Abstract

Objectives: Human T cell leukemia virus-1 (HTLV-1) infection may lead to one or both diseases including HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) or adult T cell leukemia lymphoma (ATLL). The complete interactions of the virus with host cells in both diseases is yet to be determined. This study aims to construct an interaction network for distinct signaling pathways in these diseases based on finding differentially expressed genes (DEGs) between HAM/TSP and ATLL.

Results: We identified 57 hub genes with higher criteria scores in the primary protein–protein interaction network (PPIN). The ontology-based enrichment analysis revealed following important terms: positive regulation of transcription from RNA polymerase II promoter, positive regulation of transcription from RNA polymerase II promoter involved in meiotic cell cycle and positive regulation of transcription from RNA polymerase II promoter by histone modification. The upregulated genes TNF, PIK3R1, HGF, NFKBIA, CTNNB1, ESR1, SMAD2, PPARG and downregulated genes VEGFA, TLR2, STAT3, TLR4, TP53, CHUK, SERPINE1, CREB1 and BRCA1 were commonly observed in all the three enriched terms in HAM/TSP vs. ATLL. The constructed interaction network was then visualized inside a mirrored map of signaling pathways for ATLL and HAM/TSP, so that the functions of hub genes were specified in both diseases.

Keywords: Human T-lymphotropic virus type 1, HTLV-1 associated myelopathy/tropical spastic paraparesis, Adult T-cell leukemia/lymphoma, Pathogenesis, Systems virology

Introduction

Infection with Human T cell leukemia virus-1 (HTLV-1) which normally takes an innocuous and insidious course in humans can rarely manifest as adult T cell leukemia/lymphoma (ATLL) and/or HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP). HAM/TSP results in lower back pain, limb spasticity, progressive neurological decline, and urinary disturbances [1]. Contrastingly, ATLL is characterized by unrestrained growth of T-cell precursors in blood, bone marrow, thymus, or lymph nodes. Despite thorough investigations, it is still unclear what triggers HTLV-1 infection to remain innocuous or to progress to either of these complications [2, 3]. A thorough examination of the pathogenic pathways and the disruptions in host proteins and their interactions might therefore reveal novel aspects of the disease through identification of unbalanced pathways in HTLV-1 related diseases. Identifying the key dysregulated pathways could be utilized to screen individuals...
with higher risk and reveal novel therapeutic approaches to prevent the progression to either of these diseases. Analyzing the cellular gene expression signature in both of these complications highlights common and specific deregulated pathways in HTLV-1 related diseases which are likely to provide novel insight into the pathogenesis and the management of HTLV-1 related diseases.

Several analytical approaches and softwares have been developed (https://pubs.acs.org/doi/abs/10.1021/acs.jproteome.5b01080, https://academic.oup.com/biotechnology/article-abstract/35/8/1436/5102873, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0189922) [4] as means to further analyze the high-throughput data derived from previous genomic studies to extract and conclude additional results which would only be available when the data from different studies were compared with each other. A well-known method to analyze previous datasets is to search for differentially expressed genes (DEGs) between normal individuals and those harboring the disease or between several complications associated with a single pathogen [5, 6].

In this study we performed the protein–protein interaction network-based analysis to determine DEGs between HAM/TSP and ATLL samples. The resulting networks and related hub genes were enriched in gene ontology for biological processes (BP) and the results and their implication is discussed.

Main text
Methodology
Microarray dataset
The gene expression profile GSE19080 was acquired from the public repository database gene expression omnibus (GEO) (www.ncbi.nlm.nih.gov/geo) which includes an individual platform, GPL9686. The dataset contains the results of microarray experiments using the human ImmuneArray cDNA array. The data related to ATLL (7 specimens) and HAM/TSP (12 specimens) patients were extracted and then analyzed.

Exploration of differentially expressed genes
The Data was first normalized and preprocessed with log2 transformation. Then, the differentially expressed genes (DEGs) and their value of fold changes (FC) were acquired by GEO2R, which is an interactive web tool based on GEOquery and limma packages in R computing language. The adjusted P-value < 0.05 (calculated by FDR) was selected as the criterion for selection of DEGs. The direction of dysregulation of each for DEGs was reported as upregulation (positive logFC) and downregulation (negative logFC) compared to baseline.

