Novel tumor suppressor roles for *GZMA* and *RASGRP1* in dissemination of both *Theileria annulata*-transformed macrophages and human B-lymphoma cells

Zineb Rchiad[1,2,3*], Malak Haidar[1,2,3*], Hifzur Rahman Ansari[1,4°], Shahin Tajeri[2,3°], Sara Mfarrej[1], Fathia Ben Rached[1], Abhinav Kaushik[1], Gordon Langsley[2,3#], Arnab Pain[1,5#]

[1]Pathogen Genomics Laboratory, BESE Division, King Abdullah University of Science and Technology (KAUST), Thuwal-23955-6900, Saudi Arabia.
[2]Laboratoire de Biologie Cellulaire Comparative des Apicomplexes, Faculté de Médecine, Université Paris Descartes - Sorbonne Paris Cité, France.
[3] Inserm U1016, Cnrs UMR8104, Cochin Institute, Paris, 75014 France.
[4]King Abdullah International Medical Research Center (KAIMRC), King Abdulaziz Medical City, Ministry of National Guard Health Affairs, Jeddah 21423, Saudi Arabia.
[5]Global Station for Zoonosis Control, Global Institution for Collaborative Research and Education (GI-CoRE), Hokkaido University, N20 W10 Kita-ku, Sapporo, Japan.
*Co-first authors; °Co-second authors

Running Title: Role of Key Genes in *Theileria annulata*-mediated Leukocyte Transformation

Key Words: *Theileria annulata*, *RASGRP1*; *GZMA;* Tumor suppressor; Transcriptome

Financial Support: This study was supported by the Competitive Research Grant 4 (OSR-2015-CRG4-2610) from the Office for Sponsored Research in King Abdullah University of Science and Technology (KAUST). ST was supported by a ParaFrap post-doctoral fellowship and GL acknowledges support from Labex ParaFrap (ANR-11-LABX-0024), INSERM and the CNRS.

#Corresponding authors:
Gordon Langsley; Email: gordon.langsley@inserm.fr, 27, rue du Faubourg-Saint-Jacques, 75014 Paris, office: +33-(0)1-40-51-65-92
Arnab Pain; Email: arnab.pain@kaust.edu.sa, 4700 King Abdullah University of Science and Technology, Thuwal 23955 6900, phone: +966 (0) 2808-2561

The authors declare no potential conflict of interest.

This manuscript is a research article composed of a 40 word summary, 160-word abstract, 4390 text word count, 6 figures, 1 table, 7 supplementary figures, 4 supplementary tables and 61 references.

Summary

We compared the transcriptomes of *Theileria annulata* transformed B-lymphocytes to 934 human cancer cell lines and provide functional evidence for shared tumor suppressor roles for GZMA and RASGRP1 in controlling the dissemination phenotype of both human B lymphomas and Theileria-transformed leukocytes.

Abstract

Theileria annulata is a tick-transmitted apicomplexan parasite that infects and transforms bovine leukocytes into disseminating tumors that cause a disease called tropical theileriosis. Using comparative transcriptomics we identified genes transcriptionally perturbed during *Theileria*-induced transformation. Dataset comparisons highlighted a small set of genes associated with *Theileria*-transformed leukocyte dissemination. The roles of Granzyme A (*GZMA*) and RAS guanyl-releasing protein 1 (*RASGRP1*) were verified by CRISPR/Cas9-mediated knock-down. Knocking down of *GZMA* and *RASGRP1* in attenuated macrophages led to a regain in their dissemination in Rag2/γC mice confirming their role as dissemination suppressors *in vivo*. We further evaluated the roles of *GZMA* and *RASGRP1* in human B-lymphoma cells by comparing the transcriptome of 934 human cancer cell lines to that of *Theileria*-transformed bovine host cells. We confirmed dampened dissemination potential of human B-lymphoma cells that overexpress *GZMA* and *RASGRP1*. Our results provide evidence that *GZMA* and *RASGRP1* have a novel tumor suppressor function in both *T. annulata*-infected bovine host cells and in human B-lymphomas.

INTRODUCTION

Theileria annulata is a tick-transmitted apicomplexan parasite that infects and transforms bovine leukocytes into disseminating tumors that cause a widespread disease called tropical theileriosis. In countries endemic for tropical theileriosis live attenuated vaccines are produced by multiples *in vitro* passages of virulent, transformed macrophages and vaccination protects animals from severe disease (1). Amazingly the fully transformed state can be completely reversed by drug-induced parasite death making *Theileria*-infected leukocytes a powerful cellular model to identify genes regulating cellular transformation and dissemination (2). This parasite-based reversible model of leukocyte transformation has allowed the identification of several cell signaling pathways associated with the virulence of *Theileria*-transformed leukocytes such as c-Jun NH2-terminal kinase/c-Jun/PI3 kinase signaling (3), protein kinase A (PKA) (4), transforming growth factor beta 2 (TGF-b2) (5) (6) and SMYD3/MMP9 (7). MMP-9 and c-Jun are associated with invasion, proliferation and angiogenesis in *Theileria*-mediated host cell transformation as well as in human cancer (3, 8-10). Epigenetic changes also contribute to *Theileria*-induced leukocyte transformation (12). OncomiR addiction has been described as being generated by a miR-155 feedback loop in *T. annulata*-transformed B cells (13). Similarly, miR-126-5p contributes to infected macrophage dissemination through JNK-Interacting Protein-2 (JIP2)/JNK1/AP1-mediated *MMP9* transcription (14).

It's well established that *T. annulata* modulates gene expression of its host cell and hijacks key signalling cascades. For example, RNA extracted from *T. annulata*-transformed B cells was used to screen bovine microarrays demonstrating that infection had reconfigured host cell gene expression (11). Nonetheless, a systematic and genome scale transcriptional comparison of B cells and macrophages transformed by *T. annulata* has been lacking. In this study, we used RNA-seq to define the transcriptional landscapes

of two *T. annulata*-transformed B-cell lines and a virulent *T. annulata*-transformed macrophage line (Ode) and the attenuated live vaccine directly derived from it. High stringency bioinformatic comparisons of the transcriptional landscapes identified four candidate genes (*MMP9*, *GZMA*, *RASGRP1* and *SEPP1*) as potential players in the dissemination of virulent *T. annulata*-transformed macrophages.

The infection of lymphocytes and macrophages by *Theileria annulata* causes a lymphoproliferative phenotype with properties largely similar to human cancer, most notably immortalization, independence of exogenous growth factors, uncontrolled proliferation and invasiveness. The similarity between *Theileria*-transformed leukocytes and human leukemia suggests that *Theileria-induced transformation* could be a powerful model to elucidate common mechanisms underpinning tumour virulence. In order to generalize our *Theileria*-based observations we compared the transcriptome maps of 934 human cancer cell lines to the transcriptomes of *T. annulata* transformed B-lymphocytes and provide functional evidence for shared tumor suppressor roles for *GZMA* and *RASGRP1* in controlling the dissemination phenotype of both human B lymphomas and *Theileria*-transformed leukocytes.

Materials and Methods

Cell lines

The BL3 (15), TBL3, BL20 (16), TBL20 B lymphocytes, and Ode macrophages (17) were cultured in RPMI 1640 medium supplemented with 2 mM of L-glutamine (Lonza, catalogue number 12-702F) and 10 mM Hepes (Lonza, catalogue number 17-737E), 10 % heat-inactivated FBS (Gibco, catalogue number 10082147), 100 units/ml of Penicillin and 100 µg/ml of streptomycin (Lonza, catalogue number 17-602E) and 10mM b-mercaptoethanol (Sigma-Aldrich, catalogue number M6250) for BL3/TBL3 and BL20/TBL20. The virulent (Vir) hyper-disseminating Ode cell line was used at low

104 passage (53-71), while its attenuated (Att) poorly disseminating vaccine counterpart
105 corresponded to passages 309-317. The OCI-LY19 cell line (DSMZ, ACC 528) was
106 cultured in Minimum Essential Medium Eagle - Alpha Modification (Gibco, catalog number
107 12000063) supplemented with 2.2g/L of sodium bicarbonate (Thermofisher Scientific,
108 catalog number 25080094), 20% FBS, 10 mM Hepes and 100 units/ml of Penicillin and
109 100 ug/ml of streptomycin. The RI-1 cell line (DSMZ, ACC 585) was cultured in RPMI1640
110 and supplemented with 10% FBS, 100 units/ml of Penicillin and 100 ug/ml of streptomycin
111 and 10 mM Hepes. All cell lines were incubated at 37°C with 5% CO_2. All cell lines were
112 regularly tested for mycoplasma contamination.

113 **RNA extraction**

114 Cells were seeded in 3 biological replicates at a density of $2.5x10^5$ cell/ml. RNA extraction
115 was performed using the PureLink RNA Mini Kit (Life technologies, catalogue number
116 12183018A) following the manufacturer's instructions. Briefly, cells were pelleted, lysed
117 and homogenized using a 21-gauge needle, then 70% ethanol was added to the cell
118 lysates and the samples were loaded on spin cartridges to bind RNA. After 3 washes,
119 RNA was eluted in RNase-free water. The quality of extracted RNA was verified using a
120 Bioanalyzer 2100 and quantification carried using Qubit (Invitrogen, catalogue number
121 Q10210).

