Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Short communication

Epidemiological transcriptomic data supports BCG protection in viral diseases including COVID-19

Abhay Sharma

CSIR-Institute of Genomics and Integrative Biology, Sukhdev Vihar, Mathura Road, New Delhi 110025 India

1. Introduction

Bacille Calmette-Guérin (BCG) vaccine, besides conferring adaptive immunity against tuberculosis, is considered to induce trained immunity, a general long-term stimulation of innate immune response that provides nonspecific protection against pathogens, especially respiratory tract viral infections (Netea et al., 2016, 2020; Giamarellos-Bourboulis et al., 2020). Insights from human vaccination and experimental infection models, and animal models suggest that transcriptional, epigenetic, and functional reprogramming of hematopoietic stem cells toward myelopoiesis, and epigenetic modifications of peripheral myeloid cells drive BCG induced trained immunity (Kaufmann et al., 2018; Cirovic et al., 2020; Khan et al., 2020; Arts et al., 2018). Currently, the prospect that the vaccine may limit the impact of the coronavirus disease 2019 (COVID-19) pandemic is a subject of immense interest (O'Neil and Netea, 2020; Phimister et al., 2020; Curtis et al., 2020; Netea et al., 2020; Redelman-Sidi, 2020). Epidemiological studies have shown that countries with childhood BCG vaccination policy, compared to nations without the policy, faced lower prevalence and mortality of COVID-19 (Berg et al., 2020; Escobar et al., 2020). However, at the same time, no association between childhood vaccination and protection from COVID-19 in adulthood has also been reported (Hamiel et al., 2020), fueling the speculation that BCG induced trained immunity might be short-lived (Moorlag et al., 2020). Meanwhile, a recent retrospective study of a cohort of adults who received or did not receive the vaccine in the preceding 5 years has suggested that BCG vaccination is possibly associated with a lower incidence of illness during the pandemic (Moorlag et al., 2020). Also, while interventional clinical trials are underway to test whether BCG may have an immediate beneficial effect in COVID-19, a randomized clinical trial has shown that BCG vaccination in the elderly lengthens the time to first infection and reduces the incidence of new infection, especially of respiratory tract viruses (Giamarellos-Bourboulis et al., 2020). Interestingly, both the above studies (Giamarellos-Bourboulis et al., 2020; Moorlag et al., 2020) surprisingly found no evidence of excessive systemic inflammation in vaccinated subjects, alleviating the concern that BCG trained immunity, due to
elevated innate immune response, may add to hyperinflammation in severe disease and worsen the condition (Giamarellos-Bourboulis et al., 2020). The negative evidence of inflammation has been indirectly explained by arguing that BCG vaccination may lower systemic inflammation through enhanced antiviral defense response leading to decreased viral loads (Giamarellos-Bourboulis et al., 2020; Moorlag et al., 2020). Taken together, an improved understanding of both the durability and the mechanisms of BCG induced trained immunity is urgently required, particularly in view of the concern that indiscriminate use of the vaccine may jeopardize supply required for vaccination of children to protect them against tuberculosis, and the possibility that BCG may produce deleterious effects in severe COVID-19 patients (Curtis et al., 2020). Given the potential for knowledge discovery from themed collection and integration of available transcriptomic datasets (Wang et al., 2019), an integrative analysis of human epidemiological transcriptomic data related to BCG vaccination and viral respiratory infections, along with COVID-19 transcriptomic data, is presented here toward understanding the durability and mechanism of trained immunity.

