Artificial intelligence reveals key biomarkers of necrotizing enterocolitis in the preterm infant gut microbiome

G. Casaburi¹, S. Kazi¹, S. Chew¹, B. Henrick¹,³, B. Ling², S. Frese¹,³, K. Sylvester²

1. Evolve Biosystems, Inc, Davis, CA, United States.
2. Department of Surgery, Stanford University, Stanford, California, United States of America
3. Department of Food Science and Technology, University of Nebraska, Lincoln, NE, United States.

Introduction: Necrotizing enterocolitis (NEC) is the leading cause of overall infant mortality in the United States, affecting 0.1% of newborns. The major limitation in NEC prevention dwells in the inability to predict which subset of premature infants are at risk for developing NEC. By combining functional and taxonomic data from infant gut microbiomes, an algorithm capable of discriminating NEC cases from pre-term controls was developed.

Methods: A total of 1,712 raw publicly available shotgun metagenomic datasets were collected, (cases = 253; controls = 1,459). Taxonomic and functional analyses were carried out and datasets were divided based on corrected gestational age (cGA). Several machine learning models were tested to identify functional core biomarkers distinguishing samples from infants with NEC (cases) from controls. Cell culture assays were used to assess the impact of functions and pathway intermediates on epithelial cell injury.

Results: The 29-32 weeks cGA population exhibited a significantly higher prediction accuracy among models (up to 99.8%). Intersection of models led to the identification of key proteins and functional pathways, which were then coupled with taxonomic classification to establish a collection of biomarkers able to discriminate cases from controls. The most discriminatory bacterial species was Enterobacter cloacae. In vitro assays confirmed key metabolic functions play a role in epithelial cell damage associated with NEC.

Conclusions: Large-scale data analytics approaches are increasingly playing a role in the identification of disease etiologies. Here, we find that this approach was able to identify key microbial risk factors, including bacterial functions and pathways, associated with necrotizing enterocolitis.

Keywords: NEC, Microbiome, Metagenomics, Artificial Intelligence