International Journal for Parasitology 43 (2013) 707–719

Contents lists available at SciVerse ScienceDirect

International Journal for Parasitology

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

Deep sequencing of *Trichomonas vaginalis* during the early infection of vaginal epithelial cells and amoeboid transition ☆

Sven B. Gould [a,*], Christian Woehle [a], Gary Kusdian [a], Giddy Landan [a], Jan Tachezy [b], Verena Zimorski [a], William F. Martin [a]

[a] Institute for Molecular Evolution, Heinrich-Heine-University, 40225 Düsseldorf, Germany
[b] Laboratory of Molecular and Biochemical Parasitology, Charles University Prague, 12844 Prague, Czech Republic

ARTICLE INFO

Article history:
Received 4 December 2012
Received in revised form 8 April 2013
Accepted 9 April 2013
Available online 18 May 2013

Keywords:
Trichomonas
Infection
Cytoskeleton
Oxygen stress
Gene families

ABSTRACT

The human pathogen *Trichomonas vaginalis* has the largest protozoan genome known, potentially encoding approximately 60,000 proteins. To what degree these genes are expressed is not well known and only a few key transcription factors and promoter domains have been identified. To shed light on the expression capacity of the parasite and transcriptional regulation during phase transitions, we deep sequenced the transcriptomes of the protozoan during two environmental stimuli of the early infection process: exposure to oxygen and contact with vaginal epithelial cells. Eleven 3′ fragment libraries from different time points after exposure to oxygen only and in combination with human tissue were sequenced, generating more than 150 million reads which mapped onto 33,157 protein coding genes in total and a core set of more than 20,000 genes represented within all libraries. The data uncover gene family expression regulation in this parasite and give evidence for a concentrated response to the individual stimuli. Oxygen stress primarily reveals the parasite's strategies to deal with oxygen radicals. The exposure of oxygen-adapted parasites to human epithelial cells primarily induces cytoskeletal rearrangement and proliferation, reflecting the rapid morphological transition from spindle shaped flagellates to tissue-feeding and actively dividing amoeboids.

1. Introduction

The extracellular parasite *Trichomonas vaginalis* infects the urogenital tract of approximately 3% of the world population annually and is thus the most widespread non-viral, sexually transmitted human parasite known (Schwebke et al., 2011). Although the vast majority of *T. vaginalis* infections proceed without apparent symptoms, infection with the parasite decreases fertility, elevates the risk of prostate and cervical cancer and increases the risk of acquiring HIV (Petrin et al., 1998; Ryan et al., 2011). The more severe infections with manifest symptoms, Trichomoniasis, result in urogenital tract swelling and inflammatory discharge and are commonly treated with nitroimidazole derivates, although resistant strains are on the rise (Kulda, 1999; Upcroft and Upcroft, 2001; Benchimol, 2008; Pal and Bandyopadhyay, 2012).

A crucial step of the infection process involves a dramatic morphological shift of the parasite. The free-swimming ovoid cells, which resemble the familiar image of a flagellated protozoan, transform into an amoeboid form (Fig. 1) upon contact with the urogenital tract. This process commences more or less immediately upon contact with host tissue and the transformation of a cell takes only minutes to complete (Lal et al., 2006). The parasite's morphogenesis entails at least two distinct but simultaneous processes: (i) the dramatic shape transition and (ii) the adherence to host cells.

Recent studies on *Trichomonas* surface proteins that mediate host cell adherence and interactions with host extracellular matrix have identified three classes of proteins (recently reviewed by Hirt et al. (2011) and Ryan et al. (2011)). First there is the large BspA family (*Bacteroides forsythus* surface protein A), members of which have been localised to the parasite's surface, but whose exact function remains elusive (de Miguel et al., 2010; Noël et al., 2010). The second class comprises components of the thick glycocalyx, which is thought to also directly interact with the host extracellular matrix. Human Galectin-1 was identified as an interaction partner for a specific single lipophosphoglycan of the pathogen (Okumura et al., 2008). A third, and most controversial, class of suggested adhesion proteins is those that have been suggested to have dual functions, as they include enzymes of the glycolytic pathway such as glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and

☆ *Note:* Nucleotide sequence data reported in this paper are available through the NCBI Single Read Archive Accession No. SRA059159.

* Corresponding author. Address: Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany. Tel.: +49 2118113983; fax: +49 2118113554.
E-mail address: gould@hhu.de (S.B. Gould).

0020-7519/$36.00
http://dx.doi.org/10.1016/j.ijpara.2013.04.002

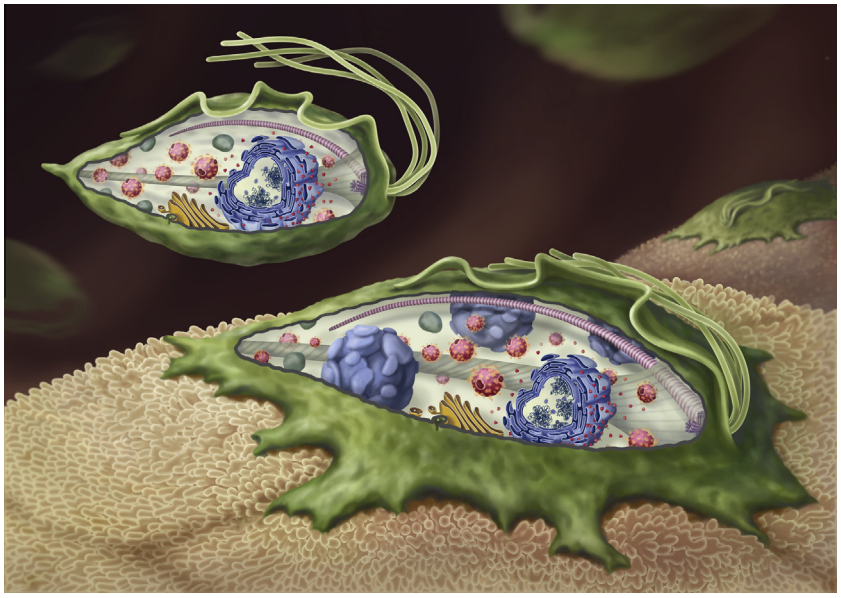


Fig. 1. The two major phenotypes of *Trichomonas vaginalis*. The human pathogen can switch between two radically different phenotypes. Phenotypic plasticity is induced through contact with tissue of the urogenital tract (light brown). After initial contact the flagellated cells become amoeboid within minutes and increase the surface area interacting with and scavenging from host tissue. Many of the amoeboid cells become multinuclear (nuclei in blue, with closely associated endoplasmic reticulum), proliferate and subsequently divide.

enzymes of pyruvate metabolism such as malic enzyme (ME), pyruvate:ferredoxin oxidoreductase (PFO) and subunits of succinyl-CoA synthetase (SCS) (Garcia and Alderete, 2007; Meza-Cervantez et al., 2011). These are common proteins with clear functions in energy metabolism (Müller et al., 2012). Their involvement in adhesion has been challenged (Addis et al., 1998; Brugerolle et al., 2000; Hirt et al., 2007; Ryan et al., 2011) and evidence for their secretion or their direct interactions with the human extracellular matrix is yet to be presented. The machinery behind trichomonad morphogenesis is not well characterised but microscopic observations and the localisation of actin and actin-associated proteins have demonstrated these cytoskeletal elements to play a crucial role (Gold and Ofek, 1992; Brugerolle et al., 1996; Bricheux et al., 1998, 2000; Kusdian et al., in press).

Trichomonas is considered a typical anaerobe. It has no oxygen requirements for growth but sometimes is designated as a microaerophile because it grows slightly better at very low oxygen tensions of ~0.25 μM, corresponding to about 1/1000th of ambient levels at 25 °C, or 250 μM (Paget and Lloyd, 1990), while oxygen tension above 60 μM has a detrimental effect (Ellis et al., 1994). The mitochondria of *Trichomonas* are designated as hydrogenosomes because they produce hydrogen as an end product of an anaerobic, fermentative energy metabolism (Müller, 1988; Müller et al., 2012). Core enzymes of the hydrogenosomal metabolism, such as PFO and Fe–Fe hydrogenases, are highly oxygen-sensitive (Williams et al., 1990; Hrdý and Müller, 1995; Page-Sharp et al., 1996). However, the parasite typically experiences oxygen stress in its natural environment, for example during the transmission from one host to the other or with fluctuating vaginal oxygen levels during the menstruation cycle (Wagner and Levin, 1978; Ellis et al., 1992; Hill et al., 2005) and hence must possess mechanisms to avoid inactivation of oxygen-sensitive enzymes and to remove reactive oxygen species (ROS). Cytosolic NADH and NADPH oxidases are among the parasite's most important oxygen scavenging enzymes (Linstead and Bradley, 1988). Glutathione, a widespread antioxidant among eukaryotes, is absent in *T. vaginalis*, with cysteine possibly acting in its place (Ellis et al., 1994). Individual proteins shown to be up-regulated during oxygen stress include superoxide dismutase (SOD) (Ellis et al., 1994; Rasoloson et al., 2001) and peroxiredoxins (Coombs et al., 2004), ubiquitous enzymes that are thought to be central to defenses against ROS (Gretes et al., 2012). Thioredoxin reductases are also present in *T. vaginalis* as a component of the hydrogenosomal thioredoxin-linked antioxidant system (Mentel et al., 2008). More recently, flavodiiron class A protein was shown to function as the major oxygen reductase responsible for respiration of hydrogenosomes (Smutná et al., 2009). This enzyme reduces oxygen to water in a four-electron reaction, while production of hydrogen peroxide was not detected. Thus, the parasite seems to utilise a sophisticated system to buffer oxygen stress, which includes a cytosolic and a hydrogenosomal antioxidant system in combination (Coombs et al., 2004; Mentel et al., 2008).