2. Results

Clustering differentially expressed genes (DEG) in whole blood (WB) of infants 2, 6, 12, 18 and 26 weeks following BCG vaccination (Loxton et al., 2017), compared to the baseline, show directionally consistent expression changes across time points (Fig. 1A). Similarly, clustering DEG in BCG stimulated peripheral blood mononuclear cells (PBMC), compared to unstimulated PBMC, of BCG naïve adults and BCG vaccinated adults - median time since vaccination, 10 years - before and after 2, 7 and 14 days of an intradermal BCG challenge (Matsumiya et al., 2015) shows directionally consistent changes across time points in general, with vaccinated subjects exhibiting higher level of transcriptomic alterations than naïve subjects (Fig. 1B). Notably, clustering of common genes between the above two clusters (Fig. 1A,B) shows that genes that are up- or down-regulated in WB weeks and months after vaccination are similarly regulated in PBMC after BCG stimulation, more so in vaccinated individuals (Fig. 1C). These results clearly suggest that, consistent with the concept of trained immunity, BCG vaccination leads to persistent changes in peripheral blood cell transcriptome. Next, clustering of DEG in WB of patients with varying severity and time

![Figure 1](image-url). Integrative transcriptomics of BCG vaccination and viral respiratory diseases. (A-1) Clustering of DEG identified under indicated conditions. Abbreviation: BCGv, BCG vaccination; BCGvcs, BCG vaccination, challenge, and stimulation; BCGnc, BCG naïve and challenged; BCGvc, BCG vaccination and challenged; S/T series, both severity and time series; S3-S1, severity in decreasing order; T1-T3, sequentially from time of patient enrolment to clinical resolution; S, severe; M, moderate; SM, severe compared to moderate; I, inpatient; O, outpatient; IO, inpatient compared to outpatient; Mi, mild; SMI, severe compared to mild; IAV, influenza A virus; RSV, respiratory syncytial virus; IBV, influenza B virus; RSV-1 to 5, five different RSV datasets; HRV, human rhinovirus; T series, time series; AR, both IAV and HRV; BR, both ABV and HRV; MO, maximum supplemental oxygen. Clusters BCGvcs (B) and Virus S/T series (D) are composed of most frequent DEG among samples. * = indicates common genes in the given clusters. Multicolored bar represents DEG directionality: red (42 genes), upregulation in all BCG and all virus samples; green (5 genes), downregulation in all BCG and all virus samples; orange (8 genes), upregulation in all BCG samples and downregulation in all virus samples; purple (54 genes), downregulation in all BCG samples and upregulation in virus samples. Sample specific details including original reference and data source, DEG identification method, and cluster values are indicated in Supplementary Table 1.
points, and in- and out-patient status of illness from infection with respiratory viruses (Dunning et al., 2018; Tang et al., 2019; de Steenhuisen Piters et al., 2016; Aschenbrenner et al., 2021), compared to healthy or appropriate patient controls, shows transcriptomic changes that are both common across viruses as well as virus specific, as expected, and differ in magnitude in consonance with the variables (Fig. 1D). Remarkably, clustering of DEG that are common between the short-term BCG group (Fig. 1A) and virus group (Fig. 1D) shows strikingly similar and contrasting pattern of altered gene expression between the BCG and the virus group as a whole; (i) a major set of genes that are upregulated in both the groups, (ii) another major set of genes that are downregulated in the BCG and upregulated in the virus group, (iii) a minor set of genes that are downregulated in both the groups, and (iv) another minor set of genes that are upregulated in the BCG and downregulated in the virus group (Fig. 1E). Merging long-term BCG group (Fig. 1B) further tightened this pattern (Fig. 1F). The gene regulation observed in the virus group (Fig. 1F) is validated in general by an independent cluster of DEG in WB or PBMC of adults, infants, and children with infection of various respiratory viruses (Mejias et al., 2013; Ioannidis et al., 2012; Herberg et al., 2013), compared to healthy controls (Fig. 1G). Further independent validation is obtained by clustering DEG in WB of adult patients in a cohort of acute respiratory illness (ARI) from infection with various viruses (Zhai et al., 2015), compared to the baseline (Fig. 1H); with day 0, representing up to 48 h of ARI onset, resembling gene regulation observed in the virus group (Fig. 1F), and later time points showing absence of differential expression of genes that were upregulated by both the BCG and the virus group (Fig. 1F), and reversal from up- to down-regulation for genes that were upregulated in the virus group and downregulated in the BCG group. Another independent validation is observed with adult patient-specific WB pattern of gene expression in COVID-19 (Ong et al., 2020); z-score transformed and normalized gene counts in severe patient, especially at time points of higher supplemental oxygen, resembling the virus group (Fig. 1F), and in the same patient during recovery or in mild patients showing a trend toward opposite regulation (Fig. 1I). Gene ontology analysis of all the genes in the final BCG-virus cluster (Fig. 1F) shows enrichment of various biological process terms (Fig. 2); mainly, the genes that are upregulated in both the groups are associated with terms related to type I interferon signaling, cytokine-mediated signaling, and defense response to virus, whereas genes that are upregulated in the virus group and downregulated in the BCG group are associated with myeloid leukocyte activation, myeloid leukocyte mediated immunity, and other related terms. Next, genes in the final BCG-virus cluster (Fig. 1F) were examined for blood cell type distribution, with the results demonstrating a clear myeloid bias for genes that are downregulated in the BCG and upregulated in the virus group, and no gross bias for genes that are upregulated in both the groups (Fig. 3A). Finally, to validate this finding in disease setting, the BCG-virus cluster genes (Fig. 1F) were intersected with DEG identified in single-cell RNA-seq analysis of PBMC from severe and severe-moderate patients of COVID-19 (Wilk et al., 2020; Arunachalam et al., 2020), compared to healthy controls; as expected from the above cell type distribution analysis (Fig. 1F), a clear myeloid bias was found for the genes that are downregulated in the BCG and upregulated in the virus group, whereas genes that are upregulated in both the groups did not starkly differ in myeloid-lymphoid distribution (Fig. 3B). Also, unlike severe patients, severe-moderate cases did not show a clear upregulation of genes that are downregulated in the BCG and upregulated in the virus group (Fig. 1F), consistent with observed differences (Fig. 1D-F,H,I) among patients with varying disease status including severity and time points.

