A Deep Learning Method Uncovers Novel Pathway Associations in Systemic Lupus Erythematosus

Eric F. Morand[1], Thomas Dörner[2], Bochao Jia[3], Yushi Liu[3], Stephanie De Bono[3], Damiano Fantini[3], Maria Silk[3], Natalia Bello Vega[3], Peter Fischer[3], Michelle A Petri[4]

[1]Centre for Inflammatory Diseases, School of Clinical Sciences at Monash Health, Melbourne, Australia; [2]DRFZ Berlin and Department of Rheumatology and Clinical Immunology, Charité University Hospital Berlin, Berlin, Germany; [3]Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN, USA; [4]Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

BACKGROUND

- Systemic lupus erythematosus (SLE) is a multisystem autoimmune disease
- Although there have been recent improvements in therapy for SLE, there remains an unmet need for targeted treatments that are pathway specific
- Artificial intelligence can facilitate drug discovery by identifying potential target molecules and pathways

OBJECTIVES

- To use the deep learning algorithm to rank genes and pathways differentially expressed in SLE by an integrated metric of expression and interactivity

REFERENCES

1. Kong Y and Yu T. Bioinformatics. 2018;34:3727-3737.
2. Valdeolivas A, et al. Bioinformatics. 2019;35:497-505.
3. http://snap.stanford.edu/decagon/.
4. https://portal.genego.com/.
5. https://cran.rproject.org/web/packages/MetaIntegrator/index.html.

KEY RESULTS

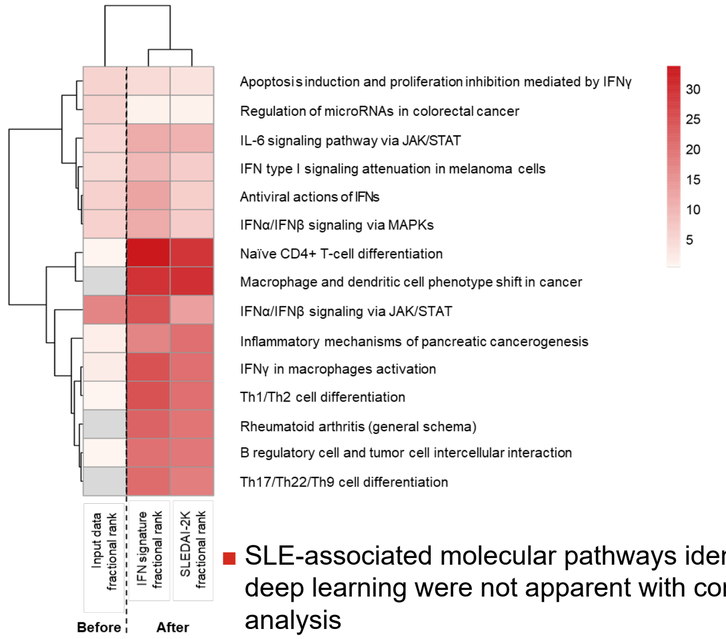
Highly Ranked Cytokines, Receptors, and Transcription Factors

Gene	Total SLEDAI-2K Fractional Rank[a]	IFN Fractional Rank[a]
IFNγ	0.9999	0.9999
IFNα[b]	0.999[c]	0.999[c]
STAT1[d]	0.9992	0.9991
IL10	0.9990	0.9987
IFNB1	0.9990	0.9989
STAT6	0.9981	0.9966
IL6	0.9975	0.9973
STAT4	0.9959	0.9944
IL12A	0.9952	0.9932
IL12B	0.9950	0.9938
STAT5A	0.9929	0.9940
STAT3	0.9915	0.9930
IFNγR2	0.9829	0.9770
IFNγR1	0.9812	0.9840
IL23A	0.9806	0.9895
IL10RA	0.9805	0.9816
IL6ST	0.9791	0.9695
STAT2	0.9772	0.9914

[a] Fractional rank (0-1) for each gene; [b] *IFNα* genes were similarly ranked; [c] Range=0.9988-0.999; [d] 0.9992=the SLEDAI-2K rank for STAT1 is higher than 99.92% of the whole genome

Deep Learning Analysis Revealed Strong Associations in Pathways Related to CD4+ T-Cell Differentiation, Macrophage, and Dendritic Cells That Were Not Evident in Conventional Analysis

Heatmap Summarizing Enriched Pathways From the Top 250 Analyzed Genes[a]



- SLE-associated molecular pathways identified with deep learning were not apparent with conventional analysis