Construction of protein–protein interaction (PPI) network
In order to construct the PPI network, the online STRING (Search Tool for the Retrieval of Interacting Genes) database version 10.5 was employed [7]. The information and interactions from various biological sources, including physical interactions, functional association, high-throughput experiments, genomic context, co-expression, databases, and text-mining were considered. The cut-off criterion was set at combined score > 0.4 to analyze the PPINs.

Identification of hub genes
The network was analyzed by Network Analyzer app in Cytoscape (3.5.1) to calculate “degree” and “betweenness” centrality measures. The number of edges of a node is assigned as the degree [8] and the number of node visiting during moving all shortest paths is defined as betweenness centrality [9]. The genes with higher degree and betweenness scores were selected as hub genes. These genes and their associated PPIN were visualized using Gephi version 0.9.1 [10].

Gene ontology analysis
Gene ontology BP was assessed by Enrichr website [11]. Top ten major functional terms were selected based on z-scores for further analysis.

Signaling network analysis
The HTLV-1-implicated signaling network was depicted based on the KEGG and WikiPathway databases. The upregulated and downregulated genes were presented with red and blue coloring in visualized pathways, respectively.

Results
Identification of DEGs and hub genes
The number of 1116 DEGs was recognized according to FDR < 0.05 by Benjamini–Hochberg procedure. Following the construction of the primary network by STRING, the network was analyzed by means of degree and betweenness centrality measures. 57 genes (Table 1) were selected as hub genes based on the aforementioned criteria. The logFC for each hub gene is reported in Additional file 1: Figure S1.

PPIN construction
The relationship between hub genes was specified using STRING database. The network was constructed, as shown in Additional file 2: Figure S2. The network consists of 57 nodes and 716 edges. The size of each node
was determined based on its degree value and the color of each node was specified according to their direction of deregulation with higher values of logFC represented by colors closer to red and lower values by blue.

**Gene ontology biological processes**
The upregulated and downregulated hub genes were individually enriched in gene ontology BP and the top functional terms were selected. Additional file 3: Figure S3a, b reveals the BP of hub genes associated with upregulated and downregulated genes, respectively. The comparison between the Gene Ontology BP of the enriched upregulated and downregulated genes revealed that positive regulation of transcription from RNA polymerase II promoter (GO: 0045944), positive regulation of transcription from RNA polymerase II promoter involved in meiotic cell cycle (GO: 0010673) and positive regulation of transcription from RNA polymerase II promoter by histone modification (GO: 1903757) are common in the samples of this analysis. Figure 1 is a circos plot demonstrating the contribution amount of proteins in each of the GO BPs (GO: 0045944, GO: 0010673 and GO: 1903757). The upregulated genes including TNF, PIK3R1, HGF, NFKBIA, CTNNB1, ESR1, SMAD2, PPARG and downregulated genes comprised of VEGFA, TLR2, STAT3, TLR4, TP53, CHUK, SERPINE1, CREB1 and BRCA1 were commonly observed in all of the three enriched terms in TSP vs. ATLL.

**Table 1** List of the up-regulated (positive logFC) and down-regulated (negative logFC) hub genes in TSP vs. ATLL

| Gene   | TSP logFC | ATLL logFC |
|--------|-----------|------------|
| TNF    | 2.08      | -0.3       |
| ITGA2  | 1.97      | -0.38      |
| PIK3R1 | 1.51      | -0.41      |
| CALM1  | 1.41      | -0.42      |
| PTK2   | 1.27      | -0.46      |
| HGF    | 1.25      | -0.47      |
| NFkBIA | 1.22      | -0.61      |
| PIK3CD | 1.11      | -0.64      |
| PRKCB  | 1.06      | -0.65      |
| EZH2   | 0.99      | -0.69      |
| CTNNB1 | 0.9       | -0.71      |
| GMPS   | 0.88      | -0.75      |
| IL8    | 0.81      | -0.76      |
| CAD    | 0.8       | -0.76      |
| YWHAZ  | 0.77      | -0.79      |
| PPP2CA | 0.76      | -0.8       |
| NOS2   | 0.74      | -0.83      |
| ESR1   | 0.74      | -0.85      |
| SMAD2  | 0.72      | -0.86      |
| RB1    | 0.68      | -0.91      |
| NGF    | 0.66      | -0.93      |
| MAPK14 | 0.63      | -0.98      |
| PTGS2  | 0.62      | -0.98      |
| PPARG  | 0.6       | -0.99      |
| CBL    | 0.58      | -1.25      |
| PARP1  | 0.51      | -1.49      |
| XPO1   | 0.48      | -1.7       |
| PHLPP2 | 0.47      |            |
| CDK62A | 0.4       |            |
| FN1    | 0.4       |            |