With approximately 160 Mbp, the genome of this protist is the largest of any protozoan genome currently available (Carlton et al., 2007). The original genome annotation identified approximately 60,000 potential protein encoding genes, which was reduced to approximately 46,000 through subtracting genes that appear to be only fragments of others (Smith and Johnson, 2011). A large portion of this redundancy of annotated genes is due to an unknown amount of duplications the genome has experienced. Approximately 65% of the genome sequence consists of repetitive elements and there are currently still >25,000 scaffolds that cannot be assembled. To what degree the parasite expresses this large arsenal of encoded genes is not known; expression evidence exists only for 10,470 genes at TrichDB (http://trichdb.org/trichdb/). Transcriptional regulation in *T. vaginalis* and in protists in general is not well characterised. Three core promoter elements have been identified in *T. vaginalis*, the initiator element (Inr) and the M3 and M5 elements (Schumacher et al., 2003; Smith et al., 2011), and one which appears to be the reverse sequence of the M3 motif and linked to the 5′ untranslated region (UTR) of histone encoding genes (Cong et al., 2010). Two additional motifs, M2 and M4, were not further characterised as the M2 showed no conserved localisation in comparison with the start codon and the M4 was in too close proximity to it (Smith et al., 2011).

In order to better understand gene regulatory mechanisms underlying the initial steps of human vaginal epithelial cell (VEC) infection by *T. vaginalis* and to identify potential core changes,

we investigated 11 transcriptomes of *T. vaginalis* under varying conditions and at different time points. To be able to distinguish between the responses to oxygen itself and contact with VECs, which grow at 15% oxygen (approximately 200 μM), we sequenced 3′ fragment libraries of anaerobically cultured cells (i) after exposure to oxygen alone, (ii) after exposure to oxygen in combination with VECs and (iii) after exposure of oxygen-adapted *Trichomonas* cells to VECs, which identified transcriptional responses specific to contact with host tissue.

2. Materials and methods

2.1. Cultures

Trichomonas vaginalis T016 was cultured in tryptone-yeast extract maltose (TYM)-medium (2.22% (w/v) tryptose, 1.11% (w/v) yeast extract, 15 mM maltose, 9.16 mM L-cysteine, 1.25 mM L(+)ascorbic acid, 0.77 mM KH_2PO_4, 3.86 mM K_2HPO_4, 10% (v/v) horse serum, 0.71% (v/v) iron solution (1% (w/v) $Fe(NH_4)_2(SO_4) \times 6H_2O$, 0.1% (w/v) 5-sulfosalicylacid)) in 15 ml tubes in the absence of oxygen at 37 °C. To prevent bacterial contamination a penicillin/streptomycin mix was added to a final concentration of 100 μg/ml to the media. Oxygen-adapted *T. vaginalis* were also grown in TYM medium but in cell culture flasks identical to those used for the human VEC line, MS-74 (see below), at 37 °C, 5% CO_2 and 15% O_2 in a Galaxy 48R (Eppendorf, Germany). At a cell density of approximately 9×10^7 cells/ml, 1 ml was transferred into 12 ml of fresh TYM medium for continuous culturing. 'Immortalised' VECs (MS-74) were cultured in 45% DMEM (Invitrogen, Germany), 45% Keratinocyte-SFM (Invitrogen) and 10% FCS in cell culture flasks with vented lids (75 cm^2, VWR, Germany) at 37 °C, 5% CO_2 and 15% O_2 in a Galaxy 48R (Eppendorf). For passaging, cells were washed twice with Dulbecco's PBS (PAA, Germany), digested with trypsin (Invitrogen) for 10 min at 37 °C, before inactivation with FCS. Cells were centrifuged and resuspended in fresh media and split 1:10 into new flasks and medium.

2.2. RNA isolation

For the oxygen stress assay (*AnOx*), 50 ml of *T. vaginalis* T016 grown in sealed tubes were transferred into cell culture flasks with vented lids and incubated at 37 °C, 5% CO_2 and 15% O_2. RNA was isolated using TRIzol® and following the manufacturer's protocol (Invitrogen) at four time points (0, 5, 30 and 120 min of oxygen exposure), whereas the RNA isolated at 0 min served as a baseline. For the oxygen stress and infection assay (*AnOxInf*) 50 ml of T016, grown identically to those used for *AnOx*, were transferred onto a monolayer of MS74 cells grown in a cell culture flask (75 cm^2) with a vented lid. Cells were harvested with a cell scraper (BD Falcon, Germany) and cell pellets produced by centrifugation at 3,500*g* for 1 min at 8 °C. RNA was then again isolated using TRIzol® at three time points (5, 30 and 120 min after infection). For the infection assay (*AdInf*), 50 ml of oxygen-adapted T016 (T016 cultured for at least five continuous days in the CO_2 incubator at 15% O_2) were exposed to a monolayer of VECs, then RNA was isolated at 5, 30 and 120 min after exposure. All experiments were performed twice (biological duplicates) and RNA duplicates pooled before DNase treatment (Fermentas, Germany) to avoid DNA contamination. Frozen RNA was then sent to Eurofins MWG (Ebersberg, Germany) for 3′ fragment library preparation and deep sequencing. For quantitative real time PCR the same RNA samples that were used for 3′ transcript library sequencing, were used for these experiments. RNA was transcribed into cDNA using the iScript™ cDNA Synthesis Kit (Bio-Rad, Germany), and served as a template. Primers are listed in Supplementary Table S1. All experiments were carried out in technical triplicates using a StepOnePlus™ and Power SYBR® Green master mix (Applied Biosystems, Germany).

2.3. Transcriptome analyses

2.3.1. Reference genomes

Genomic scaffolds of *T. vaginalis*, sequences of annotated genes and genomic features were downloaded from TrichDB V1.3 (Aurrecoechea et al., 2009), including the information as to whether they encode a signal peptide or transmembrane domains (TrichDB category "Protein Features"). The nuclear genomic sequences of *Homo sapiens* (version GRCh37.p5) and its mitochondrial genome were downloaded from the National Center for Bioinformatics Information (NCBI) website (http://www.ncbi.nlm.nih.gov). Genomic scaffolds of *Trichomonas* smaller than 1,000 nucleotides and repeated genes were discarded (Carlton et al., 2007). Genomic sequences were reduced to the predicted mRNA locations plus 100 bp of downstream sequence to cover potential 3′ UTRs. Genes labeled as "hypothetical" were manually annotated using RefSeq (as of August 2012; Pruitt et al., 2007) and NCBI BLAST (Altschul et al., 1997).

2.4. Transcriptome sequences

A total of 153,137,205 reads of a typical length of 100 bp for 11 different combinations of growth conditions and time points were obtained. The reads were deposited at the NCBI Single Read Archive (accession SRA059159). Each read was mapped onto the genomic sequences of *H. sapiens* and *T. vaginalis* using a pipeline consisting of Bowtie2 (2.0.0-beta6; Langmead and Salzberg, 2012), SAMtools (Li et al., 2009) and BEDTools (Quinlan and Hall, 2010). Non-coding sequence and RNA genes were discarded.

For comparisons of expression levels among the 11 samples, the raw read counts were normalised by the total number of reads for each sample, and scaled to match the total hit count from sample *An0*. It was noted that for expression profiling using a 3′ fragment cDNA library such as ours, there is no need to normalise for transcript length ('FPKM'; fragments per kb per million sequenced reads). To cluster highly similar genes (Table 1, bottom row), transcript sequences were aligned pairwise using NEEDLE (EMBOSS; Rice et al., 2000) and MCL (Enright et al., 2002) was used to generate clusters with 90% identity. Relative up- and down-regulation values for individual genes were calculated as the ratio of the scaled read counts of the gene at the time point and in the reference sample ("*Ad0*" for "*AdInf*", and "*An0*" for "*AnOx*" and "*AnOxInf*"). *P* values and false discovery rates (FDRs), as deposited in Supplementary Tables S2–S4, were calculated using the edgeR package (Robinson et al., 2010). Due to the lack of technical replicates for the individual time points, a fixed biological coefficient of variation (BCV) of 0.1 and a significance level of 5% for the FDR were used as an indicator for differential expression (www.bioconductor.org: edgeR user guide chapter 2.9, option 2).

Identified protein coding sequences were functionally classified using the clusters of orthologous groups (COG) database (Tatusov et al., 2003), using BLASTP (*E*-value cutoff 10^{-10}). Word clouds were generated using wordle.net and advanced settings. Protein tags were assigned manually by inspecting gene annotations with a known function or domain for the top 100 genes (see Supplementary Tables S2–S4).

For the identification of promoter sequences the 60 bp upstream regions from the ATG start codon of expressed genes were screened for the previously identified sequence motifs (Smith et al., 2011). The Inr, M3 and M5 motifs were based on the description in the original text, while the M2 and M4 were extracted from the sequence logo in a figure.

Table 1
Overview of the *Trichomonas vaginalis* transcriptomes analyzed. Eleven transcriptomes of the *T. vaginalis* strain T016 were sequenced in total. The transcriptomes of one culture grown under anaerobic conditions (*An0*) and one, in which the cells had been adapted to 15% CO_2 (*Ad0*) served as the base transcriptomes (time points 0). Three individual conditions were induced: oxygen stress (*AnOx*), oxygen stress and exposure to vaginal epithelial cells (*VECs*, *AnOxInf*) and exposure of oxygen-adapted T016 to VECs (*AdInf*). RNA was isolated at 5, 30 and 120 min after the exposure to the individual conditions.