122 **Illumina library preparation and sequencing**

123 Strand-specific RNA-sequencing (ssRNA-seq) libraries were prepared using the illumina
124 Truseq Stranded mRNA Sample Preparation Kit (Illumina, catalogue number RS-122-
125 2101) following the manufacturer's instructions. Briefly, 1ug of total RNA was used to
126 purify mRNA using poly-T oligo-attached magnetic beads. mRNA was then fragmented
127 and cDNA was synthesized using SuperScript III reverse transcriptase (Thermofisher,
128 catalogue number 18080044), followed by adenylation on the 3' end, barcoding and

adapter ligation. The adapter ligated cDNA fragments were then enriched and cleaned with Agencourt Ampure XP beads (Agencourt, catalogue number A63880). Libraries validation was conducted using the 1000 DNA kit on 2100 Bioanalyzer (Agilent Technologies, catalogue number 5067-1504) and quantified using qubit (Thermofisher, catalogue number Q32850). ssRNA libraries were sequenced on Illumina Hiseq2000 and Hiseq4000. The sequenced reads were mapped to the *Bos taurus* genome Btau 4.6.1. The quality of the sequenced libraries is shown in supplementary figure S1.

Sequencing data analysis

The quality of sequence reads and other parameters were checked using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The raw RNA-seq reads were processed for adaptor trimming by Trimmomatic (19). The strand-specific reads were mapped on to Bovine genome (bosTau7; Btau_4.6.1; GCF_000003205.5) using Tophat2 (-g 1 --library-type fr-firststrand). The samples with respective replicates were analyzed further for differential gene expression by three different tools, baySeq (20), DESeq2 (21) (fitType ="local") and CuffDiff2 (22) with default parameters unless mentioned specifically. The count values for DESeq2 and baySeq were calculated from BAM files using HTSeq-count tool (23). The transcriptome quality plots were generated by cummeRbund package (v2.14.0) in R (http://bioconductor.org/packages/release/bioc/html/cummeRbund.html). The sequencing data is available in the NCBI Gene Expression Omnibus, GEO ID: GSE135377.

Identification of differentially expressed genes after infection and attenuation by comparative transcriptome analysis

The transcriptome data was analyzed with baySeq, DESeq2 and CuffDiff2. A gene was considered as a differentially expressed gene (DEG) if it has a padj<0.05 and a fold change (FC)>2. The final list of DEGs contained genes commonly differentially expressed

in CuffDiff2, DESeq2 and baySeq. This approach minimalizes the total number of DEGs for further analysis and allows stringent selection of the most significant and reproducible DEGs.

qRT-PCR

Total RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, catalogue number 4368814) as follows: 2 μg of total RNA, 2 μL of RT buffer, 0.8 μL of 100mM dNTP mix, 2.0 μL of 10X random primers, 1 μL of MultiScribe reverse transcriptase and Nuclease-free water to a final volume of 20 μL. The reaction was incubated 10 min at 25°C, 2 h at 37°C then the enzyme inactivated at 85°C for 5 min. Real time PCR was performed in a 10 μL reaction containing 20-30 ng cDNA template, 5 μL 2X Fast SYBR Green Master Mix and 500 nM of forward and reverse primers. The reaction was run on the 7500 HT Fast Real-Time PCR System (Applied Biosystems). GAPDH was used as a housekeeping gene and the results were analyzed by the $2^{-\Delta\Delta CT}$ method. The error bars represent the SEM of 3 biological replicates. Primers were designed and assessed for secondary structures using the Primer Express Software v3.0. The primers of all genes are listed in Table S3.

Transfection

Macrophages were transfected by electroporation using the Nucleofector system (Amaxa Biosystems). A total of 5×10^5 cells were suspended in 100 μL of Nucleofector V solution mix (Lonza, VCA-1003) with 2 μg of *GZMA* and *RASGRP1* CRISPR/Cas9 plasmids and subjected to nucleofection using the cell line-specific program T-O17. The human B-lymphoma cells were transfected with 2 μg of *GZMA* (SCBT, sc-403958-ACT), *RASGRP1* (SCBT, sc-402120-ACT) and *SEPP1* (SCBT, sc-417457-ACT) activation plasmids, and the non-specific control plasmid (SCBT, sc-437275) using the T-O17

179 program. After transfection, cells were suspended in fresh complete medium and
180 incubated at 37°C with 5% CO_2.

181 **Matrigel chamber assay**

182 The invasive capacity of Ode macrophages was assessed *in vitro* using matrigel migration
183 chambers, as described in (3). The CultureCoat Medium basement membrane extract
184 (BME) 96-wells cell invasion assay was performed according to Culturex instructions
185 (Trevigen, catalog number 3482-096-K). After 24 h of incubation at 37°C, each well of the
186 top chamber was washed once in buffer. The top chamber was placed back onto the
187 receiver plate. One hundred microliters of cell dissociation solution-Calcein AM was added
188 to the bottom chamber of each well, and the mixtures were incubated at 37°C for 1 h with
189 fluorescently labeled cells to dissociate the cells from the membrane before reading at
190 485-nm excitation and 520-nm emission wavelengths.

191 **Soft agar colony forming assay**

192 A two-layer soft agar culture system was used. Cell counts were performed by ImageJ
193 software. A total of 2,500 cells were plated in a volume of 1.5 ml (0.7% bacto Agar+2×
194 RPMI 20% Fetal bovine Serum) over 1.5 ml base layer (1% bacto agar +2× RPMI 20%
195 Fetal bovine Serum) in 6-well plates. Cultures were incubated in humidified 37°C
196 incubators with an atmosphere of 5% CO_2 in air, and control plates were monitored for
197 growth using a microscope. At the time of maximum colony formation (10 days in
198 culture), final colony numbers were counted after fixation with 0.005% Cristal Violet.

199 **Intracellular levels of hydrogen peroxide (H_2O_2)**

200 Cells were seeded at $1x10^5$ cell/well in a 96 well plate and incubated in complete medium
201 for 18 h prior to the assay. Cells were then washed with PBS and incubated with 100 μL
202 of 5 M H2-DCFDA in PBS (Molecular Probes, catalogue number D399). H_2O_2 levels were

assayed on a fusion spectrofluorimeter (PackardBell) by spectrofluorimetry at 485 and 530nm excitation and emission wavelengths respectively.

In vivo mouse studies and quantification of _Theileria annulata_-transformed macrophages load in mouse tissues

T. annulata-infected macrophage cell lines (Virulent Ode passage 53, attenuated Ode passage 309, attenuated Ode transfected with *RASGRP1* CRISPR/Cas9 knock out plasmid and attenuated Ode transfected with *GZMA* CRISPR/Cas9 knock out plasmid) were injected into four groups of five Rag2γC immunodeficient mice that were equally distributed on the basis of age and sex in each group. The injection site was disinfected with ethanol and one million cells (in 200 µl PBS) were injected under the skin after gentle shaking of the insulin syringe. The mice were kept for 3 weeks and then they were humanely sacrificed and dissected. Six internal organs including heart, lung, spleen, mesentery, left kidney and liver were taken and stored in 500 *µ* L PBS in Eppendorf tubes at -20°C. The tissues were subjected to genomic DNA extraction using the QIAmp DNA mini kit (Qiagen, catalogue number 51304). DNA concentrations were measured by Nanodrop™ 1000 spectrophotometer (Thermo Fischer scientific) and before each quantitative PCR reaction samples were diluted to give a DNA concentration of 0.5 ng/*µ* L. Absolute copy numbers of a single copy *T. annulata* gene (*ama-1*, TA02980) that is representative of *T. annulata*-infected macrophage load in each tissue were estimated by the method described in (24), with some modifications. *ama-1* was cloned into pJET 1.2/blunt cloning vector using CloneJET PCR Cloning Kit (Thermo scientific, catalogue number K1232). The cloned plasmid was amplified in DH5-Alpha cells and purified with QIAfilter™ Plasmid Maxi Kit (Qiagen, catalogue number 12243). Plasmid concentration was measured using Qubit (Thermofisher, catalogue number Q32850). The primers for cloning were: forward 5'-GGAGCTAACTCTGACCCTTCG-3'and reverse 5'-

CCAAAGTAGGCCAATACGGC-3'. Quantitative PCR primers were: forward 5'-GACCGATTTCATGGCAAAGT-3'and reverse 5'-TTGGGGTCATGATGGGTTAT-3'.

