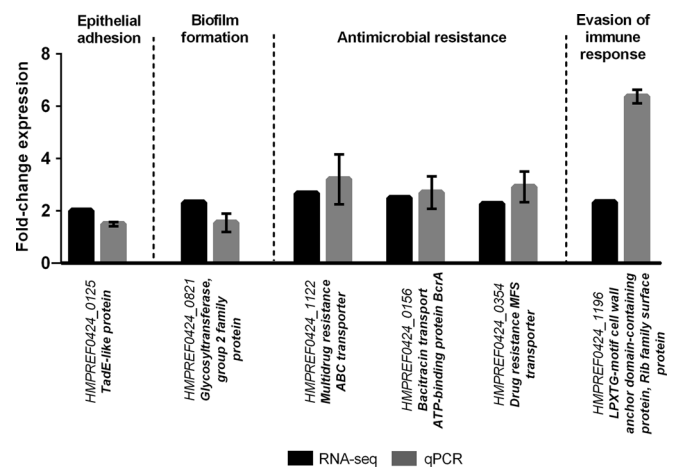
elicit protective immunity through their inter-strain size variability.<sup>22</sup>

Differential expression of vaginolysin BV-associated *G. vaginalis* biofilms

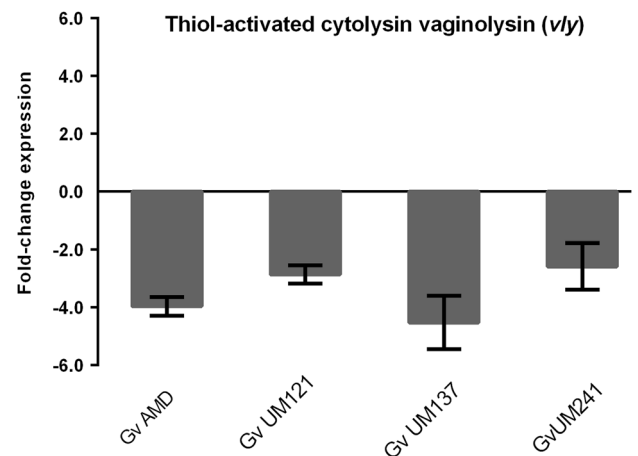
*G. vaginalis* produces a thiol-activated cholesterol-dependent cytolytic, vaginolysin (*vly*), which might induce vaginal cells lysis. Strikingly, in our experiments, the expression levels of *vly* (*HMPREF0424\_0103*) were significantly lower in biofilm cells (Fig. 1). In order to determine whether this was a strain-specific variation, we evaluated *vly* gene expression, by qPCR, in three other biofilm forming isolates, which were previously characterized.<sup>26</sup> Interestingly, as shown in Fig. 5, the downregulation of the transcription of *vly* was observed in all different isolates.

## DISCUSSION

As noted elsewhere, gene expression profiles can reveal important information about the adaptation of a bacterial species to a particular environmental niche. Therefore, adaptation to a given host environment is an extremely important factor and underlies the capacity of a colonizing species or a pathogen to persist in a



**Fig. 4** Quantification of the transcription of known virulence genes in *G. vaginalis* cultured under biofilm and planktonic conditions. Bars represent the mean and the error bars the standard error of the mean (mean  $\pm$  SEM)



**Fig. 5** Quantification of thiol-activated cytolytic vaginolysin (*vly*) transcription in *G. vaginalis* strains. Bars represent the mean and the error bars the standard error of the mean (mean  $\pm$  SEM)

host.<sup>6</sup> In the present study, we analyzed the transcriptome of a BV-associated *G. vaginalis* cultivated under biofilm and planktonic conditions. Our results demonstrated that more transcripts were increased in biofilm relative to planktonic cells. Importantly, our findings provide key insights into the development of biofilms and the pathogenicity of *G. vaginalis*, the predominant bacterial species isolated in women with BV.<sup>27,28</sup>

