Epigenetic Chronicles of COVID-19: Statistical Analysis of Epigenetics for Prognosis and Identification of Long-Term Effects

Gentilini Davide [1,2], Calzari Luciano [1] , Dragani Davide Fernando [1] , Zanotti Lucia [3] , Inglese Elvira [2,4], Danesi Romano [4,5], Cavagnola Rebecca [2] , Brusati Alberto [2] , Persani Luca [6,7] , Campi Irene [7] , Baldrighi Giulia Nicole[2], De Martino Sara [8] , Farsetti Antonella [8] , Barbi Veronica [9], Gottardi Zamperla Michela [9] , Gaetano Carlo [9], Parati Gianfranco [3,10]

[1] Bioinformatics and Statistical Genomics Unit, IRCCS Istituto Auxologico Italiano, Cusano Milanino, Italy.

[2] Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy.

[3] Department of Cardiology, Istituto Auxologico Italiano, IRCCS, Milan, Italy;

[4] Clinical Chemistry and Microbiology, ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy.

[5] Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy.

[6] Department of Medical Biotechnology and Translational Medicine, University of Milan

[7] Department of Endocrine and Metabolic Diseases, IRCCS Istituto Auxologico Italiano, Milan, Italy.

[8] Consiglio Nazionale delle Ricerche (CNR) – IASI, Rome, Italy.

[9] Laboratorio di Epigenetica, Istituti Clinici Scientifici Maugeri IRCCS, Pavia, Italy.

[10] Department of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy.

Background: COVID-19 manifests across a broad spectrum of clinical symptoms, necessitating the identification of biomarkers capable of predicting the disease's clinical trajectory and long-term outcomes. Recent research has highlighted epigenetics, particularly DNA methylation, as a promising field for such investigations [1].

Objective: This study aims to explore the epigenetic changes that characterize the host response to COVID-19 and to investigate how SARS-CoV-2 infection modulates the host's epigenetic landscape. The goal is to identify predictive biomarkers and understand the underlying mechanisms influencing disease severity and long-term health effects.

Methods: The study utilized a robust, multi-phase design involving two distinct Italian cohorts. The "acute prognosis" cohort included 187 patients, with 64 experiencing severe COVID-19 and 123 with mild disease. The "follow-up" cohort consisted of 287 participants, comprising 96 patients six months post-infection and 191 healthy controls. Epigenetic profiling was conducted using the Illumina EPIC 850K chip, allowing for comprehensive methylation analysis. Statistical methods included:

- **Differential Methylation Analysis**: Identified CpG sites with significant methylation differences using linear models adjusted for confounding factors [2].
- **Epigenetic Clock Analysis**: Assessed biological aging acceleration using established epigenetic clock algorithms.
- **Epigenetic Drift Evaluation**: Quantified the variability in methylation patterns over time, serving as a measure of epigenetic instability [3].

Results: The acute phase analysis revealed a specific epigenetic signature at hospital admission that significantly predicted severe outcomes. Patients with severe prognosis exhibited notable epigenetic drift ($p < 0.001$) and accelerated biological aging (mean age acceleration: 5.2 years, 95% CI: 3.1-7.3 years). In the follow-up cohort, 42 CpG sites displayed significant differential methylation (FDR < 0.05). Genes at these sites were involved in glutamate/glutamine metabolism and inflammatory pathways, including GLUD1, ATP1A3, and ARRB2, which are pertinent to long COVID symptoms. Additionally, recovered patients showed a slight but significant increase in biological age acceleration (mean: 2.3 years, 95% CI: 1.0-3.6 years) and increased epigenetic drift compared to controls ($p < 0.01$).

Conclusions: This study provides comprehensive insights into the epigenetic modifications associated with COVID-19, highlighting significant changes during both the acute infection and the long-term recovery phase. The findings underscore the potential of epigenetic biomarkers in predicting disease severity and long-term health outcomes. The robust design and rigorous statistical analyses lend strong validity to these results, emphasizing the importance of considering epigenetic factors in future COVID-19 research and therapeutic strategies.

References:

1- Barturen, G., Carnero-Montoro, E., Martínez-Bueno, M. *et al.* Whole blood DNA methylation analysis reveals respiratory environmental traits involved in COVID-19 severity following SARS-CoV-2 infection. *Nat Commun* **13**, 4597 (2022). https://doi.org/10.1038/s41467-022-32357-2

2- Campagna, M.P., Xavier, A., Lechner-Scott, J. et al. Epigenome-wide association studies: current knowledge, strategies and recommendations. Clin Epigenet 13, 214 (2021). https://doi.org/10.1186/s13148-021-01200-8

3- Gentilini D, Scala S, Gaudenzi G, Garagnani Pet al. Epigenome-wide association study in hepatocellular carcinoma: Identification of stochastic epigenetic mutations through an innovative statistical approach. Oncotarget. 2017 Jun 27;8(26):41890-41902. doi: 10.18632/oncotarget.17462. PMID: 28514750; PMCID: PMC5522036.