Transcriptome-based clustering of *Theileria*-transformed bovine host cells and human cancer cell lines

The processed and quality trimmed reads from TBL20/BL20 and BL3/TBL3 samples were mapped to *Bos taurus* UMD3.1 genome using HISAT2 software (25) with default settings. The mapped reads were used for gene-level TPM quantification using StringTie (Version 1.3.3b) (26, 27). The quantified genes were converted into their human ortholog Ensemble gene ID by finding one-to-one orthologs between human and *B. taurus* genomes using OMA browser (28). Subsequently, TPM values of transcripts expressed across 934 human cancer cell lines were obtained from EBI cancer cell line Expression Atlas (29). The redundant transcripts in the cancer cell line expression set were collapsed using collapseRow function from the WGCNA R package (30). Using the common human ensemble gene ID, gene expression matrices of *B. taurus* and human cell lines were merged together, which was then subjected to hierarchical clustering of samples using HCPC (31) with 3 Principal Components (nPCs), which resulted in 4 broad clusters. The sub-cluster containing the human cancer cell lines along with TBL20/BL20 and BL3/TBL3 were shortlisted for further analysis. Whereby, samples were scanned for similar gene expression profiles by computing the adjacency of each shortlisted sample with the rest of the samples, using adjacency function (method="Distance") from WGCNA R package. The resultant adjacency matrix was then subjected to flashClust (31) program for computing the dendrogram for manual inspection of TBL20/BL20 and BL3/TBL3 containing sub-cluster. A schematic of the used pipeline is presented in Fig. S6. The complete dendrogram of the 934 human cancer cell lines and *Theileria*-tranformed lymphocytes can be viewed in Fig. S7.

253 **Ethics statement**

254 The protocol (12-26) was approved by the ethics committee for animal experimentation of
255 the University of Paris-Descartes (CEEA34.GL.03312). The university ethics committee is
256 registered with the French National Ethics Committee for Animal Experimentation that
257 itself is registered with the European Ethics Committee for Animal Experimentation. The
258 right to perform the mice experiments was obtained from the French National Service for
259 the Protection of Animal Health and satisfied the animal welfare conditions defined by
260 laws (R214-87 to R214-122 and R215-10) and GL was responsible for all animal
261 experimentation, as he holds the French National Animal Experimentation permit with the
262 authorisation number (B-75-1249). This project is also covered by the KAUST IBEC
263 number 19IBEC12.

264 **RESULTS:**

265 **Differentially expressed bovine genes in *T. annulata*-transformed leukocytes**

266 The infection and full transformation of the BL20 cell line with *T. annulata* caused
267 profound transcriptional changes, as previously reported for infected BL20 cells (11).
268 Similarly, infection of BL3 cell with *T. annulata* also provoked changes in host cell gene
269 expression (Fig. S2a). Transcriptional changes between virulent compared to attenuated
270 Ode macrophages are less profound, likely because the macrophages only appear to
271 differ in dissemination potential (Fig. S2a).

272 To identify bovine genes whose transcription is perturbed by transformation and
273 attenuation of dissemination of *T. annulata*-transformed leukocytes we concentrated on
274 the most differentially expressed genes (DEGs) (Fig. 1, Table S1). Many of these genes
275 are annotated as being implicated in cell proliferation and metastasis. Amongst the top
276 five-upregulated transcripts in TBL20 is *MMP9* (matrix metallopeptidase 9), a gene highly

277 expressed in different cancer types and linked to metastasis and angiogenesis (32). *WC1-*
278 *8* is the third most upregulated gene in TBL20 lymphocytes and has been described as
279 being also upregulated in ovarian carcinoma cells (33). The most down-regulated
280 transcripts in TBL20 cells include *LAIR1* (leukocyte associated immunoglobulin like
281 receptor 1) and *VPREB* (pre-B lymphocyte 1). *LAIR1* is a strong inhibitor of natural killer
282 cell-mediated cytotoxicity and an inhibitory receptor, which down-regulates B lymphocyte
283 immunoglobulin and cytokine production (34). Down-regulation of *LAIR-1* was not
284 unexpected, as it's loss of expression is observed during B cell proliferation (35). *ZBTB32*
285 (zinc finger and BTB domain containing 32), *IL21R* (interleukin 21 receptor), and *MMP9*
286 are also among the top five up-regulated transcripts in infected TBL3 B lymphocytes. The
287 five most down-regulated transcripts in TBL3 are *KRT6C* (keratin 6C), *MATK*
288 (megakaryocyte-associated tyrosine kinase), *IGSF9B* (immunoglobulin superfamily
289 member 9B), *A2M* (alpha-2-macroglobulin) and *H2AFY2* (H2A histone family, member
290 Y2). The biological functions of these genes include inhibition of cell growth and
291 proliferation (36), repression of DNA transcription (37) and inhibition of cell adhesion and
292 migration (38), functions that are often dampened to allow continuous proliferation and
293 survival of transformed cells. We confirmed by qRT-PCR differential expression of 21
294 randomly selected genes from the BL20/TBL20 and BL3/TBL3 RNA-seq datasets (Fig.
295 S2b).

296 **Identification of key genes potentially involved in *Theileria*-mediated macrophage**
297 **dissemination**

298 The most down-regulated transcripts in attenuated Ode macrophages are *SKAP2*
299 (src kinase associated phosphoprotein 2), a gene known to promote tumor metastasis
300 through the regulation of podosome formation in macrophages (39) and *NRP2* that

regulates tumor progression by promoting TGF-β-signaling (40). Down-regulation of these genes correlates with decreased dissemination of attenuated macrophages, as previously we have described loss of TGF-β-signaling as being associated with decreased dissemination (5). By contrast, the most highly upregulated transcripts include *SEPP1* (selenoprotein P) and *PTPRT* (protein tyrosine phosphatase, receptor type T). Both *SEPP1* and *PTPRT* have been previously described as a tumor suppressor gene (41, 42). Taken together, the identity of the most strongly up- and down-regulated genes argues that our differential transcription screen could identify novel genes regulating *Theileria*-transformed macrophage dissemination.

To define the genes likely playing important roles in transformation and dissemination, we compared genes differentially expressed (DE) in TBL3, TBL20 and attenuated Ode macrophages. We assumed that the genes most likely to play a key role are upregulated after infection and downregulated upon attenuation, and vice versa. This approach identified four genes likely to play key roles in the dissemination of *Theileria*-transformed leukocytes (Fig. 2a).

The genes are *MMP9, SEPP1, GZMA* and *RASGRP1* and their biological functions have been implicated in metastasis and cell invasion (32), selenium transport (43), peptide cleavage by immune cells (44) and regulation of B cell-development and homeostasis and differentiation (45), respectively (Table 1). Differential expression of these genes was confirmed by qRT-PCR (Fig. 2b). We focused on *GZMA*, *RASGRP1* and *SEPP1*, as the role of *MMP9* in metastasis/dissemination is well established including in *Theileria*-transformed macrophages (46). CRISPR/Cas9-mediated loss of *SEPP1* in attenuated macrophages resulted in a lethal phenotype highlighting its essentiality in transformed macrophage survival and the role of SEPP1 in *Theileria*-transformed cell lines was not further evaluated.

326 **Ablation of *GZMA* and *RASGRP1* by CRISPR/Cas9 knockdown**

327 To confirm a role for *GZMA* and *RASGRP1* in dissemination of *T. annulata*-
328 transformed Ode macrophages, we knocked down their expression by CRISPR/Cas9 and
329 decreased expression led to a regain in dissemination potential, as estimated in matrigel
330 traversal assays (Fig. 3b). Both GZMA and RASGRP1 have therefore, the potential to
331 function as suppressors of tumor dissemination and consistently, knockdown of *GZMA*
332 also led to a regain in the ability of attenuated macrophages to form colonies in soft agar
333 (Fig.3c). Taken together it strongly suggests that GZMA and RASGRP1 function as tumor
334 suppressors.

335 ***GZMA* and *RASGRP1* dampen *in vivo* dissemination of Ode macrophages**

336 Similar to metastatic tumor cells *T. annulata*-transformed leukocytes also
337 disseminate in immuno-deficient mice to distant organs and form proliferative foci (47).
338 Dissemination of *Theileria*-transformed leukocytes has been previously attributed to
339 increased production of matrix metaloproteinases (MMPs) (46). As *GZMA* and *RASGRP1*
340 knockdown led to a regain in matrigel traversal we used Rag2γC immunodeficient mice to
341 test for a regain in dissemination *in vivo*. The CRISPR/Cas9-induced ablation of
342 expression of GZMA and RASGRP1 gave rise to an increase in the number of *Theileria*-
343 containing tumors in heart, lung and mesentery, while knockdown of *RASGRP1* increased
344 the number of tumors in the liver (Fig. 4). Thus, loss of *RASGRP1* and *GZMA* expression
345 led to a regain in the invasive capacity of *T. annulata*-transformed macrophages into these
346 organs.