Condition	Anaerobic	Oxygen Stress			Oxygen Stress & Infection of VECs			15% O_2	Infection of VECs		
Time point (min)	0	5	30	120	5	30	120	0	5	30	120
Name of exp. set	*An0*	*AnOx5*	*AnOx30*	*AnOx120*	*AnOxInf5*	*AnOxInf30*	*AnOxInf120*	*Ad0*	*AdInf5*	*AdInf30*	*AdInf120*
Total Reads	15,483,280.00	12,513,182.00	12,311,601.00	10,395,563.00	14,752,720.00	15,050,307.00	16,320,707.00	18,000,848.00	13,613,335.00	11,296,593.00	13,399,069.00
% Mapped[a]	99.29	99.18	99.20	98.88	98.89	99.11	98.86	97.42	98.36	98.14	98.66
Protein Hits[b]	25,109.00	24,910.00	24,949.00	24,161.00	26,098.00	26,053.00	26,625.00	29,294.00	24,319.00	23,891.00	25,852.00
⩾100 hits[c]	8,056.00	8,978.00	7,889.00	7,808.00	11,083.00	10,215.00	11,092.00	14,416.00	5,043.00	5,372.00	6,878.00
Clustered 90%[d]	23,731.00	23,605.00	23,597.00	22,922.00	24,626.00	24,570.00	24,936.00	26,339.00	22,737.00	22,163.00	23,764.00

[a] Percentage of reads that uniquely mapped onto the *T. vaginalis* genome.
[b] Fraction of protein coding genes identified.identity ("potential paralogs").
[c] Fraction of protein coding genes identified with at least 100 hits.
[d] Fraction of protein coding genes with hits after clustering those with 90% sequence.

Gene families are based on the orthologous groups deposited at TrichDB. Gene pairwise co-expression was calculated as the Pearson correlation of the within sample read count ranks. Family-wise co-expression was calculated as the median of all pairwise co-expressions within the family. Nucleotide identities were derived from pairwise global alignments of annotated transcripts using the NEEDLE program (Rice et al., 2000) and the median identity for each gene family calculated.

3. Results

3.1. Transcriptional capacity

To obtain different sets of transcriptomic libraries of the parasite, we combined the infectious *T. vaginalis* strain, T016, and the human VEC line, MS74, in the following, referred to as T016 and MS74, respectively. T016 was chosen as it represents a highly virulent isolate of *T. vaginalis* (Pereira-Neves and Benchimol, 2007). In total we sequenced 3′ fragment cDNA library transcriptomes of T016 from 11 individual conditions (Table 1), and each from biological duplicates. Altogether 153,137,205 reads were obtained, of which 143,767,773 (93.9%) mapped onto protein-coding genes of *T. vaginalis*. Only 3,887,899 reads from the libraries also containing mRNA from MS74 (*AnOxInf* and *AdInf*) mapped with their best hit onto human genes. Altogether the reads mapped onto 33,157 individual protein-coding genes of the parasite, of which 20,392 genes were expressed under all conditions and 23,879 genes with a minimum of 10 mapped transcripts under any condition. As *T. vaginalis* encodes many paralogous gene copies due to the duplication of large parts of the genome (Carlton et al., 2007), the transcript sequences of identified expressed proteins with 90% global identity were also clustered, leaving 27,719 individual protein-encoding genes being expressed when considering all libraries combined (Table 1). A set of actin genes (TVAG_337240, TVAG_054030 and TVAG_485210) was overall the highest expressed in the anaerobically grown trophozoite library (*An0*), followed by genes encoding proteins of the core carbon energy metabolism, namely pyruvate-flavodoxin oxidoreductase (TVAG_198110), phosphoenolpyruvate carboxykinase (TVAG_479540), malate dehydrogenase (TVAG_204360) and glucose-6-phosphate isomerase (TVAG_061930).

In the expression patterns presented for the individual experiments below, only those genes were considered for which at least 100 transcripts were present in each compared expression set after normalisation. Among the three conditions tested, statistical support for 94% of the top 200 up- and down-regulated genes was found (Supplementary Tables S2–S4). Further, quantitative real-time PCR on a set of exemplary genes such as rubrerythrin, cysteine protease or an ankyrin repeat-containing protein for example, further supported the deep-sequencing results (Supplementary Fig. S1).

3.2. Oxygen stress

After 5 min of oxygen stress (*AnOx5*), 33 genes were found to be up-regulated at least two-fold but none of them could directly be linked to the enzymatic machinery scavenging ROS. Genes up-regulated were for four transcription factors of the MYB family, other DNA- and RNA-binding proteins, 12 of unknown function and eight with homology to molecular switches and messengers such as a diverse range of kinases, calmodulin and a ubiquitin-conjugating enzyme (Supplementary Table S2). The expression pattern changed significantly after 30 min (Fig. 2) with 202 genes up-regulated at least two-fold. Although two MYB genes were still among the top 100, they represented different MYB genes from those observed

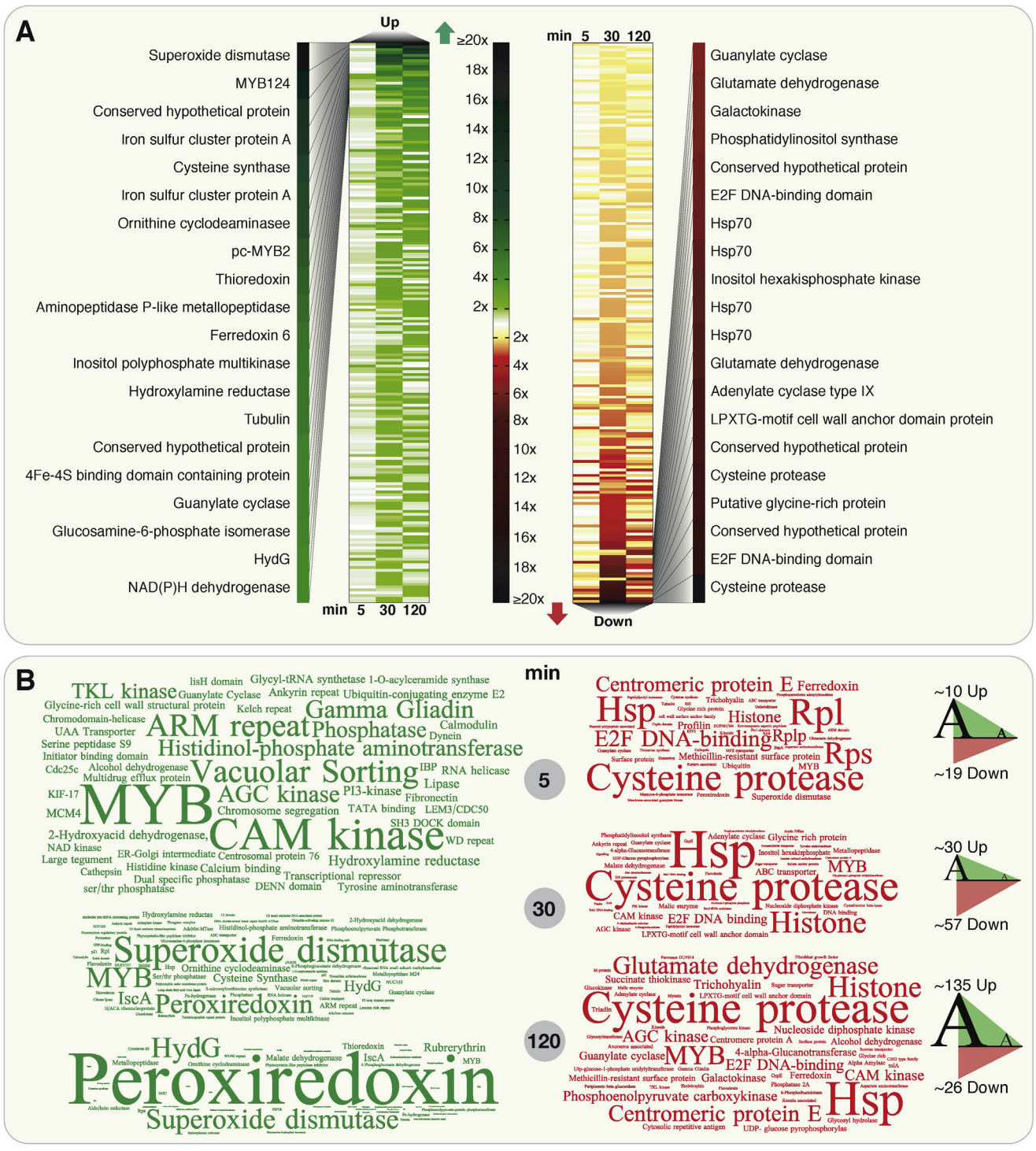


Fig. 2. The most regulated *Trichomonas vaginalis* genes during oxygen stress. (A) Distribution of the top 200 up- and down-regulated genes during oxygen stress in green and red, respectively, across all three time points (5, 30 and 120 min), with the level of transcriptional regulation indicated by the color gradient shown in the middle. The annotations of the 40 most regulated genes are given on the left and right with their individual color code presented accordingly. (B) The top 100 up- and down-regulated genes, with a predicted function or domain, represented as a word cloud. Letter height is proportional to the up- and down-regulation values, as indicated by the scale bar on the right, whereas paralogous genes (and their values) are grouped and shown as one function.

after 5 min. With an 18-fold up-regulation, a peroxiredoxin-encoding gene was overall the highest up-regulated product. The top 100 after 30 min further included many proteins known to buffer oxygen stress, including four SODs (up-regulated 13.5, 6.4, 5.5 and 4.6-fold), another member of the peroxiredoxin family (4.2-fold up), one rubrerythrin (2.6-fold up), ferredoxin6 (5.7-fold up) and thioredoxin (three-fold up). Two IscA proteins of the iron-sulfur cluster assembly machinery were also up 8.3 and 6.2-fold, which reflected the need for new iron-sulfur clusters, which are present in ferredoxin and rubrerythrin.

The trend observed after 30 min of oxygen-induced stress continued and even increased after 120 min (*AnOx120*) with 218 genes up-regulated at least two-fold (Fig. 2). The top position was held by the identical peroxiredoxin as after 30 min, with a 90-fold increase

of transcript level when compared with *An0*. Except for one ferredoxin gene, all others mentioned here experienced a further increase. These, together with another peroxiredoxin, two thioredoxin and two rubrerythrin genes that were additionally up-regulated, totalled more than a dozen potential radical scavengers among the top 100 genes with increased expression after 120 min of oxygen stress (Supplementary Table S2). Two Nfu genes, just as the two IscA genes, potentially part of the machinery providing iron-sulfur clusters, were also among the top 100 after 120 min of oxygen stress. HydE and HydG, two of the three known iron-hydrogenase maturases (Pütz et al., 2006), were also found to be up-regulated. The overall increase in the transcript level compared with *An0* was higher after 120 min then after 30 min of oxygen stress and six of the nine genes that were up-regulated by more than 10 times after 120 min, were involved in the defense of ROS or in the supply of the proteins' active centers: peroxiredoxins, SODs, rubrerythrin and enzymes of the iron-sulfur cluster assembly machinery.