Discussion
The results of this analysis indicate high dissimilarities in regulation of cell proliferation and inflammatory pathways in ATLL and HAM/TSP. While the enriched pathways regarding RNA polymerase II [12] and DNA replication in this study just allude to a surge in clonal expansion or increased viral transcription of TATA-box containing 5‘LTR of HTLV-1 [13] in the provirus-harborng cells in ATLL, the identified hub genes provide further insight into the pathogenesis of both diseases.

The SRC gene is highly overexpressed in a variety of human cancers [14]. In this study, however, ATLL was associated with lower levels of SRC expression. Counter-intuitively, the loss of src homology 2 containing tyrosine phosphatase (SHP-1) is associated with spontaneous activation in HTLV-1 infected T-cells [15]. Furthermore, another src-like tyrosine kinase LYN, which is normally
upregulated via Tax in HTLV-1 cell lines [16], is down-regulated in this study compared to HAM/TSP. This suggests that the increased LYN expression is not maintained with the loss of Tax expression in later stages of ATLL development [17]. Further studies are needed to determine the role of src-like tyrosine kinases in ATLL and HAM/TSP.

One of the major pathways in ATLL is lymphocyte activation, which is of particular importance in ATLL compared to ACs and HAM/TSP [6]. Transduction of lymphocyte activation signals occurs through the following ways: (i) activation of NF-kappa B, (ii) increase the AP1 gene expression via up-regulation of TNF [18], and (iii) enhancement the expression of NFAT gene due to up-regulation of CALM1 [19]. The diminished expression of the aforementioned genes in HAM/TSP indicates their lesser significance compared to ATLL. Therefore, it could be speculated that NF-κB pathway is likely the major signaling pathways in ATLL compared to other forms of HTLV-1 infection, as inhibition of NF-κB by a super-repressor form of IκBα (SR-IκBα) in infected T-cells in ATLL results in cell death regardless of Tax expression [20]. Furthermore, there was a significant disruption of genes related to phosphatidylinositol 3-kinase-Protein kinase B (PI3K-AKT) pathway in ATLL (PIK3CD, PIK3R1, and IRS1) which has been demonstrated to have a role in activation of NF-κB pathway [21].

Several genes like FN1, ITGA, PTK2, HGF, and CBL can affect PI3K-AKT pathway in the focal adhesion and ECM Receptor interaction pathways [22–24]. Moreover, PI3K-AKT itself is indirectly activated through the ROS pathway. This pathway can be stimulated through upregulation of PRKCB due to functions of NGF and HGF which are in turn mediated by PLC-γ and PLC-β, respectively [25–27]. The increase in NF-κB action is also important in upregulation of PTGS2, TNF, NGF, IL8, and

Fig. 1 Circos plot of the common GO biological process between up-regulated and down-regulated genes in TSP vs. ATLL. The upregulated genes including TNF, PIK3R1, HGF, NFKBIA, CTNNB1, ESR1, SMAD2, PPARG and downregulated genes containing VEGFA, TLR2, STAT3, TLR4, TPS3, CHUK, SERPINE1, CREB1 and BRCA1 were common in all the three enriched terms.
NFKBIA in ATLL [25, 28–31], as well as NOS2 [32, 33]. The NFKBIA increase in ATLL cells however, supposedly inhibit NF-κB pathway [28]. However, pathways such as PARP-1 that lead to NFKBIA degradation, are activated in ATLL to overcome the inhibition. The NF-κB pathway is less prominent in HAM/TSP, as the effective genes mentioned above are inversely expressed in it.

Inhibition of apoptosis is another pivotal distinction in signaling pathways between ATLL and HAM/TSP, marked by significant dowregulation of TP53 in ATLL [6]. TP53 loss or mutation in ATLL cells reflects their aggressive proliferation and poor prognosis as described in many other cancers [34, 35]. The upregulation of SMAD2 in ATLL facilitates metastasis and is associated with TP53 mutation [36].