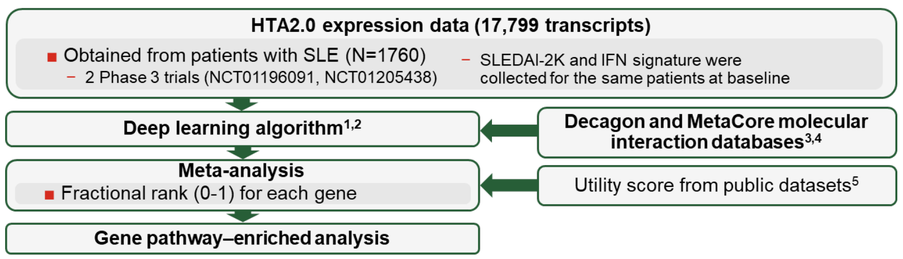
The heatmap's color intensity corresponds with the -log10 FDR-adjusted p-value.
The dendrogram shows groups of pathways with similar behavior

[a] Pathway analyses were performed using the top 250 genes with the highest fractional rank (before and after deep learning analysis) as input. Gene sets (maps) were obtained from the MetaBase database (Clarivate). Enriched pathways were detected via hypergeometric tests, and p-values were log-transformed and used to build the heat map

CONCLUSIONS

- Deep learning algorithms incorporating interactivity analyses revealed SLE-associated genes that were not apparent when using conventional bioinformatic approaches
- After the deep learning–based exercise, many IFN and JAK/STAT genes maintained very high fractional rankings, supporting the central role of the corresponding pathway in the molecular alterations of SLE
- Our deep learning-based study also unveiled new pathways that may be associated with SLE and were not apparent in analyses that were limited to the top differentially expressed genes in SLE versus healthy controls

METHODS



Assessments

- Deep learning algorithm
  - Graphical Embedded Deep Forward Network[1] and Random Walk With Restart[2] were conducted to determine the disease-specific network and the distance from each gene to other genes
- Meta-analysis
  - Each gene's utility score as the input for step 2, a measure of the gene's relevance to SLE disease activity derived from public datasets[5]
  - The final output was the fractional rank (0-1) of utilities[5] for each gene based on the distance derived using the pruned protein-protein interaction network against the prespecified outcomes, Systemic Lupus Erythematosus Disease Activity Index-2000 (SLEDAI-2K) and interferon (IFN) gene signatures based on deep neural network technique and multiplied by each gene's utility score in step 1
    - Total SLEDAI-2K analyzed as a continuous variable
    - IFN gene signatures analyzed as a continuous variable, combining 34 IFN-related genes
  - Fractional rankings of selected Janus kinase–signal transducer and activator of transcription (JAK-STAT), cytokine, and IFN genes are presented
- Pathways enriched in the top 250 genes from the input data and deep learning output were analyzed by hypergeometric tests

DISCLOSURES

- **E. Morand** has been a speaker, instructor, and/or consultant for: Amgen, AstraZeneca, Biogen, Bristol Myers Squibb, Eli Lilly and Company, EMD Serono, Genentech, GlaxoSmithKline, and Janssen; **T. Dörner** has received grant, research, and consulting support and/or speakers bureau fees from: Celgene, Chugai, Eli Lilly and Company, Janssen, Novartis, Roche, Sanofi, and UCB Pharma; **B. Jia**, **Y. Liu**, **S. de Bono**, **D. Fantini**, **M. Silk**, **N. Bello Vega**, and **P. Fischer** are employees and shareholders of: Eli Lilly and Company; **M. Petri** has received consulting support from: Eli Lilly and Company
- This study was sponsored by Eli Lilly and Company. Medical writing assistance was provided by Luke Carey, PhD, CMPP, of ProScribe – Envision Pharma Group, and was funded by Eli Lilly and Company
- Poster previously presented at European League Against Rheumatism (EULAR); Virtual; 2-5 June 2021

ABBREVIATIONS

CD4=cluster of differentiation 4; FDR=false discovery rate; HTA2.0=Human Transcriptome Array 2.0; IFN=interferon; IL=interleukin; JAK=Janus kinase; MAPK=mitogen-activated protein kinase; SLE=systemic lupus erythematosus; SLEDAI-2K=Systemic Lupus Erythematosus Disease Activity Index-2000; STAT=signal transducer and activator of transcription; Th=T helper

Scan or click the QR code or use this URL (https://lillyscience.lilly.com/congress/rhapp2021) for a list of all Lilly content presented at the congress.

Study was sponsored by Eli Lilly and Company, under license from Incyte Corporation