The down-regulation of genes under oxygen stress was less strong and peaked at 30 min by a gene encoding a cysteine protease that was down-regulated 21-fold, and which was displaced by a MYB transcription factor after 120 min, down nearly nine-fold. Next to the cysteine protease-encoding genes, another prominent gene family that was down-regulated included those encoding heat shock proteins (Hsps), in particular those of the Hsp70 family (Fig. 2). In general it was noteworthy in comparison to the up-regulation of genes, which still slightly increased after 2 h, that the down-regulation effect subsided after the 1 h.

3.3. Exposure to human epithelial cells

To identify major changes in the transcriptome of *T. vaginalis* upon contact with human vaginal epithelial tissue, and be able to separate them from the changes induced through oxygen stress alone, two additional experiments were performed. In the first experiment, anaerobically grown parasites, identical to those used for the oxygen-stress only experiment, were exposed to a monolayer of MS74 (*AnOxInf*). In the second experiment the parasite was first adapted to oxygen by growing the cells for 5 days in the same CO_2 incubator as the human cell line, before exposing them to MS74 (*AdInf*). The infection of MS74 through T016 was monitored through light microscopy and in both experimental sets, adhesion to human cells was observed to occur instantly, i.e. a large proportion of the VECs presented with an adherent parasite after 5 min. Again transcriptomes of parasites were sequenced at 5, 30 and 120 min time points.

In the AnOxInf experiments, in which MS74 was infected with anaerobically grown parasites, an overall similar pattern of up-regulated genes was observed among the Eukaryotic Orthologous Groups (KOG) categories as in the oxygen-stress experiment (Fig. 3). At all three time points the top positions were held by the same genes as in the oxygen-only experiments, a MYB transcription factor (TVAG_076270) after 5 min, and a peroxiredoxin (TVAG_455310) after 30 and 120 min. Many other genes shared identical positions among the set of genes with increased expression, and the intensity of regulation was comparable. Generally the up-regulated factors were again dominated by proteins involved in dealing with oxygen stress, also including the enzymes of the iron-sulfur cluster generating machinery (Supplementary Table S3). Among the top 200 results, 31.5% of the genes whose expression was increased, were identical between just oxygen stress and oxygen stress combined with infection, and 40.1% of the down-regulated genes (Fig. 5B). In comparison, the AdOxInf set shared only 0.7% among the up-regulated genes with the AnInf set (see below), and 9.2% of the down-regulated genes. One obvious difference was observed among the down-regulated genes: a noticeable amount of genes associated with the core-carbon metabolism, in particular ME, malate dehydrogenase and fruc-

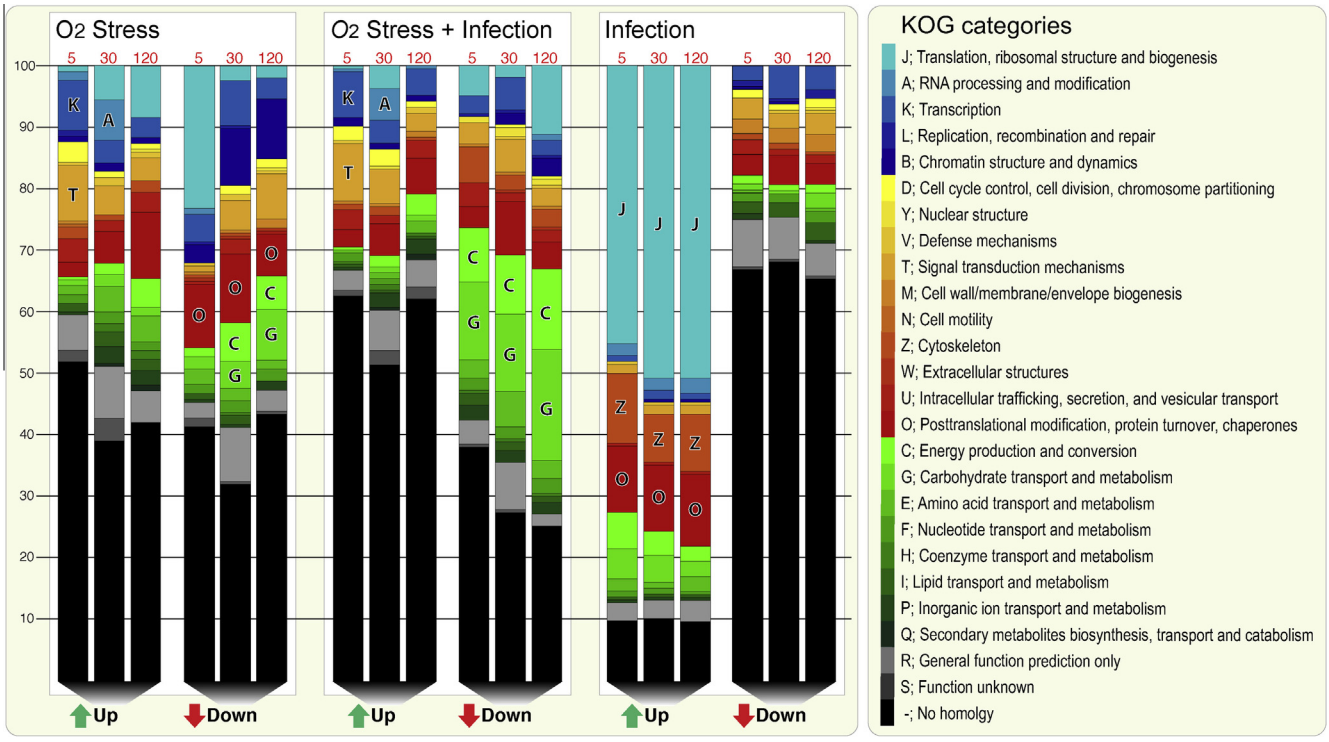


Fig. 3. Transcriptional shift of eukaryotic orthologous group categories. Regulated genes were mapped to the eukaryotic orthologous group database and each gene assigned to one of the 25 categories according to the best Basic Local Alignment Search Tool (BLAST) hit. In blue, information storage and processing; in yellow to red, cellular processes and signaling; in green, metabolism; and in grey to black, general or unknown function. Categories of particular interest are highlighted by the corresponding letter of the individual eukaryotic orthologous group category (shown on the right).

tose-bisphosphate aldolase, are increasingly down-regulated over the time of 2 h, an effect that does not occur to this degree during oxygen stress alone (Supplementary Table S3).

The transcriptome of parasites adapted to the growth conditions of the CO_2 incubator and then exposed to MS74 differed significantly at all time points from both other experimental sets, which immediately becomes apparent from the KOG categories that were up-regulated (Fig. 3 and Supplementary Table S4). After 30 and 120 min, approximately half of the top 200 genes with increased expression belong to category J (translation, ribosomal structure and biogenesis) and many of them encode ribosomal subunits, elongation factors and tRNA synthetases. Other gene families noticeably up-regulated included many encoding cysteine-like proteases, actin and the actin-binding protein profilin

Fig. 4. The top regulated genes of oxygen adapted *Trichomonas vaginalis* after exposure to human vaginal epithelial cells. (A) Distribution of the top 200 up- and down-regulated genes during oxygen stress in green and red, respectively, across all three time points (5, 30 and 120 min). The annotations of the top 40 most regulated genes are given on the left and right. (B) The word cloud representation shows the majority of up-regulated genes during infection is rather stable over the three time points (see also Fig. 5A) and guided by genes encoding actin, cysteine proteases and ribosomal proteins. For details refer to Fig. 2.

(Fig. 4). Noteworthy, the set of up-regulated genes during the first 2 h of the infection was extremely stable in comparison with the other experimental sets (Fig 5A), i.e. across all three time points approximately 200 of the genes were identical.

Next the *AdInf* data set was screened for gene families previously suggested to play a role when infection is established. For example, the large family of BspA-like proteins are thought to be important for various aspects of *T. vaginalis* pathobiology at the host-pathogen interface (Noël et al., 2010). We found expression evidence for 721 of the reported 911 putative BspA genes, albeit the vast majority was represented by a low number of mapped transcripts, with a median of 3.7 and an average of 37.2 reads/gene (max: 5446, min: 1) and compared with a general median of 19 and an average of 457 reads/gene when considering the entire data in this study. More BspA genes were up-regulated during oxygen stress (*AnOx*), and in particular when combined with exposure to VECs (*AnOxInf*), then when oxygen-adapted parasites were exposed to VECs (*AdInf*). In the latter set the vast majority of BspA genes were observed to be down-regulated and only seven were up-regulated more than two-fold (Supplementary Table S5). We searched for further potential host-interacting and surface bound proteins of *Trichomonas* by screening for up-regulated proteins with signal peptides or a single transmembrane domain. These included previously suggested and potentially important protein families such as the leishmanolysin-like, legume-like or polymorphic membrane protein-like proteins (Carlton et al., 2007). However, apart from a few individual cases (e.g. TVAG_265530; saposin, TVAG_388060; a polymorphic membrane protein, TVAG_140850) no striking

