The transcription factor and basic helix-loop-helix family member e41 BHLHE41 is differentially expressed and transcriptionally induced in models of coronavirus infection.

Shahan Mamoor¹

¹Thomas Jefferson School of Law

San Diego, CA 92101

mamoorsk@tjssl.edu

The coronavirus COVID19 pandemic is an emerging biosafety threat to the nation and the world (1). There are no treatments approved for coronavirus infection in humans (2) and there is a lack of information available regarding the basic transcriptional behavior of human cells and mammalian tissues following coronavirus infection. We mined two independent datasets (3, 4), public (3) and published (4) containing transcriptome data from infection models of the Middle East respiratory syndrome (MERS) coronavirus and human coronavirus (HCoV) to discover genes that are differentially expressed in coronaviruses and identify potential therapeutic targets and host cell vulnerabilities. We identified the transcription factor basic helix-loop-helix family member e41 BHLHE41 as a conserved differentially expressed gene following coronavirus infection. BHLHE41 may be involved in the cellular response to COVID19 infection.

Keywords: BHLHE41, coronavirus, MERS coronavirus, human coronavirus, HCoV, systems biology of viral infection, COVID19.

Viruses are classified according to a system known as the “Baltimore” classification of viruses (5) wherein the characteristics of the viral genome, including whether it is plus or minus strand, also known as positive-sense or negative-sense, whether the genome is single-stranded or double-stranded, whether their genome is composed of RNA or DNA, and whether or not they use a reverse transcriptase to replicate are used to group viruses. Coronaviruses are plus-strand RNA viruses that contain an envelope surrounding their viral particle (6). Their genome is largest of all RNA viruses, ranging from 27 to 33 kb in size (7). They obtain their name from the crown-like appearance of the viral particle imparted by the structure of the large-surface glycoprotein (7). The coronaviridae family includes seven viruses capable of infecting humans, including the severe acute respiratory distress syndrome, or SARS coronavirus (8), the Middle East respiratory syndrome coronavirus, or MERS coronavirus (9), the human coronaviruses (HCoV) 229E, OC43, HKU and NL63 (10-12), and the novel human coronavirus now known as COVID19 (13, 14). As of March 19, 2020, the World Health Organization reported 209,839 cases of COVID19 and 8778 deaths from COVID19 infection world-wide (15). There are no FDA-approved treatments for human coronavirus infection.

We used a systems-level approach to identify the genes whose expression changes most significantly following infection with multiple types of coronavirus, including MERS-CoV and HCoV 229E using two independent datasets (3, 4). From these datasets were mined transcriptome data following coronavirus infection in human cell culture and in primary human endothelial cells. Across both of these datasets, we identified the transcription factor BHLHE41 as among the genes most differentially expressed following coronavirus infection. BHLHE41 represents a transcriptional target of the host cell gene expression program following infection with coronaviruses in human cells.

Methods

We used datasets GSE100509 (3) and GSE89167 (4) for this systems-level differential gene expression analysis of coronavirus infections in conjunction with GEO2R.

The Benjamin and Hochberg method of p-value adjustment was used for ranking of differential expression but raw p-values were used for assessment of statistical significance of global differential expression. Log-transformation of data was auto-detected, and the NCBI generated category of platform annotation was used.

A statistical test was performed to evaluate the significance of difference between mRNA expression levels of BHLHE41 in HuH-7 cells with and without HCoV 229E infection using a two-tailed, unpaired t-test. A statistical test was performed to evaluate the significance of difference between mRNA expression levels of BHLHE41 in human primary microvascular endothelial cells using a one-way ANOVA with multiple comparisons, all compared to baseline infection at 0 hours. Only p-values less than 0.05 were considered statistically significant. We used PRISM for all statistical analyses (Version 8.4.0)(455).

Results

We mined two independent microarray datasets, public (3) and published (4) containing transcriptome data from models of coronavirus infection in primary human cells and cell culture. Across experimental models, we identified the transcription factor BHLHE41 as differentially expressed following coronavirus infection.

BHLHE41 is differentially expressed in primary human microvascular endothelial cells when comparing cells infected with wild-type Middle East respiratory syndrome coronavirus (MERS-CoV), icMERS-CoV EMC2012 and uninfected cells.

