Network-based enrichment analysis of differentially methylated regions in pediatric multiple sclerosis reveals key regulators of immune response and Epstein-Barr virus infection

Andrea Corona[1,2], Martina Tosi[3] , Marta Mellai[4] , Alen Zollo[1,2], Marta Allesina[4], Marta Simone[5] , Alessandra Protti[6], Angela Berardinelli[7], Antonio Gallo[8], Carlotta Canavese[9], Domizia Vecchio[10], Eleonora Cocco[11], Lucia Moiola[12], Marta Zaffira Conti[13], Martina Borghi[14], Maurizio Viri[15], Alessandra Mingione[1,2] , Pietro Annovazzi[16], Oscar Oddo[17], Roberta Lanzillo[18], Sarah Rasia[19], Stefania Bova[20], Valentina Torri Clerici[21], Stefano Sotgiu[22], Alberto Priori[2,23] , Maria Trojano[5], Maria Pia Amato[24], Roberto Giuseppe Ernesto Bergamaschi[7], Maura Pugliatti[25], Angelo Ghezzi[3], Sandra D'alfonso[3], Filippo Martinelli Boneschi [1,2, 23]

1 Laboratory of Precision Medicine of Neurological Diseases, Department of Health Science, University of Milan, Milan, Italy 2 CRC "Aldo Ravelli" for Experimental Brain Therapeutics, Department of Health Science, University of Milan, Milan, Italy 3 University Of Eastern Piedmont, Department Of Health Sciences, Novara, Italy, 4 Centro Di Ricerca Traslazionale Sulle Malattie Autoimmuni E Allergiche (Caad), Novara, Italy, 5 Azienda Ospedaliera-Universitaria Consorziale Policlinico Di Bari, Department Of Basic Medical Sciences, Neuroscience And Sense Organs; Neurophysiopathology Unit, Bari, Italy, 6 Grande Ospedale Metropolitano Niguarda Di Milano, Department of neurology, Milan, Italy, 7 Irccs Fondazione Istituto Neurologico Nazionale Mondino, Pavia, Italy, 8 Azienda Ospedaliera-Universitaria Vanvitelli, Centro Sclerosi Multipla – I Clinica Neurologica, Napoli, Italy, 9 Ospedale Infantile Regina Margherita - Città Della Salute E Della Scienza Di Torino, Neuropsichiatria Infantile , Torino, Italy, 10 University Hospital "Maggiore Della Carità", Department Of Neurology, Novara, Italy, 11 Ospedale Binaghi, Centro Regionale Sclerosi Multipla, Cagliari, Italy, 12 Ospedale San Raffaele, Milan, Italy, 13 Asst Papa Giovanni XXIII, Bergamo, Italy, 14 Fondazione Cavalieri Ottolenghi, Orbassano, Italy, 15 University Hospital "Maggiore Della Carità", Child Neurology And Psychiatry Unit, Novara, Italy, 16 Ospedale S. Antonio Abate, Centro Sclerosi Multipla , Gallarate, Italy, 17 Fondazione Istituto G. Giglio , Neurologia , Cefalù, Italy, 18 Università Degli Studi Di Napoli "Federico II, Dipartimento Di Neuroscienze, Scienze Riproduttive Ed Odontostomatologiche, Napoli , Italy, 19 ASST Spedali Civili Di Brescia, Usd Svsd Neurologia, P.O. Montichiari, Centro Regionale Sclerosi Multipla, Montichiari , Italy, 20 Ospedale Dei Bambini V Buzzi, ASST Fatebenefratelli - Sacco, Neurologia Pediatrica, Milan, Italy, 21 Istituto Neurologico C. Besta, Milan, Italy; 22 Azienda Ospedaliero Universitaria Di Sassari , Neuropsichiatria Infantile , Sassari, Italy 23 Clinical Neurology Unit, Azienda Socio-Sanitaria Territoriale Santi Paolo e Carlo and Department of Health Sciences, University of Milan, Milan, Italy, 24 Università Degli Studi Di Firenze , Neurologia, Firenze, Italy, 25 Azienda Ospedaliero Universitaria Di Ferrara , Ferrara , Italy

Background and Aim

Pediatric multiple sclerosis (PedMS) poses unique challenges in diagnosis and understanding the underlying disease mechanisms, given its complexity and the importance of early identification of the involved factors.

Leveraging epigenetic signatures detected in the previous phase of the study, specifically highly informative differentially methylated regions (DMRs), this study aims to elucidate the epigenetic mechanisms mediating the interplay between genetic susceptibility and environmental factors in PedMS onset, employing a network-based approach to investigate joint association signals.

Materials and methods

This multi-center retrospective study included 175 Italian subjects (122 PedMS patients and 53 healthy controls matched for age, sex, and ethnicity) from the PEDIGREE study group. Genomic DNA was extracted from peripheral blood. DNA methylation was analyzed using the Infinium Methylation EPIC Array v2. Bioinformatic analysis was performed in R environment and DMRs were identified using the DMRCate method. Using the NDEx tool, we performed a sub-network detection analysis, projecting genes overlapping or in close proximity to DMRs onto the high-confidence (combined score >95%) STRING v12 reference interactome. Specifically, we explored the network using the one-step neighborhood method to identify modules defined by the group of nodes connected to the query term(s) and all edges between these nodes. The results were imported into the Cytoscape 3.10.3 environment, where the g:Profiler tool was used to perform gene set enrichment analysis (GSEA) of the sub-network nodes against the Gene Ontology, Reactome and KEGG databases.