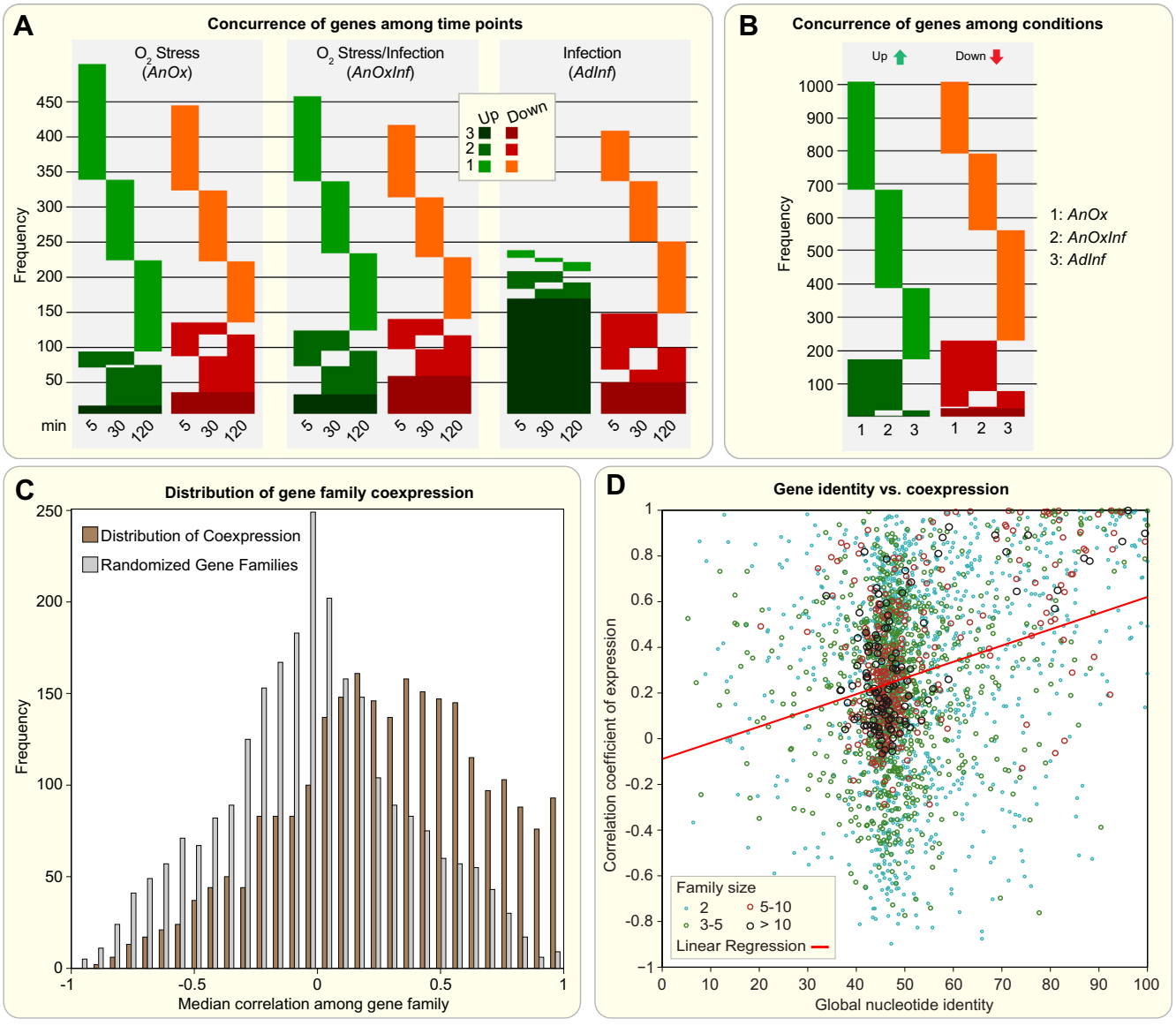
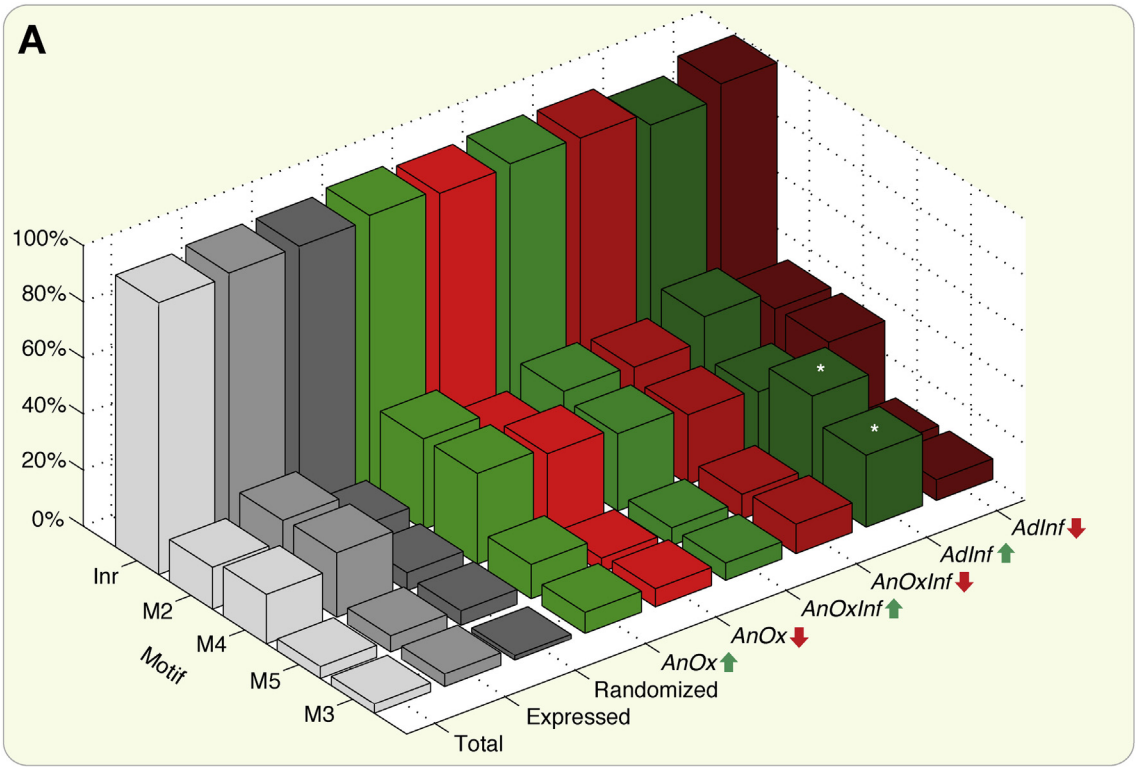


Fig. 5. Concurrence and correlated expression of *Trichomonas vaginalis* genes and gene families. (A) Concurrence of regulated genes among the individual expression sets is indicated by a color gradient. If a gene is up-regulated within only one time point this is indicated by light green; if a gene is up-regulated at all three time points, it is indicated by dark green. The highest amount of concurrence was observed among the up-regulated genes induced through the infection of MS74 with oxygen-adapted parasites (*AdInf*). (B) Co-occurrence of identical genes among the three conditions (*AnOx*, *AnOxInf*, *AdInf*), which demonstrates that oxygen alone (*AnOx*) and in combination with vaginal epithelial cells (*AnOxInf*) leads to the identical up- and down-regulation of approximately 200 genes. They differ significantly from the transcriptomes of the parasites adapted to oxygen before their exposure to vaginal epithelial cells (*AdInf*). (C) Displayed are the median expression correlations of gene family members (brown) and as a reference compared with a randomized set of gene families (grey). Expression levels of members of the same gene family correlate to a higher degree than expected by chance (Kolmogorov–Smirnow test *P* value <0.05). (D) Median co-expression per gene family plotted against median nucleotide identity. The linear regression line is shown in red ($r = \sim 0.25$; *P* value <0.05).

up-regulations were detected, especially not for entire families (Supplementary Table S6). In contrast, a protein of the tetraspanin family, TvTSP6, involved in parasite migration and potential sensor reception (de Miguel et al., 2012), was up-regulated in all of our infection libraries, as well as some hydrogenosomal proteins and proteins of the glycolytic pathway, which were reported to have a moonlighting function and serve as potential adhesion proteins, too, as mentioned in the Introduction (see also Supplementary Table S7). A surface proteome analysis revealed that 11 proteins were found more commonly present on the surface of more adherent strains (de Miguel et al., 2010). All of these genes were found expressed in T016 (a strain that was not part of the proteome analysis), with TVAG_239650 (unknown function) having the highest absolute expression. The transcriptomic response of the 11 genes is mixed and to a certain degree mirrors that of the BspA family. Genes up- or down-regulated during infection of VECs with parasites not adapted to oxygen (*AnOxInf*) did not necessarily match those up- or down-regulated during infection of VECs with oxygen-adapted parasites (Supplementary Table S8). It is noteworthy that oxygen stress alone (*AnOx*) had very little to no effect, whereas the presence of VECs in both experimental sets (*AnOxInf* and *AdInf*) induced more significant transcriptional changes.

3.4. Regulation and co-expression of gene families

One interesting aspect of trichomonad biology is the massive expansion of gene families, whose coordinated expression has not been examined hitherto in detail due to the lack of available sequencing data. The correlation of expression among 2,509 gene families was analysed, for which we found at least two members to be identified by 100 mapped reads. A higher correlation of regulation between the members of gene families was observed when compared with randomised samples (Fig. 5C), but for many – including important gene families – the patterns were non-uniform. Whereas in some families a considerable correlation between the regulations of expression was observed, in others individual copies could be very differentially expressed. For example, for ribosomal protein-encoding genes or the actin gene family, all genes are generally expressed at high levels, the co-expression coefficient was high and expression variance was low. On the



B

	Total	Expressed	Randomized	*AnOx*↑	*AnOx*↓	*AnOxInf*↑	*AnOxInf*↓	*AdInf*↑	*AdInf*↓
Frequency	59169	32973	32973	254	201	2432	238	789	2822
Inr	96.9	98.0	98.1	99.6	98.0	99.0	98.7	93.9	99.1
M2	15.0	22.2	9.5	32.3	22.4	30.3	29.4	37.9	31.6
M4	17.6	23.0	6.4	32.3	29.9	27.5	24.8	23.4	31.7
M5	4.2	5.7	5.2	12.2	5.0	6.4	8.8	34.3*	6.7
M3	3.2	4.5	1.6	7.5	6.5	6.1	10.4	25.6*	7.3

Fig. 6. *Trichomonas vaginalis* core promoter distribution. We screened the 60 bp upstream region from the start codon, for the five core promoters identified in *Trichomonas*. (A) In light grey (Total) the distribution of the promoter motifs among all genes annotated at TrichDB and medium grey (Expressed) the distribution among all genes, for which there is expression evidence. The Inr motif was found to occur at the same rate among scrambled upstream sequences (Randomized), albeit this did not take into account the distance to the start codon. The occurrence of the M3 and M5 motif is slightly increased among the genes up-regulated ⩾ two-fold during oxygen stress, but even more significantly among the *AdInf* set (asterisk). (B) Table of the values from (A) with the frequency of analyzed sequences given in the first row and the percentage of the individual motifs (shown in the second column) identified among the sets analyzed.

contrary, we found expression evidence for 30 genes of the Hsp70 family, within which one gene copy (TVAG_092490) was represented in total by 216,000 normalised reads, another copy (TVAG_130280) by only two. The same extreme difference was observed, but was not limited to, a glucosylceramidase and a synaptogamin family (Supplementary Table S9). Importantly, the expression correlation within an individual gene family was not random but was correlated with sequence identity, i.e. the more similar the sequences of two genes, the greater the chance their expression significantly correlates (Fig. 5D).