We identified BHLHE41 as differentially expressed following infection of primary human microvascular endothelial cells with wild-type Middle East respiratory syndrome coronavirus (MERS-CoV), icMERS-CoV EMC2012 when compared to non-infected cells (Table 1) (3). When sorting all of the transcripts expressed in human microvascular endothelial cells measured by microarray based on change in expression with and without infection, BHLHE41 ranked 13 out of 34127 transcripts. Differential expression of BHLHE41 in primary human microvascular

endothelial cells following infection with MERS-CoV was statistically significant (Table 1; $p=2.59E-30$).

BHLHE41 is differentially expressed in the human cell line HuH-7 when comparing cells with human coronavirus 229E and uninfected cells.

We also identified BHLHE41 as differentially expressed in the human cell line HuH-7 when comparing cells with human coronavirus 229E with uninfected cells (Table 2) (4). When sorting all of the transcripts expressed in human HuH-7 cells measured by microarray based on change in expression with and without infection with human coronavirus 229E, BHLHE41 ranked 94 out of 62976 transcripts. Differential expression of BHLHE41 in HuH-7 following infection with HCoV 229E was statistically significant (Table 2; $p=0.00020815$).

BHLHE41 is transcriptionally induced following infection of primary human microvascular endothelial cells with wild-type Middle East respiratory syndrome coronavirus (MERS-CoV), icMERS-CoV EMC2012.

We extracted exact mRNA expression values for BHLHE41 from primary human microvascular endothelial cells infection with wild-type Middle East respiratory syndrome coronavirus (MERS-CoV), icMERS-CoV, and from uninfected primary human microvascular endothelial cells in order to compare expression levels of BHLHE41 between these two groups rather than relative to the rest of the transcriptome as assessed in differential gene expression analysis. This dataset contained transcriptome information from infection of primary human microvascular endothelial cells at 12 hours, 24 hours, 36 hours and 48 hours post-infection at compared to baseline (0 hours). We also performed a statistical test to evaluate whether the difference in expression of BHLHE41 in primary human microvascular endothelial cells infection with and without MERS-CoV infection was statistically significant, BHLHE41 was expressed at significantly higher levels higher levels at 12 hours, 24 hours, 36 hours, and 48 hours post-infection with MERS-CoV as compared to cells infected at baseline (0 hours), reaching a peak at 24 hours (Figure 1; $p<0.0001$ for all comparisons).

BHLHE41 is transcriptionally induced in the human cell line HuH-7 following infection with human coronavirus 229E.

We also extracted exact mRNA expression values for BHLHE41 from HuH-7 cells following infection with human coronavirus 229E and from uninfected HuH-7 cells in order to compare expression levels of BHLHE41 between these two groups rather than relative to the rest of the transcriptome as assessed in differential gene expression analysis. We also performed a statistical test to evaluate whether the difference in expression of BHLHE41 in HuH-7 cells following infection with human coronavirus 229E as compared to uninfected HuH-7 cells was statistically significant. BHLHE41 was expressed at significantly higher levels in HuH-7 cells following infection with HCoV 229E (Figure 2; $p=0.0110$).

Discussion

COVID19 is a newly discovered coronavirus that infects humans and whose spread has lead to a global pandemic (19) with zero available therapeutic strategies. To facilitate target discovery and to further understanding of the basic transcriptional program of cells and tissues following coronavirus infection, we mined two independent microarray datasets containing transcriptome data of human cells infected with MERS-CoV and HCoV 229E, in primary human cells and in human cell culture. Across both experimental models, we identified the transcription factor BHLHE41 as one of the genes whose expression changed most significantly following infection with a coronavirus. In both experimental models analyzed, BHLHE41 expression significantly increased after infection with coronavirus.