3. Discussion

The present analysis suggests that BCG vaccination induces long-term changes in blood immune cell transcriptome that partly mimic and partly oppose transcriptomic changes induced by viral respiratory illnesses including COVID-19. The mimicking part relates to upregulation of defense response to virus, and the opposing part to downregulation of myeloid cell activation in vaccinated subjects, with upregulation of the same observed in patients. Given these, individuals vaccinated with BCG in the near or distant past would be expected to have a higher basal level of antiviral defense and a lower basal level of
myeloid cell activation, rendering them less susceptible to viral infections on account of the former, and to hyperinflammation, that characterizes severe disease (Moorlag et al., 2020), on account of the latter. This prediction is consistent with the epidemiological studies showing lower prevalence and mortality of COVID-19 in countries with childhood BCG vaccination policy (Berg et al., 2020). Regarding mechanism, the present finding is in line with the existing concept (Netea et al., 2020) that trained immunity involves reprogramming of myeloid cells, and enhanced capacity for cytokine production and antimicrobial function. However, instead of myeloid cell activation genes showing upregulation in vaccinated subjects, as might have been possibly expected on account of enhanced innate immune response that is considered to characterize trained immunity, these genes show persistent downregulation postvaccination. Notably, a retrospective cohort study (Moorlag et al., 2020) and a randomized clinical trial (Giamarellos-Bourboulis et al., 2020) have recently reported that BCG vaccination is not associated with systemic inflammation, alleviating the concern that the vaccine may produce deleterious effects in severe COVID-19 by adding to hyperinflammation. That BCG is not associated with inflammation (Giamarellos-Bourboulis et al., 2020; Moorlag et al., 2020) has been indirectly explained by arguing that BCG induced enhancement in antiviral defense would cause decreased viral loads which in turn would lead to lower systemic inflammation (Moorlag et al., 2020). In contrast, the analysis presented here would suggest that a suppressed level of myeloid cell activation may directly limit innate immunity response from causing inflammation. Notably, a recent clinical trial has found that moderate adult COVID-19 patients administered a single dose of intradermal BCG achieve faster resolution of hypoxia, and significant radiological improvement and viral load reduction, without showing evidence of BCG induced cytokine storm (Padmanabhan et al., 2020). The present observation of persistently upregulated antiviral defense response and downregulated myeloid cell activation in BCG vaccinated subjects in blood cells is in line with this new finding. However, given limitations of in silico evidence, the present analysis would require functional validation. Whether BCG induced short- or long-term gene expression changes observed under normal conditions indeed form the basis of protection against viral infections and hyperinflammatory conditions would need confirmation. In conclusion, the present evidence suggesting high durability of BCG induced trained immunity provides a rationale against possible indiscriminate use (Curtis et al., 2020) of BCG vaccination as a measure to reduce the impact of COVID-19 pandemic. Also, it offers a mechanistic reasoning to ease the concern (Curtis et al., 2020) that BCG induced trained immunity may escalate systemic inflammation and worsen the condition in severe COVID-19 patients.