To elucidate whether certain promoter elements might be predominantly associated with one of the introduced environmental changes, we extracted the 5′ UTRs of all annotated genes, screened them for the known core promoter regions and compared them with those genes whose expression within our libraries was regulated at least two-fold and for which at least 100 reads were identified. First, to generally elucidate promoter motif significance, we randomised the 60 bp upstream regions of all expressed genes and noticed that only the Inr motif ((A/T/C)(T/C)A(A/T/C)(T/A)), as defined by Smith et al. (2011), to be as frequently encoded among the scrambled sequences as it was among true promoter regions (Fig. 6). However, the latter does not take into account the distance to the start codon of the motif and which is likely of importance.

All five promoter motifs are enriched among the >33,000 expressed genes in comparison with the approximately 60,000 annotated at TrichDB. While the Inr motif, which mediates correct transcription initiation, was present in 98% of all the upstream regions of genes expressed, the M3 was found within only 4.5% and the M5 in 5.7% of the sequences (Fig. 6). The Inr motif experienced the lowest amount of change, the most significant being a decrease to 93.4% among the up-regulated genes of the infection set. The amount of M3 and M5 motifs changed significantly among all sets of genes but most significantly among those up-regulated during the 2 h of exposure to VECs. Twenty-five point five percent of the upstream regions of all genes up-regulated harbored a M3 motif and 34% a M5 motif, which approximately corresponded to a five- and six-fold increase, respectively. Even after subtraction of ribosomal subunit encoding genes, which are known to harbor these motifs (Smith et al., 2011), the elevation of M3 and M5 remains high at 14.9% and 7.5%, respectively.

4. Discussion

Eukaryotic genomes vary in size and protein-coding capacity ranges from approximately 2,000 in Encephalitozoon cuniculi to possibly even 60,000 in *T. vaginalis* (Katinka et al., 2001; Carlton et al., 2007). In the parabasalian parasite, the massive coding capacity is the result of the duplication of at least large sections of the genome and additional, partly massive, expansion of some gene families (Carlton et al., 2007). However, even if it is considered that the duplication of genes can also lead to the expression of redundant copies and proteins with identical functions, the fact that almost 28,000 remain after clustering genes with 90% nucleotide identity, demonstrates the large arsenal of proteins that the parasite can tap. Our transcriptomes encompass >30,000 expressed genes and cover 98% of the 714 genes recently discussed in the microarray-based analysis of Horváthová et al. (2012) that focused on iron-regulated genes in *Trichomonas*. The difference demonstrates that deep-sequencing transcriptomes from different conditions will expand this list even further. The overall response of *T. vaginalis* to oxygen stress and upon contact with human tissue was extensive, with hundreds of at least two-fold up- and down-regulated genes, and with individual genes experiencing a 100-fold increase in gene expression.

It has been suggested that the expansion of certain gene families is not random, but rather favors those that aid the special lifestyle of *Trichomonas* (Carlton et al., 2007). However, these gene families can only then act on the parasite, if they are all expressed. By comparing the correlation of expression within gene families, we found evidence that indeed many hundreds of the expanded families, including paralogous groups of Tyrosine Kinase-Like (TKL) kinases, actin, rabs or tubulin for example, are co-regulated (Fig. 5C). What is puzzling is why so many gene families cluster around the 45–50% identity mark when plotting global nucleotide identity versus expression correlation (Fig. 5C). They do not only correspond to a limited set of certain gene families, nor do they correlate with the size of gene families. This ‘cloud’ might represent a footprint of an evolutionary event (genome duplication?), or maybe at this level of identity a certain threshold level is reached where any two genes that are compared can have any type of co-regulation imaginable; i.e. it could also be random. It indicates that the extension of specific gene families – or the retention thereof in comparison with others – might be related to their biological significance for the parasite’s specialised lifestyle. Yet, the expansion of a gene family does not always result in high expression, as can be seen for example from the BspA family, and not always does a low or high expression level of an individual gene directly translate into a high or low amount of protein (Mair et al., 2006; Gry et al., 2009). Horváthová and colleagues (2012) also noted that among genes regulated under different iron concentrations, not all genes of a family experienced identical changes. In some cases all or the majority of paralogous copies were simultaneously regulated, in others it was only one gene of a family. Our results are congruent with those findings, for a larger sample of genes, different strains and different conditions, suggesting that this trend is general for *Trichomonas*.

The results for some genes such as those of the BspA family and for example for some of the surface-associated genes identified by de Miguel et al., 2010) are peculiar. For instance, our results for the entire BspA family confirm earlier findings for a few BspA genes, that they are up-regulated during amoeboid transition and simultaneous oxygen exposure (Noël et al., 2010). However, we found that oxygen-adapted cells do not up-regulate BspA genes upon exposure to VECs, which suggests either an oxygen-dependent component might be involved in the up-regulation of this, and perhaps other, gene families or that the strong up-regulation of, in particular, the translation machinery, in our case ‘masks’ less strongly regulated genes.

The many different strains of *T. vaginalis* isolated and examined to date disclose very different behaviors upon environmental stimuli (Ellis et al., 1992; Pereira-Neves and Benchimol, 2007; de Miguel et al., 2012). As the different behaviors result from different gene expression – demonstrated for example by the increased expression rate of TSP in highly-adherent strains (de Miguel et al., 2012) – this needs to be considered when comparing the sets of expression data from different *Trichomonas* strains. Also, albeit quantitative real time PCRs on individual genes can mirror patterns observed among large-scale generated data, these two techniques cannot always be directly compared. Due to the required normalisation of deep-sequencing data, genes with a very high expression will tend to mask the expression values of those with very low expression.

The parasite’s core energy metabolism harbors many proteins with iron-sulfur clusters and is hence very sensitive to oxygen. High levels of oxygen therefore directly hamper substrate level phosphorylation, the only pathway for ATP synthesis within the hydrogenosomes of *Trichomonas*. However, the natural habitat of *Trichomonas* is never absolutely oxygen free and, during transmission and menstruation, the parasite experiences high levels of oxygen fluctuation (Ellis et al., 1994; Hill et al., 2005). Therefore the

parasite requires an elaborate and fast system in order to deal with oxygen stress, which our data highlights. This response can be separated into two main categories: (i) synthesizing ROS scavengers and (ii) re-synthesizing proteins, which have been damaged through ROS (Supplementary Fig. S2). Among the large set of radical scavengers that were up-regulated, one particular peroxiredoxin (TVAG_455310) appeared to have the lead role after 30 and 120 min of oxygen stress and it's expression remains among the highest of all expressed genes in the oxygen-adapted parasites grown in the CO_2 incubator. The suggestion that cysteine replaces glutathione in *Trichomonas* as an antioxidant (Ellis et al., 1994) is supported by the up to nine-fold increase of two cysteine synthases (TVAG_040090 and TVAG_387920) during oxygen stress.

The response to oxygen encompasses a broad set of genes, with many members of the many radical-scavenging protein families being up-regulated at once. A large proportion of the transcriptional response appears to be dealing with the degradation of toxic ROS and shutting down energy consuming processes, in particular protein translation. Although the parasite is able to consume oxygen from the medium (Ellis et al., 1994; Chapman et al., 1999), it does not need to completely eradicate it before continuing with proliferation.

We observed a similar response in cells challenged with oxygen-stress alone and oxygen-stress combined with exposure to MS74. The extreme response to oxygen masks and possibly alters the expression shifts induced upon contact with MS74. This is underpinned by the many proteins that are down-regulated during oxygen stress, which in contrast were up-regulated during the infection of MS74 with parasites adapted in the CO_2 incubator, such as many cysteine proteases, Hsp and ribosomal subunits (Figs. 2 and 4). It shows that the cascade of events induced by oxygen stress needs to be clearly distinguished from those of host contact adhesion. Intriguingly, the oxygen stress first entails the up-regulation of transcription factors and other signaling molecules, a trend not observed during the exposure of oxygen-adapted T016 to MS74. In this latter case the translation machinery and proteins of the actin cytoskeleton are immediately among the top up-regulated genes after 5 min and this pattern does not change significantly over the 2 h analysed. It suggests that the up-regulation of these genes either requires transcription factors different from the MYB family – maybe with a WD40 repeat, as proteins with such repeats were found to be up-regulated across all three time points (Fig. 4) – or their up-regulation is independent of the up-regulation of the responsible transcription factors. That might also explain why we observed a significant increase in transcription in genes with the M3 promoter motif (Fig. 6), but no up-regulation of a recently identified M3 binding protein (TVAG_225940; Smith et al., 2011).