Furthermore, MDM2 can increase the expression of CDKN2A, which in turn increases TP53 [37]. Although, upregulation of CDKN2A secondary to the induction of oncogenes is unavoidable, EZH2 functionally inhibits the role of CDKN2A in cell cycle [38]. The decrease in expression level of TP53 is accompanied by downregulation of its target genes including Serpin1 in ATLL. Serpin1 has been linked to reduction of tumor invasion and growth [39]. In addition, the downregulation of EGR1 tumor repressor is also observed in ATLL in this study. The increase in the expression level of EGR1 in HAM/TSP along with the higher levels of TP53 and Serpin1 compared to ATLL restricts the uncontrolled proliferation of infected cells as seen in ATLL [40]. Furthermore, the analysis revealed PRKCB, ITGA2, IL8, and NOS2 genes to be significantly upregulated in ATLL. The aforementioned genes are associated with cell motility and angiogenesis and could be surmised to contribute to the metastatic properties of leukemic cells in ATLL.

Viral persistence cannot continue without significant dysregulation of the host immune response. The balance of immune response can tip to either immunosuppression or hypersensitivity in HTLV-1 infections, resulting in ATLL and HAM/TSP, respectively. The mild immunodeficiency almost exclusively associated with ATLL is partly the consequence of disruption of regulatory genes implicated in the immune response. In this study, the expression levels of STAT3, TLR2/4,
and MJ performed bioinformatics analysis. MZ‑G, S‑HM, and MT‑R interpreted the review. All authors read and approved the final manuscript.

Authors’ contributions
MZ‑G, S‑HM, MN, SY1, SY2 and RF designed the project. MZ‑G, S‑HM, MS, MF, and MJ performed bioinformatics analysis. MZ‑G, S‑HM, and MT‑R interpreted and wrote the manuscript. MT‑R and SB did systematic search and literature review. All authors read and approved the final manuscript.

Limitations
Further detailed studies help us understand other functions of the involved genes in the pathogenesis of HTLV‑1.

Acknowledgements
Many thanks to the Vice Chancellor for Research, Alborz University of Medical Sciences for supporting the study.

References
1. Bangham CR, Araujo A, Yamano Y, Taylor GP. HTLV‑1-associated myelopathy/tropical spastic paraparesis. Nat Rev Dis Prim. 2015;1:15012.
2. Bangham CR, Ratner L. How does HTLV‑1 cause adult T‑cell leukaemia/lymphoma (ATL)? Curr Opin Virol. 2015;14:93–100.
3. Bangham CRM, Human T. Cell leukemia virus type 1: persistence and pathogenesis. Annu Rev Immunol. 2018;36:43–71.
4. Leung MK, Delong A, Alipanahi B, Frey BJ. Machine learning in genomic medicine: a review of computational problems and data sets. Proc IEEE. 2016;104(1):176–97.
5. Azimzadeh Jamalkandi S, Mozhgani SH, Gholami Pourbadie H, Mirzaie M, Noorbakhsh F, Vaziri B, et al. Systems biomedicine of rabies delineates the affected signaling pathways. Front Microbiol. 2016;7:1688.
6. Mozhgani SH, Zarei‑Ghobadi M, Meymoori‑Rad M, Mokhtari‑Azad T, Mirzaie M, Sheikh M, et al. Human T‑lymphotropic virus 1 (HTLV‑1) pathogenesis: a systems virology study. J Cell Biochem. 2018;119(5):3968–79.
7. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta‑Cepas J, et al. STRING v10: protein–protein interaction networks, integrated over weighted biological networks based on edge‑betweenness centrality. Nucleic Acids Res. 2014;43(D1):D447–52.
8. Diestel R. Graph theory. Grad Texts in Math. 2005;101.
9. Yoon J, Blumer A, Lee K. An algorithm for modularity analysis of directed and weighted biological networks based on edge-betweenness centrality. Bioinformatics. 2006;22(4):496–505.
10. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. Icwm. 2009;8:361–2.
11. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1):W90–7.
12. Butler JE, Kadonaga JT. The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev. 2002;16(20):2583–92.
13. Yoshida M, Satou Y, Yasunaga J, Fujisawa J, Matsuoka M. Transcriptional control of spliced and unspliced human T‑cell leukemia virus type 1 bZIP factor (HBZ) gene. J Virol. 2008;82(19):9359–68.

Funding
This study was funded by Alborz University of Medical Sciences (ABZUMS).

Availability of data and materials
The dataset analyzed in this study (GSE19080) was accessed through gene expression omnibus public repository (www.ncbi.nlm.nih.gov/geo).