4. Methods

Relevant datasets were identified by extensively searching NCBI’s PubMed and Gene Expression Omnibus (GEO), with a preference for large and comparative studies, and were manually curated and annotated. Original author-identified gene sets were used, if available in full. Otherwise, differentially expressed genes (DEG) with adjusted p value significance, without any log2 fold change cutoff, were identified from GEO datasets by using GEO2R (Barrett et al., 2013). Details regarding sets of DEG used in the analysis, including GEO accession numbers and references, are individually presented in Supplementary Table 1 and 4, as indicated in the relevant figure legend (Figs. 1, 3). Dataset associated gene symbols were used as given, with duplicates removed using excel. Heatmapper (Babicki et al., 2016) was used to generate heatmaps, with genes clustered using the criteria scale type none, average linkage as clustering method, and Euclidean as distance measurement method (Fig. 1A,B,D) or using manual arrangement for distinguishing up- and down-regulation of genes (Fig. 1C,E-I, Fig. 2B). Gene ontology biological process enrichment was determined using TopFun of ToppGene Suite (Chen et al., 2005). For enrichment significance, q-value FDR B &Y was considered. Full enrichment results giving details including p and q values, hit count in query list, hit count in genome, and hit in query list are presented in Supplementary Table 2 as indicated in the relevant figure legend (Fig. 2). Heatmapper (Babicki et al., 2016) was used to display genes belonging to enriched gene ontology terms (Fig. 2). The list of cell type elevated genes in myeloid and lymphoid lineages in
blood analysis (Uhlen et al., 2019) was used for cell type classification. As indicated in the legend of the relevant figure (Fig. 3A), the distribution of individual DEG among cell type elevated gene sets is listed in Supplementary Table 3. Heatmap (Babicki et al., 2016) was used to depict genes belonging to different cell types (Fig. 3A).

5. Data availability

All data is available in the manuscript or the Supplementary Materials.

CRediT authorship contribution statement

Abhay Sharma: Conceptualization, Data curation, Formal analysis, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2021.145574.

References

Arts, R.J.W., Moorlag, S.J.C.F.M., Novakovic, B., Li, Y., Wang, S.-Y., Oosting, M., Kumar, V., Vashist, R.J., Wijmenga, C., Joosten, L.A.B., Reuten, C.B.E.M., Benc, C. S., Aaby, P., Koopmans, M.P., Stunnenberg, H.G., van Crevel, R., Netea, M.G., 2018. BCG vaccination protects against experimental viral infection in humans through the induction of cytokines associated with trained immunity. Cell Host Microbe 23 (1), 89–90.e5.