347 **Induced expression of *GZMA*, *RASGRP1* and *SEPP1* reduces human B-lymphoma**
348 **cell dissemination**

In order to extend the roles of GZMA, RASGRP1 and SEPP1 to human cancer we sought human tumor cells displaying transcriptional signatures similar to *T. annulata*-transformed leukocytes. To this end, the transcriptional profiles of 934 human cancer cells were obtained from the EBI cancer cell line expression atlas (29) and their profiles compared to those of *Theileria*-transformed TBL3 and TBL20 B-lymphocytes (Fig. 5a). Among the subset human B lymphomas OCI-LY19 and RI-1 displayed the transcriptional signature most similar to TBL3 and TBL20 and therefore, were used to test if *GZMA, RASGRP1* and *SEPP1* can act as suppressors of dissemination in certain types of human cancer. CRISPR-mediated transcriptional activation of *GZMA*, *RASGRP1* and *SEPP1* resulted in decreased matrigel traversal of the OCI-LY19 B lymphoma. By contrast, only upregulation of *GZMA* showed a statistically significant decrease in traversal of the RI-1 B lymphoma (Fig. 5c). Clearly, *GZMA* has the potential to act as a suppressor in two independent human B-lymphomas, whereas *RASGRP1* and *SEPP1* may only function as suppressors in specific B-lymphomas.

DISCUSSION

In this study, we provide a holistic view of the transcriptional landscape of two *T. annulata*-transformed B cell lines, TBL3 and TBL20, and in addition the landscape of virulent versus attenuated Ode macrophages. In order to find genes with commonly perturbed transcription the different datasets were compared using three independent pipelines to identify just four genes, as potential regulators of tumor dissemination. In addition to *MMP9* three other genes (*SEPP1*, *GZMA* and *RASGRP1*) were identified as potentially having a role in dissemination. SEPP1 is a major selenoprotein involved in selenium transport and cellular defense against oxidative stress (48). Attenuated

372 macrophages did not survive CRISPR/Cas9-knockdown of *SEPP1* implying death might
373 be due to a failure to control excessive oxidative stress, since attenuated macrophages
374 display high levels of H_2O_2 (49).

375 RASGRP1-activated Ras family proteins possess both pro- and anti-oncogenic
376 properties, depending on the downstream effector pathway and cellular context; reviewed
377 in (50). Our transcription profiling showed that most of the members of the *RASGRP* gene
378 family (*RASGRP1, 2 & 4*) are significantly downregulated in TBL3 and TBL20 (Fig. S3).
379 GZMA is a serine protease that contributes to killing of both tumors and pathogen-infected
380 cells via a caspase-independent pathway (44). *GZMA* expression induces reactive oxygen
381 species (51) and attenuated macrophages are known to be more oxidatively stressed than
382 virulent macrophages (49). Indeed, H_2O_2 output was reduced in attenuated macrophages
383 following CRISPR/Cas9-mediated *GZMA* knockdown (Fig. S4). Furthermore, it has been
384 shown that *RASGRP1*-deficient CD8 T cells exhibit markedly reduced expression of
385 GZMB (52) . This led us to investigate whether loss of *RASGRP1* could perhaps provoke
386 a drop in *GZMA* expression rendering attenuated macrophages doubly deficient in
387 dissemination *in vivo*. As hypothesized, we found that expression of *GZMA* decreased
388 after *RASGRP1* knockdown (Fig. S5). Moreover, *GZMA* and *RASGRP1* expression is
389 repressed by TGF-β (52, 53) and the role of TGF-β in regulating dissemination of
390 *Theileria*-transformed macrophages is very well established (4-6). Expression of *GZMA*
391 and *RASGRP1* was decreased in attenuated Ode macrophages treated with TGF-β
392 (Table S4). Taken altogether, it suggests that one way TGF-β promotes dissemination of
393 *Theileria*-transformed leukocytes could be via repression of both *GZMA* and *RASGRP1*
394 transcription and their impact on dissemination confirmed *in vivo* in mice. In addition to the
395 role of *GZMA, RASGRP1* and *SEPP1* in regulating tumorigenesis of *Theileria*-transformed
396 leukocytes, all three genes also dampened the capacity of the human OCI-LY19 B-

lymphoma to traverse matrigel and GZMA also dampened traversal of human RI-1 B lymphoma cells.

GZMA is known to cleave APEX1 (apurinic/apyrimidinic endodeoxyribonuclease 1) after Lys31 and destroys its oxidative repair functions. APEX1 is involved in NK-cell-mediated killing via GZMA (51, 54) and can suppress the activation of PARP1 during repair of oxidative DNA damage (55), and prevent oxidative stress by negatively regulating Rac1/GTPase activity (56). Additionally, APEX1 directly reduces the redox-sensitive cysteine residues of target transcription factors, enhancing their DNA binding and transcriptional activity. Analysis of our deep RNA-seq data revealed that *APEX1* is downregulated in attenuated macrophages and its expression increases after TGF-β treatment, along with an important downregulation of *GZMA* and *RASGRP1* (Table S4). It is over-expressed in many cancers (57) (58) (59) and has been implicated in growth, migration and invasion of colon cancer both *in vitro* and *in vivo* (60). Interestingly, APEX-1 protects melanoma cells from H_2O_2 induced apoptosis (57). The established function of APEX1 in human cancer sustains our novel observations on the tumor suppressor roles of *GZMA* and *RASGRP1.* The ensemble of our results allows us to propose a model, where TGF-β modulates tumor redox balance-mediated dissemination by regulating a GZMA/RASGRP1/APEX1 pathway (Fig. 6).

This study has revealed new players in dissemination and oxidative stress regulation of *Theileria*-transformed leukocytes and has provided evidence for similar roles for GZMA and RASGRP1 in transcriptionally matched human B-lymphoma cell lines. The similarity between *Theileria*-induced B cell transformation and human B-lymphomas has unveiled novel therapeutic targets to treat cancer.

Acknowledgements.

This study was supported by a Competitive Research Grant from the Office for Sponsored Research (OSR-2015-CRG4-2610) at King Abdullah University of Science and Technology (KAUST) awarded to AP and GL. ZR acknowledges KAUST for awarding her PhD studentship. ST was supported by a post-doctoral fellowship from ParaFrap (ANR-11-LABX-0024) and in addition to ANR-11-LABX-0024 GL also acknowledges core support from INSERM and the CNRS. We thank members of the Bioscience Core Laboratory (BCL) at KAUST for producing the raw sequencing datasets. Franck Letourneur of the genomics platform at the Cochin institute (GENOM'IC) for quantifying the pJET-*ama-1* plasmid.

Author contributions.

AP and GL conceived and designed the study. ZR prepared the ssRNAseq libraries, SM, MH and ZR ran qRT-PCR reactions. MH performed the soft agar colony formation assay, intracellular H_2O_2 levels and invasion assays. ST performed the mouse dissemination assays, HRA, AK and ZR performed data analysis. ZR and MH prepared the figures and ZR prepared the first draft of the manuscript with input from FBR that was then edited by MH, GL and AP.

References

1. **Nene V, Morrison WI.** 2016. Approaches to vaccination against Theileria parva and Theileria annulata. Parasite Immunol **38:**724-734.
2. **Tretina K, Gotia HT, Mann DJ, Silva JC.** 2015. Theileria-transformed bovine leukocytes have cancer hallmarks. Trends Parasitol **31:**306-314.
3. **Lizundia R, Chaussepied M, Huerre M, Werling D, Di Santo JP, Langsley G.** 2006. c-Jun NH2-terminal kinase/c-Jun signaling promotes survival and metastasis of B lymphocytes transformed by Theileria. Cancer Res **66:**6105-6110.
4. **Haidar M, Echebli N, Ding Y, Kamau E, Langsley G.** 2015. Transforming growth factor beta2 promotes transcription of COX2 and EP4, leading to a prostaglandin E2-driven autostimulatory loop that enhances virulence of Theileria annulata-transformed macrophages. Infect Immun **83:**1869-1880.
5. **Chaussepied M, Janski N, Baumgartner M, Lizundia R, Jensen K, Weir W, Shiels BR, Weitzman JB, Glass EJ, Werling D, Langsley G.** 2010. TGF-b2 induction regulates invasiveness of Theileria-transformed leukocytes and disease susceptibility. PLoS Pathog **6:**e1001197.
6. **Haidar M, Whitworth J, Noe G, Liu WQ, Vidal M, Langsley G.** 2015. TGF-beta2 induces Grb2 to recruit PI3-K to TGF-RII that activates JNK/AP-1-signaling and augments invasiveness of Theileria-transformed macrophages. Sci Rep **5:**15688.
7. **Cock-Rada AM, Medjkane S, Janski N, Yousfi N, Perichon M, Chaussepied M, Chluba J, Langsley G, Weitzman JB.** 2012. SMYD3 promotes cancer invasion by epigenetic upregulation of the metalloproteinase MMP-9. Cancer Res **72:**810-820.
8. **Hofmann UB, Westphal JR, Van Muijen GN, Ruiter DJ.** 2000. Matrix metalloproteinases in human melanoma. J Invest Dermatol **115:**337-344.
9. **Vleugel MM, Greijer AE, Bos R, van der Wall E, van Diest PJ.** 2006. c-Jun activation is associated with proliferation and angiogenesis in invasive breast cancer. Hum Pathol **37:**668-674.
10. **Adamson R, Logan M, Kinnaird J, Langsley G, Hall R.** 2000. Loss of matrix metalloproteinase 9 activity in Theileria annulata-attenuated cells is at the transcriptional level and is associated with differentially expressed AP-1 species. Mol Biochem Parasitol **106:**51-61.
11. **Kinnaird JH, Weir W, Durrani Z, Pillai SS, Baird M, Shiels BR.** 2013. A Bovine Lymphosarcoma Cell Line Infected with Theileria annulata Exhibits an Irreversible Reconfiguration of Host Cell Gene Expression. PLoS One **8:**e66833.
12. **Robert McMaster W, Morrison CJ, Kobor MS.** 2016. Epigenetics: A New Model for Intracellular Parasite-Host Cell Regulation. Trends Parasitol **32:**515-521.
13. **Marsolier J, Pineau S, Medjkane S, Perichon M, Yin Q, Flemington E, Weitzman MD, Weitzman JB.** 2013. OncomiR addiction is generated by a miR-155 feedback loop in Theileria-transformed leukocytes. PLoS Pathog **9:**e1003222.