Here, we showed that BV-associated *G. vaginalis* biofilm cells alter their gene expression profile, namely transcript levels of genes involved in metabolism (with downregulation of genes associated with glycolysis and carbon metabolism) and translation (with downregulation of genes encoding ribosomal proteins), as also reported for other microorganisms such as *S. epidermidis*,<sup>11</sup> *S. aureus*,<sup>10</sup> and *S. mutans*.<sup>12</sup> In *G. vaginalis* biofilms, cell density is substantially higher than in planktonic culture.<sup>4</sup> As a consequence, most biofilm cells are likely to encounter restricted availability of nutrients.<sup>29</sup> Similar to what was found for *S. mutans*<sup>30</sup> and *S. pneumoniae*,<sup>13</sup> we also observed that the transcripts encoding ABC transporter proteins were elevated in biofilm cells. In addition, our study revealed that transcripts of genes involved in the synthesis of peptidoglycan and cell wall were also greater in biofilms. This has also been shown for *P. aeruginosa*<sup>31</sup> and *S. aureus*.<sup>10</sup> It has been hypothesized that the cell envelope is a highly dynamic and



active component of biofilm cells, contributing to its persistence.<sup>10,13</sup> However, the reasons for the overexpression of genes involved in cell wall biogenesis require further investigation.

Notably, transcripts of other potential virulence genes, previously annotated by Yeoman and colleagues,<sup>22</sup> were also more abundant in biofilm cells, with the exception of *vly*. Several studies have highlighted the role of *vly* gene in *G. vaginalis* virulence.<sup>32,33</sup> The *vly* gene belongs to the cholesterol-dependent cytolysins (CDCs), a family of pore-forming toxins, which cause cytotoxicity on vaginal epithelium.<sup>33</sup> Our previous findings showed that planktonic cultures of BV isolates of *G. vaginalis* expressed twofold more *vly* than planktonic cultures of non-BV isolates.<sup>26</sup> Herein, we found that *vly* transcript levels were higher in planktonic than in biofilms cells. The low levels of expression of *vly* in biofilms might reflect the more chronic nature of vaginal colonization by BV-associated *G. vaginalis* and serve as a means towards preventing a host immune response. Similarly, Resch and colleagues<sup>10</sup> showed that the production of various *S. aureus* toxins were significantly upregulated in planktonic rather than in biofilm cells, suggesting that toxins may not be conducive to biofilm persistence in the host.

Similar to what was observed in other microorganisms, BV-associated *G. vaginalis* biofilm phenotype might induce a quiescent mode of growth that is less sensitive to antibiotics, as the efficacy of many antibiotics relies on active cell metabolism and the cell-wall construction process.<sup>11</sup> Here, we observed that efflux pumps and ABC transporters, reported as mechanisms responsible for antimicrobial resistance,<sup>34</sup> were upregulated in biofilms cells. Comparable evidence for the role of efflux pumps in biofilm resistance has been found in several microorganisms such as *Pseudomonas aeruginosa*,<sup>35</sup> *Escherichia coli*,<sup>36</sup> and *Candida albicans*.<sup>37</sup>

Taken together, these data indicated that BV-associated *G. vaginalis* changes its transcriptomic profile when growing as a biofilm. These changes are likely important for biofilm persistence and, consequently, for the virulence of this bacterium. Furthermore, the fact that *vly* is downregulated in biofilms represents an important finding, that might contribute towards a better understanding of the pathogenesis of BV. However, this study is limited by the fact that the growth medium did not contain all of the factors found *in vivo*, and some *in vivo* cues may turn on the expression of biofilm-related genes. Nevertheless, as animal models for BV are lacking, *in vitro* models can be very informative, and are key to furthering our understanding of virulence potential of *G. vaginalis*. In conclusion, our findings showed that the gene expression profile of BV-associated *G. vaginalis* biofilms characterizes a distinct physiologic status that may promote the chronic and recurrent nature of BV.

## MATERIALS AND METHODS

### Bacterial strains

*G. vaginalis* strain AMD, isolated from a woman diagnosed with BV based on Amsel criteria at VCU Women's Health Clinic,<sup>38</sup> was used for RNA-seq analysis. *G. vaginalis* strains UM121, UM137, and UM241, also isolated from women with BV based on Amsel and Nugent criteria,<sup>26</sup> were used for subsequent analysis.