BHLHE41 is also known as Dec2 and has multiple important roles in both the adaptive and innate immune system that may be relevant in coronavirus pathogenesis (16-19). BHLHE41 is required for commitment to the T_H2 lineage of T-lymphocytes and promotes production of the cytokines IL-4, IL-5 and IL-13 (16). BHLHE41 is important for the development and self-renewal of the B-1a subset of B-lymphocytes (18). In the innate immune system, BHLHE41 regulates the identity and self-renewal of alveolar macrophages through

1 directly binding the genomic locus and suppressing the expression of lineage-inappropriate
2 genes in alveolar macrophages (19). BHLHE41 also has functions in the DNA repair system
3 though its ability to suppress expression of the MLH 1 mismatch repair enzyme (20, 21).

4 We found that BHLHE41 was differentially expressed in human cells following infection
5 with human coronavirus as well as in human cells after infection with Middle East respiratory
6 syndrome coronavirus. Induction of BHLHE41 expression by coronaviridae may be important
7 for viral infection of human host cells by transactivating a network of genes controlled by a
8 single transcription factor that together support enhanced viral infectivity. Alternatively or in
9 conjunction, the physiological roles of BHLHE41 may important indirectly through its ability to
10 control the differentiation of the T_H2 subset of T-lymphocytes (16), the B-1a subset of B-
11 lymphocytes (18), and of tissue-resident innate immune cells, lung alveolar macrophages (19).
12 BHLHE41 expression should be evaluated in the cells and tissues of patients infected with the
13 novel coronavirus COVID19 to assess whether COVID19 infection similarly leads to induction of
14 BHLHE41 as seen after infection with MERS-CoV and HCoV 229E. Modulation of BHLHE41
15 expression may represent a therapeutic strategy in coronavirus infection broadly and
16 specifically in COVID19 infection.
17
18
19
20
21
22
23
24
25
26
27
28

References

1. World Health Organization, 2020. Coronavirus disease 2019 (COVID-19): situation report, 49.
2. Wang, Y. and Zhu, L.Q., 2020. Pharmaceutical care recommendations for antiviral treatments in children with coronavirus disease 2019. *World Journal of Pediatrics*, pp.1-4.
3. GSE100509. Baric R, Sims A, Heller N, Waters KM, Einfeld AJ, Kawaoka Y. PNNL. 902 Battelle Blvd. Richland, USA.
4. Poppe, M., Wittig, S., Jurida, L., Bartkuhn, M., Wilhelm, J., Müller, H., Beuerlein, K., Karl, N., Bhujju, S., Ziebuhr, J. and Schmitz, M.L., 2017. The NF- κ B-dependent and-independent transcriptome and chromatin landscapes of human coronavirus 229E-infected cells. *PLoS pathogens*, 13(3), p.e1006286.
5. Baltimore, D., 1971. Expression of animal virus genomes. *Bacteriological reviews*, 35(3), p. 235.
6. Sawicki, S.G., Sawicki, D.L. and Siddell, S.G., 2007. A contemporary view of coronavirus transcription. *Journal of virology*, 81(1), pp.20-29.
7. Lai, M.M. and Cavanagh, D., 1997. The molecular biology of coronaviruses. In *Advances in virus research* (Vol. 48, pp. 1-100). Academic Press.
8. Peiris, J.S.M., Lai, S.T., Poon, L.L.M., Guan, Y., Yam, L.Y.C., Lim, W., Nicholls, J., Yee, W.K.S., Yan, W.W., Cheung, M.T. and Cheng, V.C.C., 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. *The Lancet*, 361(9366), pp.1319-1325.
9. Lu, G., Hu, Y., Wang, Q., Qi, J., Gao, F., Li, Y., Zhang, Y., Zhang, W., Yuan, Y., Bao, J. and Zhang, B., 2013. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. *Nature*, 500(7461), pp.227-231.
10. van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J. and Berkhout, B., 2004. Identification of a new human coronavirus. *Nature medicine*, 10(4), pp.368-373.
11. van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J. and Berkhout, B., 2004. Identification of a new human coronavirus. *Nature medicine*, 10(4), pp.368-373.