14. **Haidar M, Rchiad Z, Ansari HR, Ben-Rached F, Tajeri S, Latre De Late P, Langsley G, Pain A.** 2018. miR-126-5p by direct targeting of JNK-interacting protein-2 (JIP-2) plays a key role in Theileria-infected macrophage virulence. PLoS Pathog **14:**e1006942.
15. **Theilen GH, Rush JD, Nelson-Rees WA, Dungworth DL, Munn RJ, Switzer JW.** 1968. Bovine leukemia: establishment and morphologic characterization of continuous cell suspension culture, BL-1. J Natl Cancer Inst **40:**737-749.
16. **Morzaria SP, Roeder PL, Roberts DH, Chasey D, Drew TW.** 1984. Characteristics of a continuous suspension cell line (BL20) derived from a calf with sporadic bovine leukosis.
17. **Singh S, Khatri N, Manuja A, Sharma RD, Malhotra DV, Nichani AK.** 2001. Impact of field vaccination with a Theileria annulata schizont cell culture vaccine on the epidemiology of tropical theileriosis. Vet Parasitol **101:**91-100.
18. **Goh S, Ngugi D, Lizundia R, Hostettler I, Woods K, Ballingall K, MacHugh ND, Morrison WI, Weir W, Shiels B, Werling D.** 2016. Identification of Theileria lestoquardi Antigens Recognized by CD8+ T Cells. PLoS One **11:**e0162571.
19. **Bolger AM, Lohse M, Usadel B.** 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30:**2114-2120.
20. **Hardcastle TJ, Kelly KA.** 2010. baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics **11:**422.
21. **Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL.** 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol **14:**R36.
22. **Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L.** 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol **31:**46-53.
23. **Anders S, Pyl PT, Huber W.** 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics **31:**166-169.
24. **Gotia HT, Munro JB, Knowles DP, Daubenberger CA, Bishop RP, Silva JC.** 2016. Absolute Quantification of the Host-To-Parasite DNA Ratio in Theileria parva-Infected Lymphocyte Cell Lines. PLoS One **11:**e0150401.
25. **Kim D, Langmead B, Salzberg SL.** 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods **12:**357-360.
26. **Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL.** 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc **11:**1650-1667.
27. **Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL.** 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol **33:**290-295.
28. **Altenhoff AM, Glover NM, Train CM, Kaleb K, Warwick Vesztrocy A, Dylus D, de Farias TM, Zile K, Stevenson C, Long J, Redestig H, Gonnet GH, Dessimoz C.** 2018. The OMA orthology database in 2018: retrieving

evolutionary relationships among all domains of life through richer web and programmatic interfaces. Nucleic Acids Res **46:**D477-D485.

29. **Papatheodorou I, Fonseca NA, Keays M, Tang YA, Barrera E, Bazant W, Burke M, Fullgrabe A, Fuentes AM, George N, Huerta L, Koskinen S, Mohammed S, Geniza M, Preece J, Jaiswal P, Jarnuczak AF, Huber W, Stegle O, Vizcaino JA, Brazma A, Petryszak R.** 2018. Expression Atlas: gene and protein expression across multiple studies and organisms. Nucleic Acids Res **46:**D246-D251.
30. **Langfelder P, Horvath S.** 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics **9:**559.
31. **Lê S, Josse J, Husson F.** 2008. FactoMineR: An R Package for Multivariate Analysis. 2008 **25:**18.
32. **Yu Q, Stamenkovic I.** 2000. Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis. Genes Dev **14:**163-176.
33. **Mangala LS, Zuzel V, Schmandt R, Leshane ES, Halder JB, Armaiz-Pena GN, Spannuth WA, Tanaka T, Shahzad MM, Lin YG, Nick AM, Danes CG, Lee JW, Jennings NB, Vivas-Mejia PE, Wolf JK, Coleman RL, Siddik ZH, Lopez-Berestein G, Lutsenko S, Sood AK.** 2009. Therapeutic Targeting of ATP7B in Ovarian Carcinoma. Clin Cancer Res **15:**3770-3780.
34. **Merlo A, Tenca C, Fais F, Battini L, Ciccone E, Grossi CE, Saverino D.** 2005. Inhibitory receptors CD85j, LAIR-1, and CD152 down-regulate immunoglobulin and cytokine production by human B lymphocytes. Clin Diagn Lab Immunol **12:**705-712.
35. **van der Vuurst de Vries AR, Clevers H, Logtenberg T, Meyaard L.** 1999. Leukocyte-associated immunoglobulin-like receptor-1 (LAIR-1) is differentially expressed during human B cell differentiation and inhibits B cell receptor-mediated signaling. Eur J Immunol **29:**3160-3167.
36. **Kim SO, Avraham S, Jiang S, Zagozdzon R, Fu Y, Avraham HK.** 2004. Differential expression of Csk homologous kinase (CHK) in normal brain and brain tumors. Cancer **101:**1018-1027.
37. **Perche PY, Vourc'h C, Konecny L, Souchier C, Robert-Nicoud M, Dimitrov S, Khochbin S.** 2000. Higher concentrations of histone macroH2A in the Barr body are correlated with higher nucleosome density. Curr Biol **10:**1531-1534.
38. **Kurz S, Thieme R, Amberg R, Groth M, Jahnke HG, Pieroh P, Horn LC, Kolb M, Huse K, Platzer M, Volke D, Dehghani F, Buzdin A, Engel K, Robitzki A, Hoffmann R, Gockel I, Birkenmeier G.** 2017. The anti-tumorigenic activity of A2M-A lesson from the naked mole-rat. PLoS One **12:**e0189514.
39. **Tanaka M, Shimamura S, Kuriyama S, Maeda D, Goto A, Aiba N.** 2016. SKAP2 Promotes Podosome Formation to Facilitate Tumor-Associated Macrophage Infiltration and Metastatic Progression. Cancer Res **76:**358-369.
40. **Grandclement C, Pallandre JR, Valmary Degano S, Viel E, Bouard A, Balland J, Remy-Martin JP, Simon B, Rouleau A, Boireau W, Klagsbrun M, Ferrand C, Borg C.** 2011. Neuropilin-2 expression promotes TGF-beta1-

mediated epithelial to mesenchymal transition in colorectal cancer cells. PLoS One **6:**e20444.