Trichomonas responds to oxygen with a dramatic down-regulation of the translation machinery (KOG category J, Fig. 3) and many genes encoding proteins of the KOG category O (posttranslational modification, turnover and chaperones), in particular Hsps (Figs. 2 and 3). Approximately 60% of a cell's ATP turnover is dedicated to protein synthesis, folding and posttranslational modification (Stouthamer, 1973). For *Trichomonas*, oxygen stress in direct consequence most likely also results in a depletion of the ATP pool, as core components of the hydrogenosomal metabolism are direct oxygen targets. Therefore the parasite first battles oxygen stress in a concerted effort, including the decrease of ATP consumption, before switching the cell's metabolism back to normal. The down-regulation of chaperones such as the ATPase Hsp70 perhaps reflects not only a lower demand for protein folding due to a decrease in protein synthesis, but additional means to save ATP. In accordance with this, we see that parasites that have been adapted to the oxygen conditions (*AdInf*) do the contrary when exposed to MS74: they massively up-regulate the translational machinery, including the chaperones required for subsequent protein folding (Fig. 4).

There have been reports that several common proteins of carbon and energy metabolism moonlight as surface proteins that help the parasite to adhere to VECs (Garcia and Alderete, 2007; Mundodi et al., 2008; Meza-Cervantez et al., 2011). Specifically, the proteins ME, PFO, two subunits of SCS, GAPDH and enolase have been implicated in this function and were designated as adhesion proteins. In the oxygen-adapted and human epithelia-challenged cells (*AdInf* set), we did not observe a specific induction of transcripts beyond what was observed for enzymes of energy metabolism in general (Supplementary Table S7). Thus, if those enzymes are moonlighting for adhesion, which has been questioned (Addis et al., 1998; Brugerolle et al., 2000; Hirt et al., 2007; Ryan et al., 2011), then it does not entail a specific increase in their transcripts upon contact with VECs. In our view their up-regulation is better explained by them being part of the general up-regulation of the ATP-synthesis machinery, in particular for protein synthesis as mentioned above, rather than those enzymes acting as potential adhesion proteins on the cell surface.

The transcriptomic response of *Trichomonas* to contact with VECs underpins three fundamental processes thought to be essential for the parasite to successfully establish an infection and this is increasing the expression of genes involved in (i) protein synthesis (proliferation), (ii) phenotypic plasticity and (iii) host cell degradation (Figs. 3 and 4). Several cysteine protease-encoding genes were up-regulated across all three time points among the *AdInf* libraries – most dominantly TVAG_355480 with an up-regulation of approximately 90-fold – supporting their suspected central role regarding virulence, cytoadherence, hemolysis and cytotoxicity of *T. vaginalis* (Sommer et al., 2005; Hirt et al., 2011; Ramón-Luing et al., 2011). The major up-regulation of actin and actin-associated genes provides further evidence for this part of the cytoskeleton playing a lead role during the amoeboid transition as suggested previously (Brugerolle et al., 1996; Bricheux et al., 2000; Kusdian et al., in press). The importance of the cytoskeletal transformation of *Trichomonas* during infection is in general reflected by the increase in genes of the KOG categories N (cell motility) and Z (cytoskeleton) (Fig. 3). Flagellate-amoeboid morphogenesis is rare in nature but also occurs in another human parasite, *Naegleria fowleri* (Fulton, 1993). Finally, *Trichomonas* has been observed to not only perform binary fission, which is widely thought to be the norm, but to also become multinuclear (Yusof and Kumar, 2011); possibly allowing a faster rate of proliferation. We noticed that 60 min into the infection, some cells significantly increase their overall mass and become multinuclear, a phenotype we refer to as a juggernaut, and which is reminiscent of the schizont stage of apicomplexan parasites. This type of rapid proliferation would require massive protein synthesis and would help to explain the strong increase in genes needed for protein translation, fueled by the availability of rich substrate scavenged from the host tissue.

To summarise, *Trichomonas* expresses a core set of >24,000 genes and responds to the different environmental stimuli by the concerted regulation of hundreds of genes, with the transcript level of some increasing more than 100-fold and others decreasing more than 50-fold. The expression of the many gene families is co-regulated and the further a gene copy has diverged within a family, the less its expression correlates with that of the original "mother copy". The massive up-regulation of ROS buffering proteins allows the parasite to survive elevated oxygen levels that allows the parasite to overcome periods of higher-than-normal oxygen levels, which are naturally part of its lifestyle. Upon contact with VECs the parasite switches to an actin-cytoskeleton driven amoeboid and induces proliferation, which is likely fueled by scavenging host tissue. The ease and speed with which morphogenesis can be

induced in *T. vaginalis* offers opportunities to study locomotion shifts from tubulin-based swimming to actin-based gliding.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, Germany) grants (GO 1825/3-1 and MA1426/19-1) to S.B.G. and W.F.M. and the support of the "Stiftung zur Erforschung infektiös-immunologischer Erkrankungen" to SBG.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijpara.2013.04.002.

References

Addis, M.F., Rappelli, P., Delogu, G., Carta, F., Cappuccinelli, P., Fiori, P.L., 1998. Cloning and molecular characterization of a cDNA clone coding for *Trichomonas vaginalis* alpha-actinin and intracellular localization of the protein. Infect. Immun. 66, 4924–4931.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Carlton, J.M., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Miller, J.A., Morrison, H.G., Nayak, V., Pennington, C., Pinney, D.F., Roos, D.S., Ross, C., Stoeckert Jr., C.J., Sullivan, S., Treatman, C., Wang, H., 2009. GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens *Giardia lamblia* and *Trichomonas vaginalis*. Nucleic Acids Res. 37, D526–D530.

Benchimol, M., 2008. The hydrogenosome as a drug target. Curr. Pharm. Des. 14, 872–881.

Bricheux, G., Coffe, G., Pradel, N., Brugerolle, G., 1998. Evidence for an uncommon alpha-actinin protein in *Trichomonas vaginalis*. Mol. Biochem. Parasitol. 95, 241–249.

Bricheux, G., Coffe, G., Bayle, D., Brugerolle, G., 2000. Characterization, cloning and immunolocalization of a coronin homologue in *Trichomonas vaginalis*. Eur. J. Cell Biol. 79, 413–422.

Brugerolle, G., Bricheux, G., Coffe, G., 1996. Actin cytoskeleton demonstration in *Trichomonas vaginalis* and in other trichomonads. Biol. Cell 88, 29–36.

Brugerolle, G., Bricheux, G., Coffe, G., 2000. Immunolocalization of two hydrogenosomal enzymes of *Trichomonas vaginalis*. Parasitol. Res. 86, 30–35.

Carlton, J.M., Hirt, R.P., Silva, J.C., Delcher, A.L., Schatz, M., Zhao, Q., Wortman, J.R., Bidwell, S.L., Alsmark, U.C., Besteiro, S., Sicheritz-Ponten, T., Noel, C.J., Dacks, J.B., Foster, P.G., Simillion, C., Van de Peer, Y., Miranda-Saavedra, D., Barton, G.J., Westrop, G.D., Müller, S., Dessi, D., Fiori, P.L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuera, F.D., Simoes-Barbosa, A., Brown, M.T., Hayes, R.D., Mukherjee, M., Okumura, C.Y., Schneider, R., Smith, A.J., Vanacova, S., Villalvazo, M., Haas, B.J., Pertea, M., Feldblyum, T.V., Utterback, T.R., Shu, C.L., Osoegawa, K., de Jong, P.J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S.B., Logsdon Jr., J.M., Henze, K., Gupta, A., Wang, C.C., Dunne, R.L., Upcroft, J.A., Upcroft, P., White, O., Salzberg, S.L., Tang, P., Chiu, C.H., Lee, Y.S., Embley, T.M., Coombs, G.H., Mottram, J.C., Tachezy, J., Fraser-Liggett, C.M., Johnson, P.J., 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science 315, 207–212.

Chapman, A., Linstead, D.J., Lloyd, D., 1999. Hydrogen peroxide is a product of oxygen consumption by *Trichomonas vaginalis*. J. Biosci. 24, 339–344.

Cong, P., Luo, Y., Bao, W., Hu, S., 2010. Genomic organization and promoter analysis of the *Trichomonas vaginalis* core histone gene families. Parasitol. Int. 59, 29–34.

Coombs, G.H., Westrop, G.D., Suchan, P., Puzova, G., Hirt, R.P., Embley, T.M., Mottram, J.C., Müller, S., 2004. The amitochondriate eukaryote *Trichomonas vaginalis* contains a divergent thioredoxin-linked peroxiredoxin antioxidant system. J. Biol. Chem. 279, 5249–5256.

de Miguel, N., Lustig, G., Twu, O., Chattopadhyay, A., Wohlschlegel, J.A., Johnson, P.J., 2010. Proteome analysis of the surface of *Trichomonas vaginalis* reveals novel proteins and strain-dependent differential expression. Mol. Cell. Proteomics 9, 1554–1566.

de Miguel, N., Riestra, A., Johnson, P.J., 2012. Reversible association of tetraspanin with *Trichomonas vaginalis* flagella upon adherence to host cells. Cell. Microbiol. 14, 1797–1807.

Ellis, J.E., Cole, D., Lloyd, D., 1992. Influence of oxygen on the fermentative metabolism of metronidazole-sensitive and resistant strains of *Trichomonas vaginalis*. Mol. Biochem. Parasitol. 56, 79–88.

Ellis, J.E., Yarlett, N., Cole, D., Humphreys, M.J., Lloyd, D., 1994. Antioxidant defences in the microaerophilic protozoan *Trichomonas vaginalis*: comparison of metronidazole-resistant and sensitive strains. Microbiology 140, 2489–2494.

Enright, A.J., Van Dongen, S., Ouzounis, C.A., 2002. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. 30, 1575–1584.

Fulton, C., 1993. *Naegleria* – a research partner for cell and developmental biology. J. Eukaryot. Microbiol. 40, 520–532.

Garcia, A.F., Alderete, J., 2007. Characterization of the *Trichomonas vaginalis* surface-associated AP65 and binding domain interacting with trichomonads and host cells. BMC Microbiol. 7, 116.

Gold, D., Ofek, I., 1992. Adhesion of *Trichomonas vaginalis* to plastic surfaces: requirement for energy and serum constituents. Parasitology 105, 55–62.