### Planktonic growth

Planktonic cells were grown in sBHI [Brain-heart infusion supplemented with 2% (wt/wt) gelatin (Liofilchem, Rosetodegli Abruzzi, Italy), 0.5% (wt/wt) yeast extract (Liofilchem) and 0.1% (wt/wt) soluble starch (Panreac, Barcelona, Spain)] for 24 h at 37 °C with 10% CO<sub>2</sub> (Shel Lab, Cornelius, Oregon, USA).<sup>26</sup> At this time, planktonic cells were still in the exponential growth phase. Thereafter, 18 mL of planktonic cells were harvested by centrifugation (20 min, 7197×g) and suspended in 1 mL of RNA protect [diluted 2:1 in phosphate-buffered saline (PBS); QIAGEN, Germany].

### Biofilm formation

For biofilm formation, the cell concentration of 24 h old cultures was assessed by optical density at 600 nm (Model Sunrise, Tecan, Switzerland) and was further diluted in order to obtain a final concentration of approximately 10<sup>6</sup> CFU/mL. After homogenization, 200 µL of *G. vaginalis* suspensions were dispensed into each well of three 96-well flat-bottom tissue culture plates (Orange Scientific, Braine L'Alleud, Belgium). The tissue culture plates were then incubated at 37 °C in 10% CO<sub>2</sub>. After 24 h, the culture medium covering the biofilms was removed, replaced by fresh sBHI and allowed to grow, under the same conditions, for an additional 24 h. This time was required for this strain to develop a notable biofilm. Forty-eight hour biofilms were then washed once with 1 × PBS, scraped from the bottom of 96-well plates in sBHI and pooled together. Finally, biofilm cells were harvested by centrifugation (20 min, 7197×g) and suspended in 1 mL of RNA protect (as described above).

### RNA extraction

Total RNA was extracted using a combination of mechanical lysis (3.0 mm zirconium beads, Sigma-Aldrich Inc., St. Louis, MO, USA) and the columns of the RNeasy Mini kit (QIAGEN), as optimized before (França *et al.*<sup>49</sup>). To remove genomic DNA, TURBO DNA-free™ kit (Ambion, Austin, TX, USA) was used as indicated by the manufacturer followed by acid-phenol:chloroform:isoamyl alcohol (125:24:1) treatment. RNA integrity was determined using an Experion™ automated electrophoresis system (Bio-Rad, Hercules, CA, USA), and samples with RNA Quality Indicator (RQI) above eight were selected for cDNA library preparation.

### cDNA library preparation and sequencing

cDNA libraries were constructed using the kit ScriptSeq™ Complete Kit—low input (Illumina, San Diego, WI, USA), which already includes the kit for ribosomal (rRNA) depletion: Ribo-Zero™ Kit (Bacteria)—Low Input (Illumina, Madison, WI, USA). The construction of the libraries was rigorously validated by qPCR and Hi-Sensitivity D1K TapeStation (Agilent 2200 TapeStation). Libraries were then multiplexed and sequencing data generated from paired-end reads (2 × 150 bp) using a MiSeq system (Illumina).

### RNA-sequencing data analysis

After sequencing, adapters were trimmed by MiSeq® internal software during the base calling. Quality, ambiguity and length trimming, as well as mapping to the reference genome, and normalization of gene expression were performed using CLC Genomics Workbench version 8 (MA, USA). Quality, ambiguity and length trimming were performed using the CLC genomics workbench default settings. RNA-seq reads were aligned to the reference genome of *G. vaginalis* strain 409-05 (GenBank accession number NC\_013721). Gene expression was normalized using RPKM, that account for both library size and gene length, as described by Mortazavi and colleagues.<sup>39</sup> To detect significant gene expression alterations, Kal's test<sup>40</sup> with FDR<sup>41</sup> correction was applied. A *p*-value ≤ 0.05 was considered statistically significant. Transcripts uniquely expressed in each condition were identified using BioinfoGP.<sup>42</sup> Data were deposited at Gene Expression Omnibus database (accession number GSE8012, available at: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80127>).