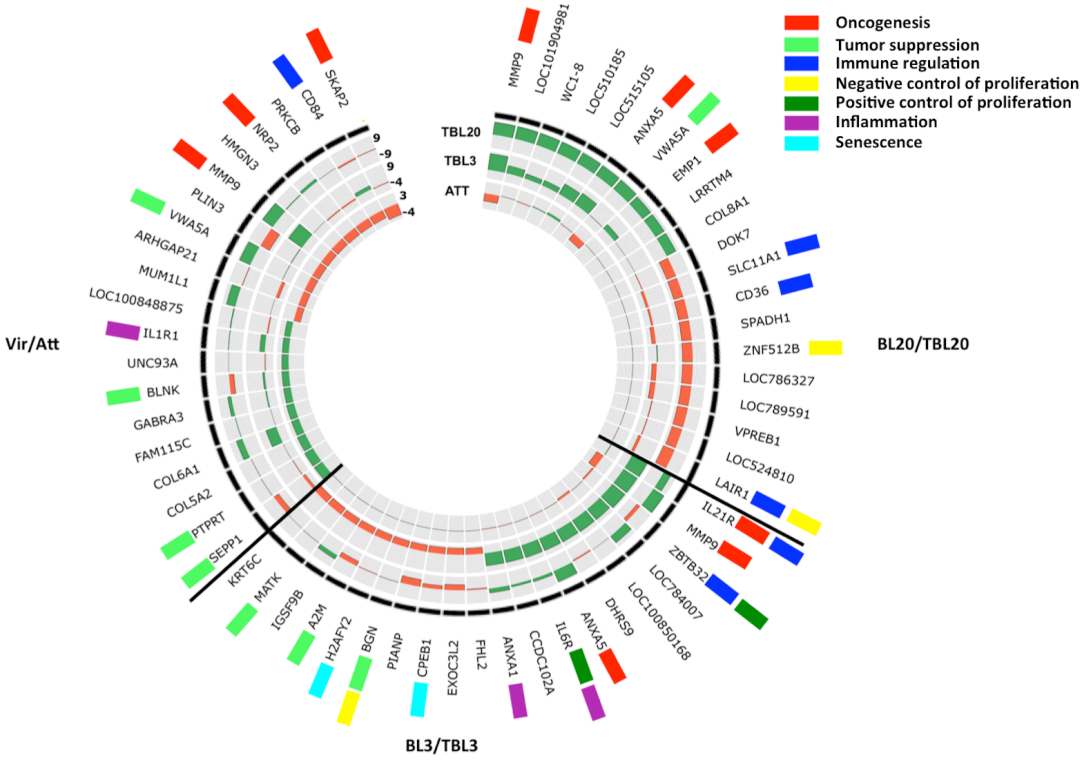
41. **Scott A, Wang Z.** 2011. Tumour suppressor function of protein tyrosine phosphatase receptor-T. Biosci Rep **31:**303-307.
42. **Short SP, Whitten-Barrett C, Williams CS.** 2016. Selenoprotein P in colitis-associated carcinoma. Mol Cell Oncol **3:**e1075094.
43. **Burk RF, Hill KE, Awad JA, Morrow JD, Kato T, Cockell KA, Lyons PR.** 1995. Pathogenesis of diquat-induced liver necrosis in selenium-deficient rats: assessment of the roles of lipid peroxidation and selenoprotein P. Hepatology **21:**561-569.
44. **Chowdhury D, Lieberman J.** 2008. Death by a thousand cuts: granzyme pathways of programmed cell death. Annu Rev Immunol **26:**389-420.
45. **Priatel JJ, Chen X, Zenewicz LA, Shen H, Harder KW, Horwitz MS, Teh HS.** 2007. Chronic immunodeficiency in mice lacking RasGRP1 results in CD4 T cell immune activation and exhaustion. J Immunol **179:**2143-2152.
46. **Somerville RP, Adamson RE, Brown CG, Hall FR.** 1998. Metastasis of Theileria annulata macroschizont-infected cells in scid mice is mediated by matrix metalloproteinases. Parasitology **116 (Pt 3):**223-228.
47. **Fell AH, Preston PM, Ansell JD.** 1990. Establishment of Theileria-infected bovine cell lines in scid mice. Parasite Immunol **12:**335-339.
48. **Hill KE, Dasouki M, Phillips JA, 3rd, Burk RF.** 1996. Human selenoprotein P gene maps to 5q31. Genomics **36:**550-551.
49. **Metheni M, Echebli N, Chaussepied M, Ransy C, Chereau C, Jensen K, Glass E, Batteux F, Bouillaud F, Langsley G.** 2014. The level of H(2)O(2) type oxidative stress regulates virulence of Theileria-transformed leukocytes. Cell Microbiol **16:**269-279.
50. **Cox AD, Der CJ.** 2003. The dark side of Ras: regulation of apoptosis. Oncogene **22:**8999-9006.
51. **Martinvalet D, Zhu P, Lieberman J.** 2005. Granzyme A induces caspase-independent mitochondrial damage, a required first step for apoptosis. Immunity **22:**355-370.
52. **Thomas DA, Massague J.** 2005. TGF-beta directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. Cancer Cell **8:**369-380.
53. **Takami M, Cunha C, Motohashi S, Nakayama T, Iwashima M.** 2018. TGF-beta suppresses RasGRP1 expression and supports regulatory T cell resistance against p53-induced CD28-dependent T-cell apoptosis. Eur J Immunol **48:**1938-1943.
54. **Fan Z, Beresford PJ, Zhang D, Xu Z, Novina CD, Yoshida A, Pommier Y, Lieberman J.** 2003. Cleaving the oxidative repair protein Ape1 enhances cell death mediated by granzyme A. Nat Immunol **4:**145-153.
55. **Peddi SR, Chattopadhyay R, Naidu CV, Izumi T.** 2006. The human apurinic/apyrimidinic endonuclease-1 suppresses activation of poly(adp-ribose) polymerase-1 induced by DNA single strand breaks. Toxicology **224:**44-55.

618 56. **Ozaki M, Suzuki S, Irani K.** 2002. Redox factor-1/APE suppresses oxidative
619 stress by inhibiting the rac1 GTPase. FASEB J **16:**889-890.
620 57. **Yang S, Irani K, Heffron SE, Jurnak F, Meyskens FL, Jr.** 2005. Alterations in
621 the expression of the apurinic/apyrimidinic endonuclease-1/redox factor-1
622 (APE/Ref-1) in human melanoma and identification of the therapeutic
623 potential of resveratrol as an APE/Ref-1 inhibitor. Mol Cancer Ther **4:**1923-
624 1935.
625 58. **Moore DH, Michael H, Tritt R, Parsons SH, Kelley MR.** 2000. Alterations in
626 the expression of the DNA repair/redox enzyme APE/ref-1 in epithelial
627 ovarian cancers. Clin Cancer Res **6:**602-609.
628 59. **Kakolyris S, Kaklamanis L, Engels K, Fox SB, Taylor M, Hickson ID,**
629 **Gatter KC, Harris AL.** 1998. Human AP endonuclease 1 (HAP1) protein
630 expression in breast cancer correlates with lymph node status and
631 angiogenesis. Br J Cancer **77:**1169-1173.
632 60. **Kim MH, Kim HB, Yoon SP, Lim SC, Cha MJ, Jeon YJ, Park SG, Chang IY,**
633 **You HJ.** 2013. Colon cancer progression is driven by APEX1-mediated
634 upregulation of Jagged. J Clin Invest doi:10.1172/JCI65521.
635 61. **Barrett CW, Reddy VK, Short SP, Motley AK, Lintel MK, Bradley AM,**
636 **Freeman T, Vallance J, Ning W, Parang B, Poindexter SV, Fingleton B,**
637 **Chen X, Washington MK, Wilson KT, Shroyer NF, Hill KE, Burk RF,**
638 **Williams CS.** 2015. Selenoprotein P influences colitis-induced tumorigenesis
639 by mediating stemness and oxidative damage. J Clin Invest **125:**2646-2660.

640 **Table 1: Biological functions of DEGs potentially playing key roles in *T. annulata*-**
641 **mediated leukocyte transformation and dissemination**

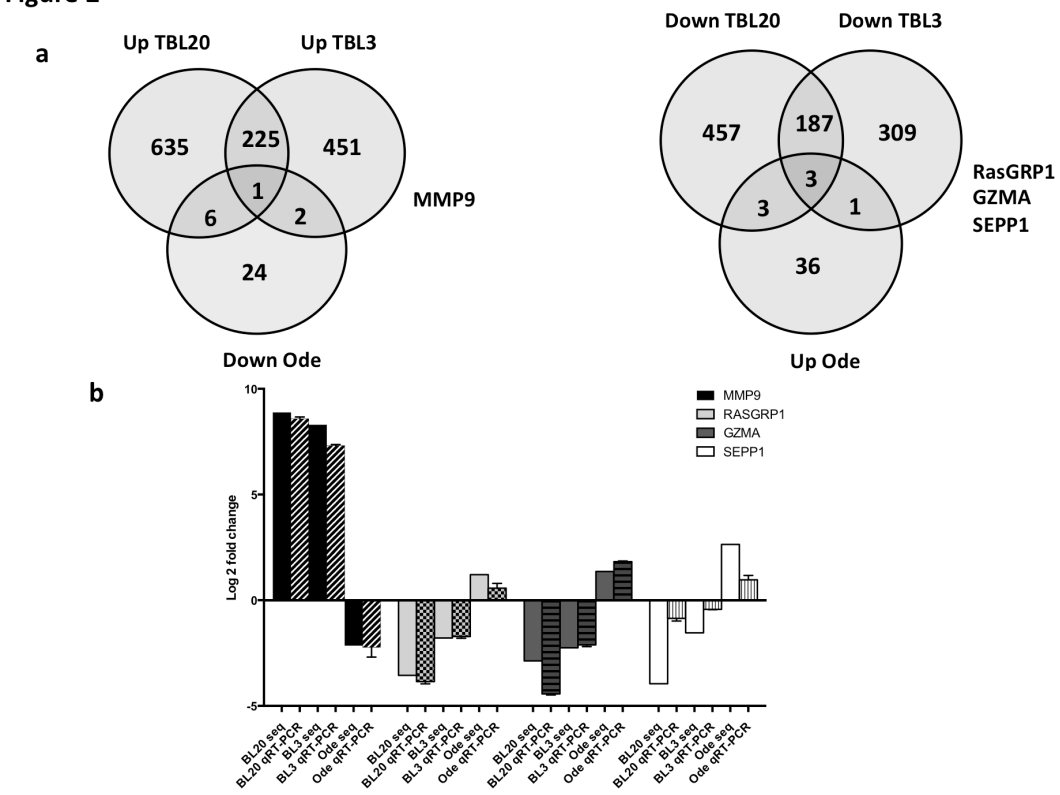
Gene symbol	Log2 FC (TBL20)	Adj p value	Log2 FC (TBL3)	Adj p value	Log2 FC (ATT)	Adj p value	Biological functions	References on biological functions
MMP9	8.87	0	8.29	0	-2.13	5.14 E-94	Metastasis formation, cancer cells invasion	(32)
SEPP1	-3.95	8.74 E-174	-1.54	3.75 E-248	2.63	1.01 E-140	Transports selenoprotein, tumor suppressor	(48), (42), (61)
GZMA	-2.87	1.45 E-05	-2.25	1.97 E-41	1.02	9.14 E-21	Plays a role in killing pathogen infected cells and cancer cells, induces increase in ROS	(44), (51)
RASGRP1	-3.55	0	-1.78	1.97 E-41	1.2	2.15 E-15	Required for correct functioning of lymphocytes in chronic infections,	(45)