Gretes, M.C., Poole, L.B., Karplus, P.A., 2012. Peroxiredoxins in parasites. Antioxid. Redox Signaling 17, 608–633.

Gry, M., Rimini, R., Strömberg, S., Asplund, A., Pontén, F., Uhlén, M., Nilsson, P., 2009. Correlations between RNA and protein expression profiles in 23 human cell lines. BMC Genomics 10, 365.

Hill, D.R., Brunner, M.E., Schmitz, D.C., Davis, C.C., Flood, J.A., Schlievert, P.M., Wang-Weigand, S.Z., Osborn, T.W., 2005. In vivo assessment of human vaginal oxygen and carbon dioxide levels during and post menses. J. Appl. Physiol. 99, 1582–1591.

Hirt, R., Noel, C.J., Sicheritz-Ponten, T., Tachezy, J., Fiori, P., 2007. *Trichomonas vaginalis* surface proteins: a view from the genome. Trends Parasitol. 23, 540–547.

Hirt, R., de Miguel, N., Nakjang, S., Dessi, D., Liu, Y.C., Diaz, N., Rappelli, P., Acosta-Serrano, A., Fiori, P.L., Mottram, J.C., 2011. *Trichomonas vaginalis* pathobiology new insights from the genome sequence. Adv. Parasitol. 77, 87–140.

Horváthová, L., Šafaříková, L., Basler, M., Hrdy, I., Campo, N.B., Shin, J.W., Huang, K.Y., Huang, P.J., Lin, R., Tang, P., Tachezy, J., 2012. Transcriptomic identification of iron-regulated and iron-independent gene copies within the heavily duplicated *Trichomonas vaginalis* genome. Genome Biol. Evol. 4, 1017–1029.

Hrdý, I., Müller, M., 1995. Primary structure of the hydrogenosomal malic enzyme of *Trichomonas vaginalis* and its relationship to homologous enzymes. J. Eukaryot. Microbiol. 42, 593–603.

Katinka, M.D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., Vivarès, C.P., 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414, 450–453.

Kulda, J., 1999. Trichomonads, hydrogenosomes and drug resistance. Int. J. Parasitol. 29, 199–212.

Kusdian, G., Woehle, C., Martin, W.F., Gould, S.B., in press. The actin-based machinery of *Trichomonas vaginalis* mediates flagellate-amoeboid transition and migration across host tissue. Cell. Microbiol. doi:10.1111/cmi.12144.

Lal, K., Noel, C., Field, M., Goulding, D., Hirt, R., 2006. Dramatic reorganisation of *Trichomonas* endomembranes during amoebal transformation: a possible role for G-proteins. Mol. Biochem. Parasitol. 148, 99–102.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.Genome Project Data Processing Subgroup, 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.

Linstead, D., Bradley, S., 1988. The purification and properties of two soluble reduced nicotinamide: acceptor oxidoreductases from *Trichomonas vaginalis*. Mol. Biochem. Parasitol. 27, 125–133.

Mair, G.R., Braks, J.A., Garver, L.S., Wiegant, J.C., Hall, N., Dirks, R.W., Khan, S.M., Dimopoulos, G., Janse, C.J., Waters, A.P., 2006. Regulation of sexual development of *Plasmodium* by translational repression. Science 313, 667–669.

Mentel, M., Zimorski, V., Haferkamp, P., Martin, W., Henze, K., 2008. Protein import into hydrogenosomes of *Trichomonas vaginalis* involves both N-terminal and internal targeting signals: a case study of thioredoxin reductases. Eukaryot. Cell 7, 1750–1757.

Meza-Cervantez, P., González-Robles, A., Cárdenas-Guerra, R.E., Ortega-López, J., Saavedra, E., Pineda, E., Arroyo, R., 2011. Pyruvate:ferredoxin oxidoreductase (PFO) is a surface-associated cell-binding protein in *Trichomonas vaginalis* and is involved in trichomonal adherence to host cells. Microbiology 157, 3469–3482.

Müller, M., 1988. Energy metabolism of protozoa without mitochondria. Annu. Rev. Microbiol. 42, 465–488.

Müller, M., Mentel, M., van Hellemond, J.J., Henze, K., Woehle, C., Gould, S.B., Yu, R.Y., van der Giezen, M., Tielens, A.G., Martin, W.F., 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol. Mol. Biol. Rev. 76, 444–495.

Mundodi, V., Kucknoor, A.S., Alderete, J.F., 2008. Immunogenic and plasminogen-binding surface-associated alpha-enolase of *Trichomonas vaginalis*. Infect. Immun. 76, 523–531.

Noël, C.J., Diaz, N., Sicheritz-Ponten, T., Safarikova, L., Tachezy, J., Tang, P., Fiori, P.L., Hirt, R.P., 2010. *Trichomonas vaginalis* vast BspA-like gene family: evidence for functional diversity from structural organisation and transcriptomics. BMC Genomics 11, 99.

Okumura, C.Y., Baum, L.G., Johnson, P.J., 2008. Galectin-1 on cervical epithelial cells is a receptor for the sexually transmitted human parasite *Trichomonas vaginalis*. Cell. Microbiol. 10, 2078–2090.

Page-Sharp, M., Behm, C.A., Smith, G.D., 1996. *Tritrichomonas foetus* and *Trichomonas vaginalis*: the pattern of inactivation of hydrogenase activity by oxygen and activities of catalase and ascorbate peroxidase. Microbiology 142, 207–211.

Paget, T.A., Lloyd, D., 1990. *Trichomonas vaginalis* requires traces of oxygen and high concentrations of carbon dioxide for optimal growth. Mol. Biochem. Parasitol. 41, 65–72.

Pal, C., Bandyopadhyay, U., 2012. Redox-active antiparasitic drugs. Antioxid. Redox Signal. 17, 555–582.

Pereira-Neves, A., Benchimol, M., 2007. Phagocytosis by *Trichomonas vaginalis*: new insights. Biol. Cell 99, 87–101.
Petrin, D., Delgaty, K., Bhatt, R., Garber, G., 1998. Clinical and microbiological aspects of *Trichomonas vaginalis*. Clin. Microbiol. Rev. 11, 300–317.
Pruitt, K.D., Tatusova, T., Maglott, D.R., 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. 35, D61–D65.
Pütz, S., Dolezal, P., Gelius-Dietrich, G., Bohacova, L., Tachezy, J., Henze, K., 2006. Fe-hydrogenase maturases in the hydrogenosomes of *Trichomonas vaginalis*. Eukaryot. Cell 5, 579–586.
Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.
Ramón-Luing, L.D., Rendón-Gandarilla, F.J., Puente-Rivera, J., Ávila-González, L., Arroyo, R., 2011. Identification and characterization of the immunogenic cytotoxic TvCP39 proteinase gene of *Trichomonas vaginalis*. Int. J. Biochem. Cell Biol. 43, 1500–1511.
Rasoloson, D., Tomková, E., Cammack, R., Kulda, J., Tachezy, J., 2001. Metronidazole-resistant strains of *Trichomonas vaginalis* display increased susceptibility to oxygen. Parasitology 123, 45–56.
Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16, 276–277.
Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.
Ryan, C.M., de Miguel, N., Johnson, P.J., 2011. *Trichomonas vaginalis*: current understanding of host-parasite interactions. Essays Biochem. 51, 161–175.
Schumacher, M.A., Lau, A.O., Johnson, P.J., 2003. Structural basis of core promoter recognition in a primitive eukaryote. Cell 115, 413–424.
Schwebke, J.R., Hobbs, M.M., Taylor, S.N., Sena, A.C., Catania, M.G., Weinbaum, B.S., Johnson, A.D., Getman, D.K., Gaydos, C.A., 2011. Molecular testing for *Trichomonas vaginalis* in women: results from a prospective US clinical trial. J. Clin. Microbiol. 49, 4106–4111.
Smith, A., Johnson, P., 2011. Gene expression in the unicellular eukaryote *Trichomonas vaginalis*. Res. Microbiol. 162, 646–654.
Smith, A.J., Chudnovsky, L., Simoes-Barbosa, A., Delgadillo-Correa, M.G., Jonsson, Z.O., Wohlschlegel, J.A., Johnson, P.J., 2011. Novel core promoter elements and a cognate transcription factor in the divergent unicellular eukaryote *Trichomonas vaginalis*. Mol. Cell. Biol. 31, 1444–1458.
Smutná, T., Gonçalves, V.L., Saraiva, L.M., Tachezy, J., Teixeira, M., Hrdy, I., 2009. Flavodiiron protein from *Trichomonas vaginalis* hydrogenosomes: the terminal oxygen reductase. Eukarot. Cell 8, 47–55.
Sommer, U., Costello, C.E., Hayes, G.R., Beach, D.H., Gilbert, R.O., Lucas, J.J., Singh, B.N., 2005. Identification of *Trichomonas vaginalis* cysteine proteases that induce apoptosis in human vaginal epithelial cells. J. Biol. Chem. 280, 23853–23860.
Stouthamer, A.H., 1973. A theoretical study on the amount of ATP required for synthesis of microbial cell material. Antonie Van Leeuwenhoek 39, 545–565.
Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A., 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4, 41.
Upcroft, P., Upcroft, J.A., 2001. Drug targets and mechanisms of resistance in the anaerobic protozoa. Clin. Microbiol. Rev. 14, 150–164.
Wagner, G., Levin, R., 1978. Oxygen tension of the vaginal surface during sexual stimulation in the human. Fertil. Steril. 30, 50–53.
Williams, K.P., Leadlay, P.F., Lowe, P.N., 1990. Inhibition of pyruvate:ferredoxin oxidoreductase from *Trichomonas vaginalis* by pyruvate and its analogues. Comparison with the pyruvate decarboxylase component of the pyruvate dehydrogenase complex. Biochem. J. 268, 69–75.
Yusof, A., Kumar, S., 2011. Ultrastructural changes during asexual multiple reproduction in *Trichomonas vaginalis*. Parasitol. Res. 110, 1823–1828.