Declarations
Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher details
Received: 9 December 2020   Accepted: 11 March 2021
Published online: 23 March 2021
14. Irby RB, Yeatman TJ. Role of Src expression and activation in human cancer. Oncogene. 2000;19(49):5636–42.

15. Migone TS, Cacalano NA, Taylor N, Yi T, Waldmann TA, Johnston JA. Recruitment of SH2-containing protein tyrosine phosphatase SHP-1 to the interleukin 2 receptor: loss of SHP-1 expression in human T-lymphotropic virus type I-transformed T cells. Proc Natl Acad Sci USA. 1998;95(7):3845–50.

16. Uchiumi F, Semb a K, Yamashita Y, Fujisawa J, Yoshida M, Inoue K, et al. Characterization of the promoter of the human family gene lyn and its trans activation by human T-cell leukemia virus type I-encoded p40tax. Mol Cell Biol. 1992;12(9):3784.

17. Grassmann R, Aboud M, Jeang K-T. Molecular mechanisms of cellular transformation by HTLV-1 Tax. Oncogene. 2005;24(39):5976–85.

18. Qiao Y, He H, Jonsson P, Sinha I, Zhao C, Dahlman-Wright K. AP-1 is a key regulator of proinflammatory cytokine TNFalpha-mediated triple-negative breast cancer progression. J Biol Chem. 2016;291(10):5068–79.

19. Bioczek T, Lisek M, Ferenc B, Zylinska L. Cross talk among PMCA, calcineurin and NFAT transcription factors in control of calmodulin gene expression in differentiating PC12 cells. Biochem Biophys Acta. 2017;1860(4):502–15.

20. Hironaka N, Mochida K, Mori N, Maeda M, Yamamoto N, Yamaoka S. Boczek T, Lisek M, Ferenc B, Zylinska L. Cross talk among PMCA, calcineurin and NFAT transcription factors in control of calmodulin gene expression in differentiating PC12 cells. Biochem Biophys Acta. 2017;1860(4):502–15.

21. Bai D, Ueno L, Vogt P. Akt-mediated regulation of NFkappaB and the essentialness of NFkappaB for the oncogenicity of PI3K and Akt. Int J Cancer. 2009;125(12):2863–70.

22. Xia H, Nho RS, Kahn J, Kleidon J, Henke CA. Focal adhesion kinase is upstream of phosphatidylinositol 3-kinase/Akt/eNOS/NO pathway. J Biol Chem. 2004;279(31):33024–34.

23. Liyayova MS, Ma K, Lipkowitz S. Molecular pathways: cbl protein family. J Clin Oncol. 2015;21(18):1789–94.

24. Kim MS, Shuto LP, Granaskearan A, Lin Z, Rystedt JE, Ulrich JD, et al. Nerve growth factor (NGF) regulates activity of nuclear factor of activated T-cells (NFAT) in neurons via the phosphatidylinositol 3-kinase (PI3K)-Akt-glycogen synthase kinase 3beta (GSK3beta) pathway. J Biol Chem. 2014;289(45):31349–60.

25. Chou WY, Chuang KH, Sun D, Lee YH, Kao PH, Lin YY, et al. Inhibition of PKC-induced COX-2 and IL-8 expression in human breast cancer cells by glucosamine. J Cell Physiol. 2015;230(9):2240–51.

26. Kondratyev A, Popova O, Severin S, Choladze M, Shmyrev I, Tubasheva I, et al. Nerve growth factor stimulates calcium-dependent protein kinase C translocation in PC12 cells. FEBS Lett. 1990;264(1):75–7.

27. Rusanescu G, Gotch T, Tian X, Feig LA. Regulation of Ras signaling specificity by protein kinase C. Mol Cell Biol. 2001;21(8):2650–8.

28. Hoesei B, Schmid JA. The complexity of NF-kappaB signaling in inflammation and cancer. Mol Cancer. 2013;12:86.

29. Manna SK, Ramesh GT. Interleukin-8 induces nuclear transcription factor-kappaB through a TRAF6-dependent pathway. J Biol Chem. 2005;280(8):7010–21.

30. Ahn SH, Park H, Ahn YH, Kim S, Cho MS, Kang JI, et al. Necrotic cells influence migration and invasion of glioblastoma via NF-kappaB/AP-1-mediated IL-8 regulation. Sci Rep. 2016;6:24552.