### Biological interactions

In order to determine the function of differentially expressed genes, gene ontology (GO)<sup>43</sup> and KEGG pathway<sup>44</sup> assignments and enrichment analysis were performed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (version 10).<sup>16</sup> In addition, UniProt repository<sup>45</sup> was used to determine the function of proteins that were not identified by STRING. Classes with *p*-value ≤ 0.05, FDR-adjusted, were considered statistically significant for enrichment. Further analysis was carried out using Cytoscape (version 3.2.1),<sup>46</sup> in which a gene interaction network including all differentially expressed genes and neighbors created by STRING<sup>16</sup> was imported. Gene clusters (regions of high connectivity) were obtained in Cytoscape with the MCODE plugin.<sup>47</sup> Default parameters (score value above two and at least four nodes) were used as the cutoff criteria for network module screening. Thereafter, an enrichment analysis of clusters was performed using STRING with a threshold of *p* < 0.05, FDR-adjusted.<sup>16</sup>

## Quantitative PCR

In order to validate RNA-seq data, qPCR was performed to quantify the transcription of eight randomly selected genes, by using the same total RNA utilized for libraries construction (technical validation) and new total RNA obtained from independent experiments performed under the same biological conditions (biological validation). Furthermore, the gene expression profile of known virulence genes was also addressed. Oligonucleotide primers were designed using Primer3<sup>48</sup> having *G. vaginalis* 409-05 genome as template (Table S3). The same amount of total RNA (300 ng/μL) was reverse transcribed using the RevertAid™ First Strand cDNA synthesis kit (Fermentas, Thermo Fisher Scientific, Stanford, CA, USA), as previously optimized,<sup>49</sup> and random primers (NZYTech, Lisbon, Portugal) as priming strategy. The qPCR reaction was prepared by mixing together 5 μL of iQ SYBR green supermix (Bio-Rad), 2 μL of 1:800 diluted cDNA, 0.5 μL of 5 μM forward and reverse primers and water up to 10 μL. The run was performed in a CFX96™ thermal cycler (Bio-Rad) with the following cycling parameters: 3 min at 95 °C, followed by 45 cycles of 10 s at 95 °C, 10 s at 60 °C, and 15 s at 72 °C. Reaction efficiency was determined by the dilution method.<sup>50</sup> At 60 °C all set of primers used had the highest and more similar efficiencies. Normalized gene expression was determined by using the delta C<sub>t</sub> method ( $E^{\Delta C_t}$ ), a variation of the Livak method, where  $\Delta C_t = C_t$  (reference gene) –  $C_t$  (target gene) and  $E$  stands for the reaction efficiency experimentally determined. A non-reverse transcriptase control was included in each reaction. Three biologic replicates of each condition were analyzed.

## ACKNOWLEDGEMENTS

This work was presented at Biofilms7 meeting (26–28 June 2016, Porto, Portugal) and was awarded the Biofilms7—SPM Young Researcher Award. N.C. is an Investigador FCT. This work was funded by Fundação para a Ciência e a Tecnologia (FCT) by the strategic project of UID/BIO/04469/2013 unit and the project RECI/BBB-EBI/0179/2012 (FCOMP-01-0124-FEDER-027462), and by the National Institutes of Health; P60 MD002256 “VCU NIMHD Comprehensive Center of Excellence”. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The following authors had an individual FCT fellowship: J.C. (SFRH/BD/93963/2013) and A.F. (SFRH/BPD/99961/2014).

## AUTHOR CONTRIBUTIONS

J.C. and A.F. performed the experiments. J.C., K.R.B. and M.G.S. performed the bioinformatics analysis. N.C. and K.K.J. designed the study. J.C. and N.C. drafted the manuscript. All authors critically reviewed and approved the final manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## REFERENCES