Arachchilage, P.S., Wimmers, F., Mok, C.K.P., Perera, R.A.P.M., Scott, M., Hagan, T., Sigal, N., Feng, Y., Bristow, L., Tak-Yin Tsang, O., Wang, D., Goller, J., Pellegrini, K. L., Kaemhn, D., Alaedarzadeh, G., Leung, W.S., Chan, I.M.C., Chik, T.S.H., Choi, C.Y.C., Huerta, C., Paine McCullough, M., Ly, H., Anderson, E., Edupuganti, S., Upadhyay, A.A., Bosinger, S.E., Maecker, H.T., Khatib, P., Roupheal, N., Peiris, M., Pulendran, B., 2020. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. Science 369 (6508), 1210–1220.

Aschenbrenner, A.C., Mouktaroudi, M., Kremer, B., Antonakos, N., Oestreich, M., Gzikelli, K., Nues-Urmena, M., Saridaki, M., Bonagoro, I., Reucht, N., Bassler, K., Doulos, S., Knoll, R., Pecht, T., Kapellou, T., Rosina, N., Krueger, C., Herbert, M., Holstein, I., Horne, A., Gremse, D., Agrawal, S., Bahn, K., van Ijzendoorn, B., Dews, A., Lenkeit, L., Brue, N., Gerretsen, J., Gierlichen, J., Becker, M., Haendler, K., Krat, M., Theis, H., Mengiste, S., Droggiti, D.-I., Abdorreza, F., Dahan, K., van Ijzendoorn, B., Redelman-Sidi, G., 2020. Could BCG be used to protect against COVID-19? Nat. Rev. Immunol. 20, 335–337.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, M., Wohrart, D.S., 2016. Heatmapter: Web-enabled heat map for all. Nucleic Acids Res. 44 (W1), W147–W153.

Barrère, T., Willhite, S.E., Ledsoux, P., Evangeliou, C., Kim, L.F., Tomashovskiy, M., Marshall, K.A., Phillips, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Servoa, N., Davis, S., Sobevala, A., 2013. NCBI GEO: archive for functional genomics data sets-update. Nat. Rev. Data. (Database issue), D991–D995.

Berg, M.K., Yu, Q., Salvador, C.E., Melani, I., Kitayama, S., 2020. Mandated bacillus Declaration of Competing Interest Writing - original draft, Writing - review & editing.

Hamel, U., Koer, E., Youngster, I. 2020. SARS-CoV-2 rates in BCG-vaccinated and unvaccinated young adults. JAMA 323 (22), 2340. https://doi.org/10.1001/jama.2020.8189.

Mejias, A., Villafuerte, L., Chen, J., Bardes, E.E., Aronow, B.J., Jegga, A.G., 2009. ToppGene Suite for gene list analysis and enrichment. Nucleic Acids Res. 37 (Web server issue), W305–W311.

Cirovic, B., de Bree, L.J.G., Groh, I., Blok, B.A., Chan, J., van der Velden, W.J.F.M., Bremmers, M.E.J., van Crevel, R., Handler, K., Picelli, S., Schute-Schrepping, J., Klei, K., Oosting, M., Koeken, V.A.C.M., van Ingen, J., Li, Y., Benn, C.S., Schulze, J.L., Joosten, L.A.B., Curtis, N., Netea, M.G., Schultze, J.L., 2020. BCG vaccination in humans elicits trained immunity via the hematopoietic progenitor compartment. Cell Host Microbe 28 (2), 322–335.

Curtis, N., Sparrow, A., Ghebreyesus, T.A., van de Veerdonk, F.L., Banchereau, J., 2013. Whole blood gene expression profiles to assess immunity to mild versus severe COVID-19 infection in humans. Cell Host Microbe 23 (1), 673–687.

https://doi.org/10.1016/j.gene.2021.145574.