12. Woo, P.C., Lau, S.K., Chu, C.M., Chan, K.H., Tsoi, H.W., Huang, Y., Wong, B.H., Poon, R.W., Cai, J.J., Luk, W.K. and Poon, L.L., 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *Journal of virology*, 79(2), pp.884-895.
13. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R. and Niu, P., 2020. A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*.
14. Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D.Y., Chen, L. and Wang, M., 2020. Presumed asymptomatic carrier transmission of COVID-19. *Jama*.
15. World Health Organization, 2020. Coronavirus disease 2019 (COVID-19): situation report, 59.
16. Yang, X.O., Angkasekwinai, P., Zhu, J., Peng, J., Liu, Z., Nurieva, R., Liu, X., Chung, Y., Chang, S.H., Sun, B. and Dong, C., 2009. Requirement for the basic helix-loop-helix transcription factor Dec2 in initial T H 2 lineage commitment. *Nature immunology*, 10(12), p.1260.
17. Montagner, M., Enzo, E., Forcato, M., Zanconato, F., Parenti, A., Rampazzo, E., Basso, G., Leo, G., Rosato, A., Bicciato, S. and Cordenonsi, M., 2012. SHARP1 suppresses breast cancer metastasis by promoting degradation of hypoxia-inducible factors. *Nature*, 487(7407), pp.380-384.
18. Kreslavsky, T., Vilagos, B., Tagoh, H., Poliakova, D.K., Schwickert, T.A., Wöhner, M., Jaritz, M., Weiss, S., Taneja, R., Rossner, M.J. and Busslinger, M., 2017. Essential role for the transcription factor Bhlhe41 in regulating the development, self-renewal and BCR repertoire of B-1a cells. *Nature immunology*, 18(4), p.442.
19. Rauschmeier, R., Gustafsson, C., Reinhardt, A., Noelia, A., Tortola, L., Cansever, D., Subramanian, S., Taneja, R., Rossner, M.J., Sieweke, M.H. and Greter, M., 2019. Bhlhe40 and Bhlhe41 transcription factors regulate alveolar macrophage self-renewal and identity. *The EMBO journal*, 38(19).
20. Inaguma, S., Riku, M., Hashimoto, M., Murakami, H., Saga, S., Ikeda, H. and Kasai, K., 2013. GLI1 interferes with the DNA mismatch repair system in pancreatic cancer through BHLHE41-mediated suppression of MLH1. *Cancer research*, 73(24), pp.7313-7323.
21. Baker, S.M., Plug, A.W., Prolla, T.A., Bronner, C.E., Harris, A.C., Yao, X., Christie, D.M., Monell, C., Arnheim, N., Bradley, A. and Ashley, T., 1996. Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. *Nature genetics*, 13(3), pp.336-342.

Rank	ID	p-value	F	Gene	Gene name
13	A_23_P139500/ NM_030762	2.59E-30	2424.4026	BHLHE41	basic helix-loop-helix family, member e41

Table 1: BHLHE41 is differentially expressed in primary human microvascular endothelial cells when comparing cells infected with wild-type Middle East respiratory syndrome coronavirus (MERS-CoV), icMERS-CoV EMC2012 and non-infected cells.

The rank of differential expression globally, the probe/transcript ID, the p-value of differential expression globally, F, a statistic used when more than two groups are compared in differential gene expression analysis (rather than the moderated t-statistic “t”), the gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	Gene	Gene name
94	36848/ NM_030762	0.00020815	8.21	1.105171	BHLHE41	basic helix-loop-helix family, member e41

Table 2: BHLHE41 is differentially expressed in the human cell line HuH-7 when comparing cells with human coronavirus-229E and non-infected cells.

The rank of differential expression globally, the probe/transcript ID, the p-value of differential expression globally, the moderated t-statistic “t”, the gene and gene name are listed in this chart.

Figure 1: BHLHE41 is transcriptionally induced following infection of primary human microvascular endothelial cells with wild-type Middle East respiratory syndrome coronavirus (MERS-CoV), icMERS-CoV.

Expression of BHLHE41 at the mRNA level is graphically represented with the mean value marked and the p-value listed for each comparison relative to baseline infection at 0 hours.

BHLHE41 HuH-7 cells: HCoV 229E

Figure 2: BHLHE41 is transcriptionally induced in the human cell line HuH-7 following infection with human coronavirus 229E.

Expression of BHLHE41 at the mRNA level is graphically represented with the mean value marked and the p-value listed, the result of a test evaluating the statistical significance of difference in expression of BHLHE41 with and without HCoV infection.