- Marrazzo, J. M. Vaginal biofilms and bacterial vaginosis: of mice and women. *Infect. Dis.* **207**, 1481–1483 (2013).
- Swidsinski, A. et al. Adherent biofilms in bacterial vaginosis. *Obstet. Gynecol.* **106**, 1013–1023 (2005).
- Marrazzo, J. M. et al. Characterization of vaginal flora and bacterial vaginosis in women who have sex with women. *J. Infect. Dis.* **185**, 1307–1313 (2002).
- Patterson, J. L., Girerd, P. H., Karjane, N. W. & Jefferson, K. K. Effect of biofilm phenotype on resistance of *Gardnerella vaginalis* to hydrogen peroxide and lactic acid. *Am. J. Obstet. Gynecol.* **197**, 170.e171–177 (2007).
- Verstraelen, H. & Swidsinski, A. The biofilm in bacterial vaginosis: implications for epidemiology, diagnosis and treatment. *Curr. Opin. Infect. Dis.* **26**, 86–89 (2013).
- Jefferson, K. K. What drives bacteria to produce a biofilm? *FEMS Microbiol. Lett.* **236**, 163–173 (2004).
- Swidsinski, A., Loening-Baucke, V., Swidsinski, S. & Verstraelen, H. Polymicrobial *Gardnerella* biofilm resists repeated intravaginal antiseptic treatment in a subset of women with bacterial vaginosis: a preliminary report. *Arch. Gynecol. Obstet.* **291**, 605–609 (2015).
- Muzny, C. A. & Schwabe, J. R. Biofilms: an underappreciated mechanism of treatment failure and recurrence in vaginal infections. *Clin. Infect. Dis.* **61**, 601–606 (2015).
- Cegelski, L., Marshall, G. R., Eldridge, G. R. & Hultgren, S. J. The biology and future prospects of antivirulence therapies. *Nat. Rev. Microbiol.* **6**, 17–27 (2008).
- Resch, A., Rosenstein, R., Nerz, C. & Gotz, F. Differential gene expression profiling of *Staphylococcus aureus* cultivated under biofilm and planktonic conditions. *Appl. Environ. Microbiol.* **71**, 2663–2676 (2005).
- Yao, Y., Sturdevant, D. E. & Otto, M. Genomewide analysis of gene expression in *Staphylococcus epidermidis* biofilms: insights into the pathophysiology of *S. epidermidis* biofilms and the role of phenol-soluble modulins in formation of biofilms. *J. Infect. Dis.* **191**, 289–298 (2005).
- Shemesh, M., Tam, A. & Steinberg, D. Differential gene expression profiling of *Streptococcus mutans* cultured under biofilm and planktonic conditions. *Microbiology* **153**, 1307–1317 (2007).
- Yadav, M. K. et al. Gene expression profile of early in vitro biofilms of *streptococcus pneumoniae*. *Microbiol. Immunol.* **56**, 621–629 (2012).
- Punta, M. et al. The Pfam protein families database. *Nucleic Acids Res.* **40**, D290–D301 (2012).
- Nancy, Y. Y. et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608–1615 (2010).
- Franceschini, A. et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* **41**, D808–D815 (2013).
- Cline, M. S. et al. Integration of biological networks and gene expression data using cytoscape. *Nat. Protoc.* **2**, 2366–2382 (2007).
- Schoen, C., Kischkies, L., Elias, J. & Ampattu, B. J. Metabolism and virulence in *neisseria meningitidis*. *Front. Cell. Infect. Microbiol.* **4**, eCollection 2014 (2014).
- Lévesque, C. M. et al. Involvement of sortase anchoring of cell wall proteins in biofilm formation by *streptococcus mutans*. *Infect. Immun.* **73**, 3773–3777 (2005).
- Frese, S. A. et al. Molecular characterization of host-specific biofilm formation in a vertebrate gut symbiont. *PLoS Genet.* **9**, e1004057 (2013).
- Kostakioti, M., Hadjifrangiskou, M. & Hultgren, S. J. Bacterial biofilms: development, dispersal, and therapeutic strategies in the dawn of the postantibiotic era. *Cold Spring Harb. Perspect. Med.* **3**, a010306 (2013).
- Yeoman, C. J. et al. Comparative genomics of *Gardnerella vaginalis* strains reveals substantial differences in metabolic and virulence potential. *PLoS one* **5**, e12411 (2010).
- Tomich, M., Fine, D. H. & Figurski, D. H. The TadV protein of *Actinobacillus actinomycetemcomitans* is a novel aspartic acid prepilin peptidase required for maturation of the Flp1 pilin and TadE and TadF pseudopilins. *J. Bacteriol.* **188**, 6899–6914 (2006).
- Harwich, M. et al. Drawing the line between commensal and pathogenic *Gardnerella vaginalis* through genome analysis and virulence studies. *BMC Genomics* **11**, 375 (2010).
- Lindahl, G., Stalhammar-Carlemalm, M. & Areschoug, T. Surface proteins of *Streptococcus agalactiae* and related proteins in other bacterial pathogens. *Clin. Microbiol. Rev.* **18**, 102–127 (2005).
- Castro, J. et al. Using an in-vitro biofilm model to assess the virulence potential of bacterial vaginosis or non-bacterial vaginosis *Gardnerella vaginalis* isolates. *Sci. Rep.* **5**, 1164 (2015).
- Gardner, H. L. & Dukes, C. D. *Haemophilus vaginalis* vaginitis: a newly defined specific infection previously classified non-specific vaginitis. *Am. J. Obstet. Gynecol.* **69**, 962–976 (1955).
- Kusters, J. G., Reuland, E. A., Bouter, S., Koenig, P. & Dorigo-Zetsma, J. W. A multiplex real-time PCR assay for routine diagnosis of bacterial vaginosis. *Eur. J. Clin. Microbiol. Infect. Dis.* **34**, 1779–1785 (2015).
- Mikkelsen, H., Duck, Z., Lilley, K. S. & Welch, M. Interrelationships between colonies, biofilms, and planktonic cells of *Pseudomonas aeruginosa*. *J. Bacteriol.* **189**, 2411–2416 (2007).
- Nagayama, K. et al. Role of ABC transporter proteins in stress responses of *Streptococcus mutans*. *Oral Health Dent. Manag.* **13**, 359–365 (2014).
- El Zoeiby, A., Sanschagrin, F., Havugimana, P. C., Garnier, A. & Levesque, R. C. In vitro reconstruction of the biosynthetic pathway of peptidoglycan cytoplasmic precursor in *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett.* **201**, 229–235 (2001).
- Cauci, S. et al. Pore-forming and haemolytic properties of the *Gardnerella vaginalis* cytolysin. *Mol. Microbiol.* **9**, 1143–1155 (1993).
- Gelber, S. E., Aguilar, J. L., Lewis, K. L. & Ratner, A. J. Functional and phylogenetic characterization of Vaginolysin, the human-specific cytolysin from *Gardnerella vaginalis*. *J. Bacteriol.* **190**, 3896–3903 (2008).
- Soto, S. M. Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. *Virulence* **4**, 223–229 (2013).
- Hancock, R. E. Resistance mechanisms in *Pseudomonas aeruginosa* and other nonfermentative gram-negative bacteria. *Clin. Infect. Dis.* **27**, S93–S99 (1998).
- Lynch, S. V. et al. Role of the *rapA* gene in controlling antibiotic resistance of *Escherichia coli* biofilms. *Antimicrob. Agents Chemother.* **51**, 3650–3658 (2007).
- Ramage, G., Bachmann, S., Patterson, T. F., Wickes, B. L. & Lopez-Ribot, J. L. Investigation of multidrug efflux pumps in relation to fluconazole resistance in *Candida albicans* biofilms. *J. Antimicrob. Chemother.* **49**, 973–980 (2002).
- Patterson, J. L., Stull-Lane, A., Girerd, P. H. & Jefferson, K. K. Analysis of adherence, biofilm formation and cytotoxicity suggests a greater virulence potential of