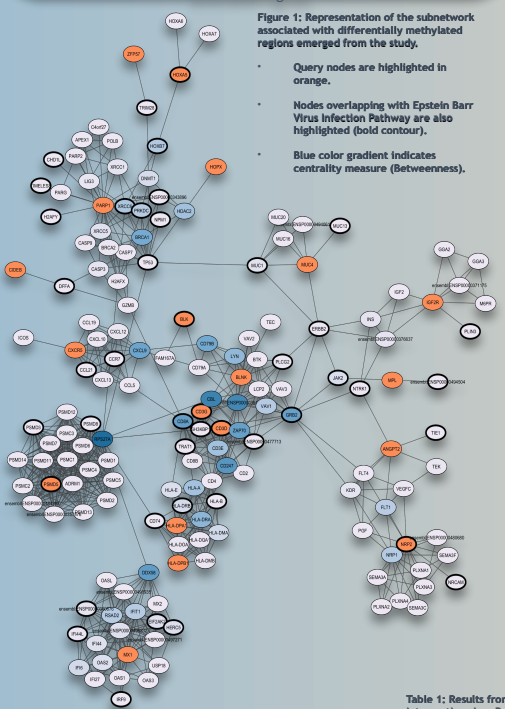

Figure 1: Representation of the subnetwork associated with differentially methylated regions emerged from the study.

- **Query nodes are highlighted in orange.**
- **Nodes overlapping with Epstein Barr Virus Infection Pathway are also highlighted (bold contour).**
- **Blue color gradient indicates centrality measure (Betweenness).**

Term name	Query size	Term size	Intersection size	P-value	Source
Epstein-Barr virus infection	**120**	**198**	**44**	**3.23E-40**	**KEGG**
positive regulation of immune system process	**145**	**1066**	**57**	**9.85E-33**	**Gene Ontology Biological Process**
immune system process	**145**	**2798**	**83**	**1.57E-32**	**Gene Ontology Biological Process**
proteasome regulatory particle	**145**	**22**	**17**	**2.60E-31**	**Gene Ontology Cellular Component branch**
immune response	**145**	**2008**	**71**	**5.41E-31**	**Gene Ontology Biological Process**
regulation of immune system process	**145**	**1526**	**63**	**2.31E-30**	**Gene Ontology Biological Process**
proteasome accessory complex	**145**	**25**	**17**	**1.05E-29**	**Gene Ontology Cellular Component branch**
positive regulation of immune response	**145**	**732**	**46**	**2.34E-28**	**Gene Ontology Biological Process**
positive regulation of biological process	**145**	**6235**	**111**	**3.14E-28**	**Gene Ontology Biological Process**
TCR signaling	**131**	**120**	**29**	**7.14E-28**	**Reactome pathways**
regulation of immune response	**145**	**897**	**49**	**1.03E-27**	**Gene Ontology Biological Process**
cell surface receptor signaling pathway	**145**	**2803**	**77**	**5.86E-27**	**Gene Ontology Biological Process**
Downstream TCR signaling	**131**	**98**	**26**	**6.54E-26**	**Reactome pathways**
lymphocyte activation	**145**	**787**	**45**	**6.99E-26**	**Gene Ontology Biological Process**
Cytokine Signaling in Immune system	**131**	**777**	**54**	**1.03E-25**	**Reactome pathways**
immune System	**131**	**2079**	**82**	**1.46E-25**	**Reactome pathways**
positive regulation of response to stimulus	**145**	**2293**	**69**	**1.61E-25**	**Gene Ontology Biological Process**
cell activation	**145**	**1098**	**50**	**1.05E-24**	**Gene Ontology Biological Process**
leukocyte activation	**145**	**951**	**47**	**1.85E-24**	**Gene Ontology Biological Process**

Table 1: Results from GSEA performed on the DMRs-associated sub-network of 165 genes. For each term query size, term size, intersection size, P-value and source database is shown.

Results

Analyses identified 55 DMRs, mapping on 91 genes, 33 of which were included in downstream analyses. NDEx revealed a sub-network of 165 nodes and 680 edges (figure 1). GSEA (table 1) indicated "**Epstein-Barr virus infection**" as the top ranked term (KEGG, p-value $3.23*10^{-40}$) with an overlap of 44 over 120 (term size 198). Other enriched terms among the top ten best ranking were **"immune system process"** (GOBP, p-value $1.57*10^{-32}$) and **"TCR signaling"** (Reactome pathways, p-value $7.14*10^{-28}$). Centrality measures indicates in addition to the already known ***PARP1*** and ***BLNK*** new interactors genes putatively involved in POMS like ***GRB2***, **CBL**, ***RPS27A, DDX58, CD247, CD8A, ZAP70, CXCL9, CD79B***.

Conclusions

Our findings highlight the potential role of epigenetic modifications, particularly DMRs, in mediating the complex interplay between genetic predisposition and environmental triggers, such as Epstein-Barr virus infection, in the onset of PedMS in this Italian cohort. The identified sub-network and enriched biological processes underscore the involvement of immune system dysregulation and lymphocyte activation in the disease pathogenesis. Building on evidence that several identified genes are key regulators of B cells, with a strong enrichment of genes specifically related to EBV infection processes, our hypothesis posits that epigenetic burden arising from prenatal and peri-natal exposures as well as from parental imprinting, in combination with genetic predisposition, may be pivotal in early MS onset.