Figure 1

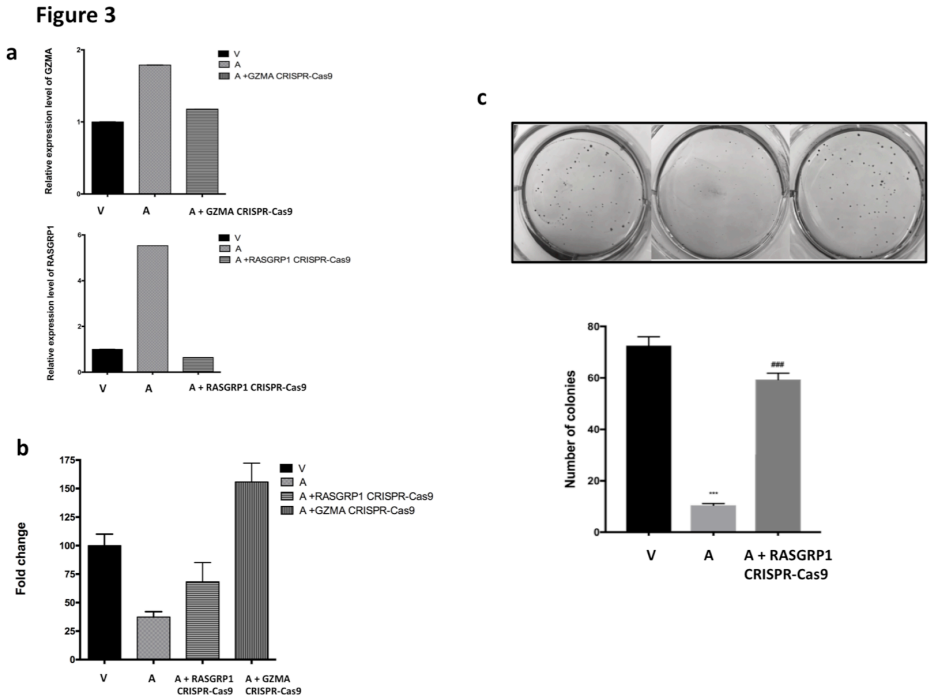


642 **Figure 1: Top 20 differentially expressed genes in Theileria-transformed bovine**
643 **host cells. A** Circos plot showing the top 10 up- and down-regulated DEGs in BL3/TBL3,
644 BL20/TBL20 and Attenuated versus Virulent (Att/Vir) Ode macrophages. The circular
645 heatmap represents the FC of the top DE genes in BL20/TBL20, BL3/TBL3, and Att/Vir
646 Ode in the outer, middle and inner rings, respectively, where green reflects the level of up-
647 regulation and red down-regulation. The genes with biological functions related to
648 tumorigenesis and immune regulation are tagged with colored rectangles. Genes with no
649 tag are hypothetical genes or have no known function in tumorigenesis and immune
650 regulation.

Figure 2



651 **Figure 2: Inversely DEGs in TBL20, TBL3 and Att Ode leukocytes.** (a) Venn diagrams
652 illustrating the genes inversely DE in TBL3, TBL20 and Attenuated Ode macrophages. (b)
653 qRT-PCR confirmation of DEGs potentially playing key roles in leukocyte transformation
654 and dissemination. The reactions were set in 3 biological replicates and the fold-change
655 calculated with the $2^{\Delta\Delta ct}$ method. The error bars represent SEM.



656 **Figure 3: Colony formation on soft agar.** (a) qRT-PCR confirmation of *GZMA* (top
657 panel) and *RASGRP1* (bottom panel) knockdown. b) Matrigel chamber assay showing a
658 regain in Matrigel traversal after *RASGRP1* and *GZMA* knockdown. c) Increased colony
659 formation in soft agar following *RASGRP1* knockdown. Non-transfected virulent
660 disseminating Ode macrophages are indicated by V, and non-transfected poorly
661 disseminating attenuated Ode macrophages by A. Error bars represent SD of 3 biological
662 replicates. *** and ### represent student t test $p<0.001$ compared to virulent and
663 attenuated Ode macrophages, respectively.

Figure 4

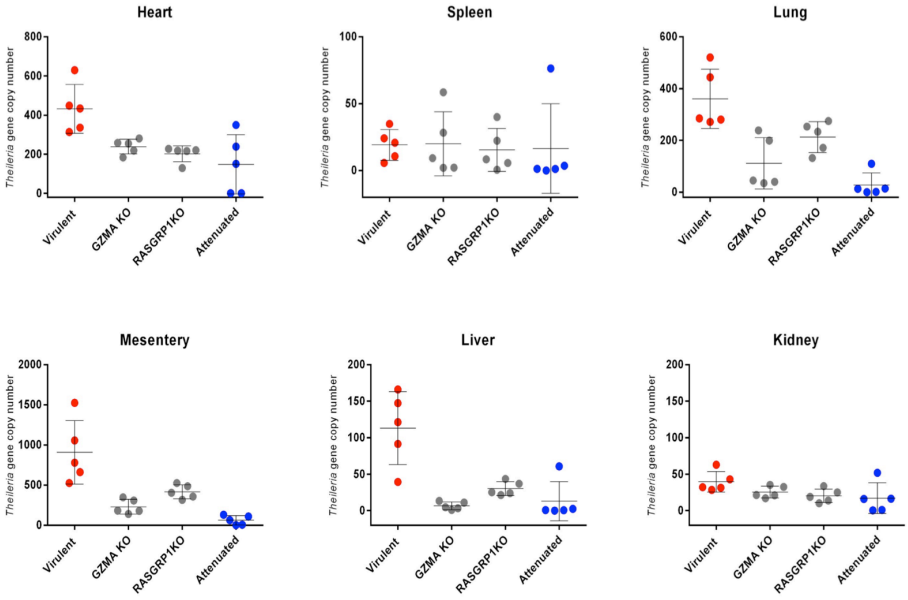


Figure 4: Effect of *GZMA* and *RASGRP1* knockdown on transformed macrophage dissemination *in vivo.* Panels represent the copy number of the single copy *T. annulata* gene (*ama-1*, TA02980) in six internal organs: heart, lung, spleen, mesentery, left kidney and liver. Transformed macrophages were injected into five Rag2γC immunodeficient mice and plotted values represent the mean of obtained *T. annulata*-specific *ama1* gene copy number. Error bars represent SD of 5 biological replicates.