- Gardnerella vaginalis* relative to other bacterial-vaginosis-associated anaerobes. *Microbiology* **156**, 392–399 (2010).
39. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
  40. Kal, A. J. *et al.* Dynamics of gene expression revealed by comparison of serial analysis of gene expression transcript profiles from yeast grown on two different carbon sources. *Mol. Biol. Cell* **10**, 1859–1872 (1999).
  41. Pawitan, Y., Michiels, S., Koscielny, S., Gusnanto, A. & Ploner, A. False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* **21**, 3017–3024 (2005).
  42. Oliveros, J. C. VENNY. An interactive tool for comparing lists with Venn Diagrams, available at <http://bioinfogp.cnb.csic.es/tools/venny/index.html> (2007).
  43. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* **25**, 25–29 (2000).
  44. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**, D277–280 (2004).
  45. Pundir, S., Martin, M. J. & O'Donovan, C. UniProt tools. *Curr. Protoc. Bioinformatics* **53**, 1.29.21–21.29.15 (2016).
  46. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
  47. Bader, G. D. & Hogue, C. W. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4**, 2 (2003).
  48. Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **132**, 365 (2000).
  49. França, A., Freitas, A. I., Henriques, A. F. & Cerca, N. Optimizing a qPCR Gene expression quantification assay for *S. epidermidis* biofilms: a comparison between commercial kits and a customized protocol. *PLoS one* **7**, e37480 (2012).
  50. Pfaffl, M. W. in *AZ of Quantitative PCR* (ed Bustin S. A., Series 5) Ch. 3 (International University Line, 2004).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies the paper on the *npj Biofilms and Microbiomes* website (doi:10.1038/s41522-017-0012-7).