Mozhui, K., Huang, S., Parnell, G.P., Nalos, M., Holubova, M., Chew, T., Booth, D., Johnstone, H., van der Syg, G., Maertzdorf, J., Kaufmann, E., Sanz, J., Dunn, J.L., Khan, N., Mendoza-Luca, L.E., Pacis, A., Tealeps, F., Banchereau, J., Brett, S.J., Moffatt, M.F., van Deuren, R.C., van Werkhoven, C.H., Jaeger, M., Debisarun, P., Cleophas, M.C., Smeenk, S., Oosting, M., van de Veerdonk, F.L., Joosten, L.A.B., Oever, J.T., van der Meer, J.W.M., Curtis, N., Aaby, P., Stabel-Benn, C., Giamarellos-Bourboulis, E.J., Lazi, N., Edwards, J., Curtis, N., van der Velden, R., Netea, M.G., 2020. Safety and immunogenicity of the recombinant Mycobacterium bovis BCG vaccine VPM1002 in HIV-unexposed newborn infants in south Africa. Clin. Vaccine Immunol. 27 (4), 513–519.

Ascheder, R.J., Xavier, R.J., 2016. Defining trained immunity and its role in health and disease. Science 352, aaf1098.

Yungster, I., Schwartz, J., Netea, M.G., 2020. Mandated bacillus Declaration of Competing Interest Writing - original draft, Writing - review & editing.

Nuccitelli, F., Klee, K., Oosting, M., Koeken, V.A.C.M., van Ingen, J., Li, Y., Benn, C.S., Schulze, J.L., van Crevel, R., Drogitti, D.-I., Benitez, G., Papadopoulos, A., Netea, M.G., 2020. Activation: Randomized clinical trial of BCG vaccination against infection in the elderly. Cell 183 (2), 315–323.e9.

Cirovic, B., de Bree, L.J.G., van Crevel, R., Handler, K., Picelli, S., Schute-Schrepping, J., Klei, K., Oosting, M., Koeken, V.A.C.M., van Ingen, J., Li, Y., Benn, C.S., Schulze, J.L., Joosten, L.A.B., Curtis, N., Netea, M.G., Schultze, J.L., 2020. BCG vaccination in humans elicits trained immunity via the hematopoietic progenitor compartment. Cell Host Microbe 28 (2), 322–335.

Hamel, U., Koer, E., Youngster, I. 2020. SARS-CoV-2 rates in BCG-vaccinated and unvaccinated young adults. JAMA 323 (22), 2340. https://doi.org/10.1001/jama.2020.8189.
associated with severe disease and fatality in patients with influenza infection. Nat. Commun. 10, 3422.

Uhlen, M., Karlsson, M.J., Zhong, W., Tebani, A., Pou, C., Mikes, J., Lakshmikanth, T., Forsström, B., Edfors, F., Odeberg, J., Mardinoglu, A., Zhang, C., von Felitzen, K., Mulder, J., Sjostedt, E., Hober, A., Okvold, P., Zwahlen, M., Ponten, F., Lindskog, C., Sivertsson, A., Fagerberg, L., Brodin, P., 2019. A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. Science 366 (6472), eaax9198. https://doi.org/10.1126/science:aax9198.

Wang, Z., Lachmann, A., Ma’ayan, A., 2019. Mining data and metadata from the gene expression omnibus. Biophys. Rev. 11, 103–110.

Wilk, A.J., Rustagi, A., Zhao, N.Q., Roque, J., Martinez-Colón, G.J., McKechnie, J., Ivison, G.T., Ranganath, T., Vergara, R., Hollis, T., Simpson, L.J., Grant, F., Subramanian, A., Rogers, A.J., Blish, C.A., 2020. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. Nat. Med. 26, 1070–1076.

Zhai, Y., Franco, L.M., Atmar, R.L., Quarles, J.M., Arden, N., Bucanas, K.L., Wells, J.M., Niño, D., Wang, X., Zapata, G.E., Shaw, C.A., Belmont, J.W., Couch, R.B., 2015. Host transcriptional response to influenza and other acute respiratory viral infections – A prospective cohort study. PLoS Pathog. 11, e1004869.