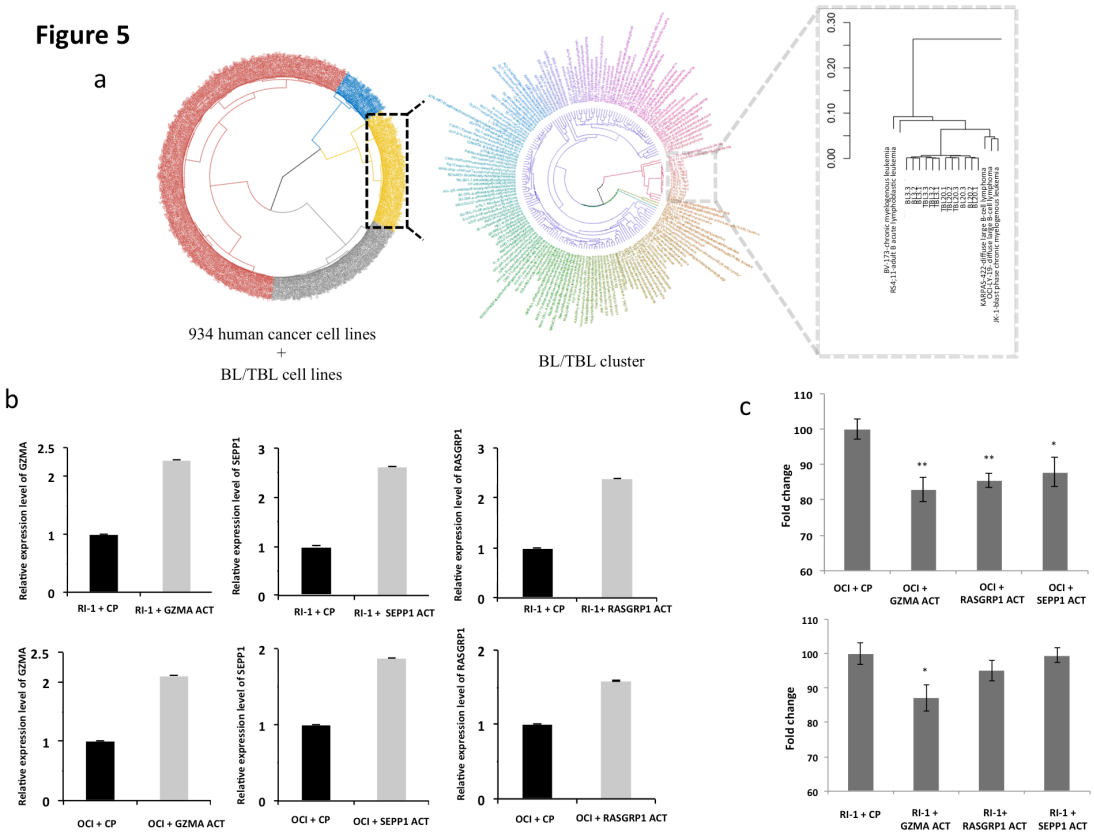
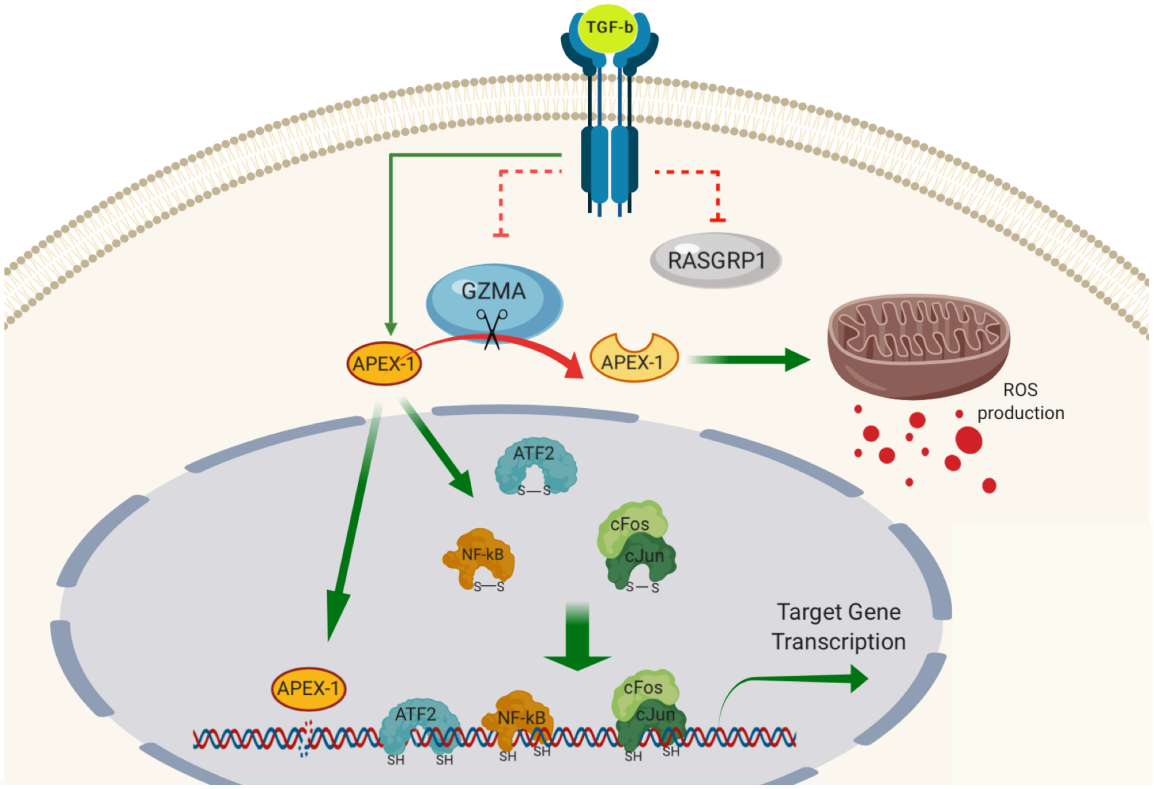


Figure 5: Effect of GZMA, RASGRP1 and SEPP1 activation on human B-lymphomas. (a) (Left panel) PCA based hierarchical clustering of 934 human cancer cell lines and *T. annulata*-transformed bovine B cells and their non-infected counterparts. The cluster containing bovine cells mainly contains leukemic human cancer cell lines. The samples are colored by cluster ID. (Middle cluster) The samples from the BL3/TBL3 and BL20/TBL20 sub-cluster were re-clustered by comparing the similarity of their gene expression profiles. The sample labels are colored by their similarity to each other. (Right panel) The sub-cluster containing bovine cell lines and human cancer cell lines with similar gene expression profile. (b) qRT-PCR determination of *GZMA*, *RASGRP1* and *SEPP1* expression after CRISPR-mediated gene activation in RI-1 (top panel) and OCI-LY19 (bottom panel). The error bars represent SEM of 3 biological replicates. (c) Matrigel chamber assay illustrating the dissemination potential of RI-1 and OCI-LY19

682 following activation of *RASGRP1* and *GZMA* transcription. Errors bars represent SEM
683 values of 3 biological replicates * and ** represent student t test $p<0.05$ and $p<0.005$,
684 respectively.

Figure 6



685 **Figure 6: Scheme of proposed TGF/GZMA/RASGRP1/APEX1 pathway**

In virulent hyper-disseminating macrophages enhanced TGF-β2 signaling results in upregulated *APEX1* transcription and suppressed *GZMA/RASGRP1* transcription. By contrast, in macrophages attenuated for dissemination reduced TGF-β2 signaling leads to lower *APEX1* transcription and increased in *GZMA* expression. Higher levels of GZMA cleave lower levels of APEX-1 ablating its oxidative repair functions leading to increased ROS output that typifies attenuated macrophages. In virulent, macrophages TGF-β2–induced APEX-1 reduces oxidized forms of target transcription factors such as NF-κB, ATF2, cFos and cJun to increase their DNA binding activity that augments transcription of their target genes to promote cellular transformation and tumor dissemination.

Supplementary figure legends

Figure S1: Sequencing quality of all samples.

(a) Clustering of all samples. (b) Density plot representing FPKM distribution of all samples.

Figure S2: Differentially expressed genes in TBL20, TBL3 and Attenuated Ode leukocytes.

(a) Histogram showing the number of up- (dark grey) and down- (light grey) regulated genes in all 3 datasets. The area-proportional Venn diagrams represent the intersection between the lists of DEGs from CuffDiff2 (green), DESeq2 (black) and baySeq (blue). The intersection between the 3 pipelines reflects the number of up- and down-regulated genes. The list of DEGs is listed in Table S2. (b) qRT-PCR confirmation of randomly selected genes in TBL20 and TBL3. The reactions were set in 3 biological replicates and the fold change calculating with the $2^{\Delta\Delta ct}$ method. The error bars represent SEM..

Figure S3: Log2FC values of RASGRP1-4 in TBL20 and TBL3

RNAseq Log2FC values from DESeq2 of TBL20 and TBL3 compared to BL20 and BL3, respectively.

Figure S4: Effect of *GZMA* knockdown on H_2O_2 output.

H_2O_2 output by virulent (V), attenuated (A), and attenuated Ode macrophages after CRISPR/Cas9-mediated *GZMA* knockdown. Error bars represent SD of 3 biological replicates. ** and ## represent $p<0.01$ compared to virulent and attenuated Ode macrophages, respectively.

Figure S5: Effect of *RASGRP1* knockdown on *GZMA* expression.

qRT-PCR of GZMA in virulent (V), attenuated (A) and attenuated Ode macrophages after CRISPR/Cas9-mediated *RASGRP1* knockdown. Error bars represent SD of 3 biological replicates *** and ### represent p<0.001 compared to virulent and attenuated Ode macrophages, respectively.

Figure S6: Schematic of *T. annulata* host cells and human cancer cell lines transcriptome clustering

Figure S7: Complete cluster of 934 human cancer cell lines and BL20/TBL20, BL3/TBL3

Table S1: Top 5 up- and down-regulated DEGs in infected and attenuated cell lines.

This list was generated based on FC values of DEGs listed in Table S2.

Table S2: List of DEGs in infected and attenuated cell lines.

For details on the methods used to generate this list, please see in the materials and methods section.

Table S3: List of qRT-PCR primers.

Table S4: List of DEGs in attenuated Ode macrophages after TGF-β2 treatment.

Genes listed as DE from DESeq2