



OPEN

# Novel candidate drugs in anti-tumor necrosis factor refractory Crohn's diseases: in silico study for drug repositioning

Min Seob Kwak<sup>✉</sup>, Hun Hee Lee, Jae Myung Cha, Hyun Phil Shin, Jung Won Jeon & Jin Young Yoon

Biologicals like anti-tumor necrosis factor (TNF) therapy for Crohn's disease (CD) are safe and effective but there is a significant rate of primary and secondary nonresponse in the patients. In this study, we applied a computational approach to discover novel drug therapies for anti-TNF refractory CD in silico. We use a transcriptome dataset (GSE100833) for the anti-TNF refractory CD patients from NCBI GEO. After co-expression analysis, we specifically investigated the extent of protein–protein interactions among genes in clusters based on a protein–protein interaction database, STRING. Pathway analysis was performed using the cIEnrich function based on KEGG gene sets. Co-expressed genes in cluster 1, 2, 3, 4, up or down-regulated genes and all differentially expressed genes are highly connected. Among them, cluster 1, which is highly enriched for chemokine signaling, also showed enrichment for cytokine–cytokine receptor interaction and identifies several drugs including cyclosporin with known efficacy in CD. Vorinostat, histone deacetylase inhibitors, and piperlongumine, which is known to have inhibitory effect on activity of NF- $\kappa$ B, were also identified. Some alkaloids were also selected as potential therapeutic drugs. These findings suggest that they might serve as a novel therapeutic option for anti-TNF refractory CD and support the use of public molecular data and computational approaches to discover novel therapeutic options for CD.

Crohn's disease (CD) involves chronic and progressive transmural inflammation of the bowel characterized by repeated periods of remission and deterioration<sup>1</sup>. Pharmacologic management of CD currently consists of 5-aminosalicylic acid, corticosteroids, purine analogs azathioprine, and 6-mercaptopurine, and biologics including anti-tumor necrosis factor (TNF)- $\alpha$  inhibitors. Although the medical armamentarium continuously expands, some patients remain refractory to current therapeutic strategies.

Biologicals like anti-TNF agents (e.g., infliximab and adalimumab) are safe and effective but there is a significant rate of primary and secondary nonresponse affecting about 36–40% of patients<sup>2–4</sup>. Currently, anti- $\alpha$ 4-integrins, natalizumab and vedolizumab, are generally well tolerated, and a therapeutic option available for those patients<sup>5,6</sup>. Another numerous other agents for IBD treatment are currently under investigation, including Janus kinase inhibitors, anti-mucosal vascular address in cell adhesion molecule-1 agents, an anti-SMAD7 anti-sense oligonucleotide, an anti-interleukin-12/23 monoclonal antibody, and a sphingosine-1-phosphate receptor-1 selective agonist. However, these are limitations that make this treatment not always satisfactory. In addition, other therapeutic options with different mechanisms of action are required. Accordingly, additional novel drugs, which have potentially favorable clinical effects in these patients, are needed.

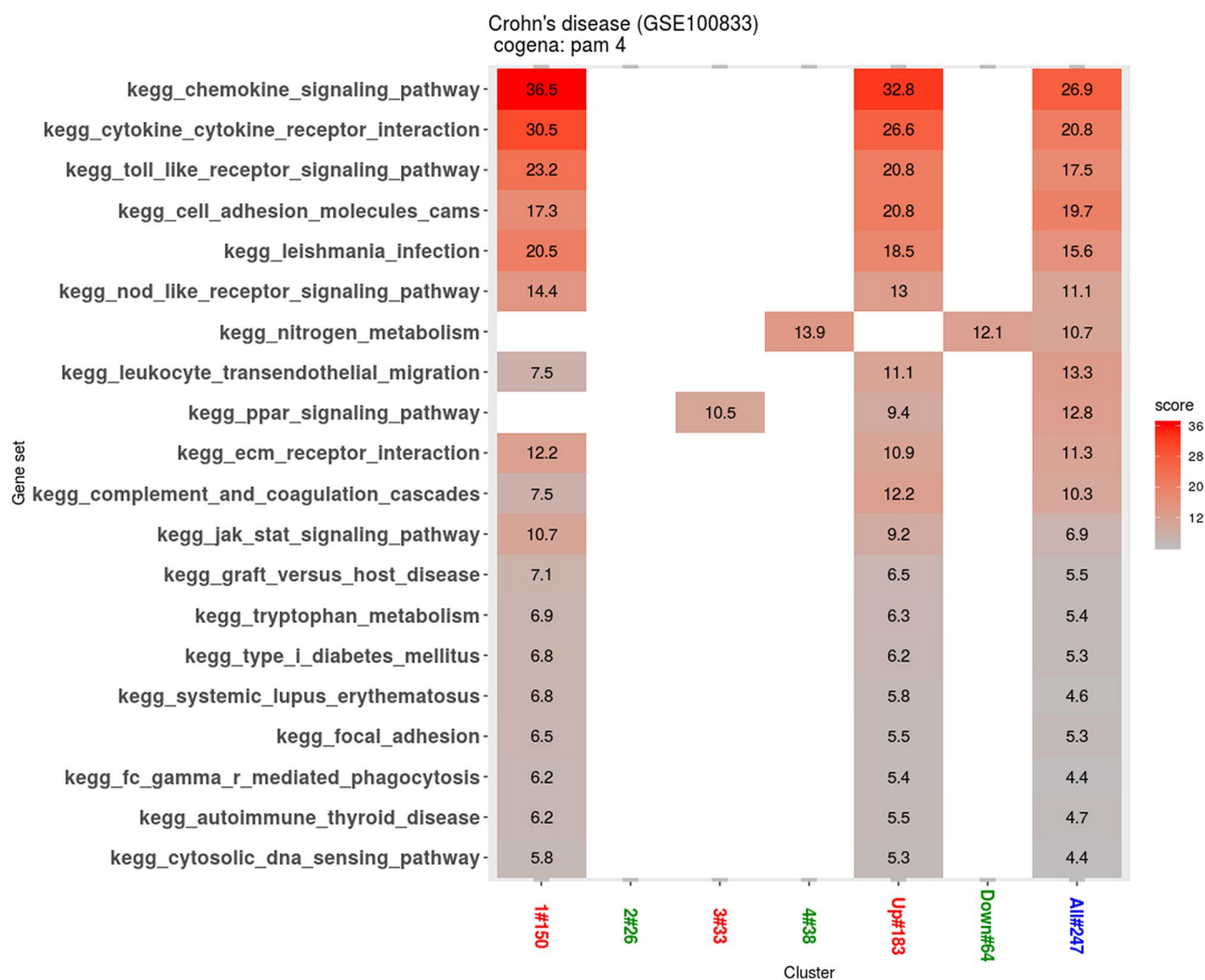
In this study, we applied a computational approach to discover novel drug therapies for CD in silico using publicly available molecular data measuring gene expression in CD samples and 164 small-molecule drug compounds.

Department of Internal Medicine, Kyung Hee University Hospital at Gangdong, Kyung Hee University College of Medicine, 892 Dongnam-ro, Gandong-gu, Seoul 05278, Republic of Korea. ✉email: kwac63@khu.ac.kr

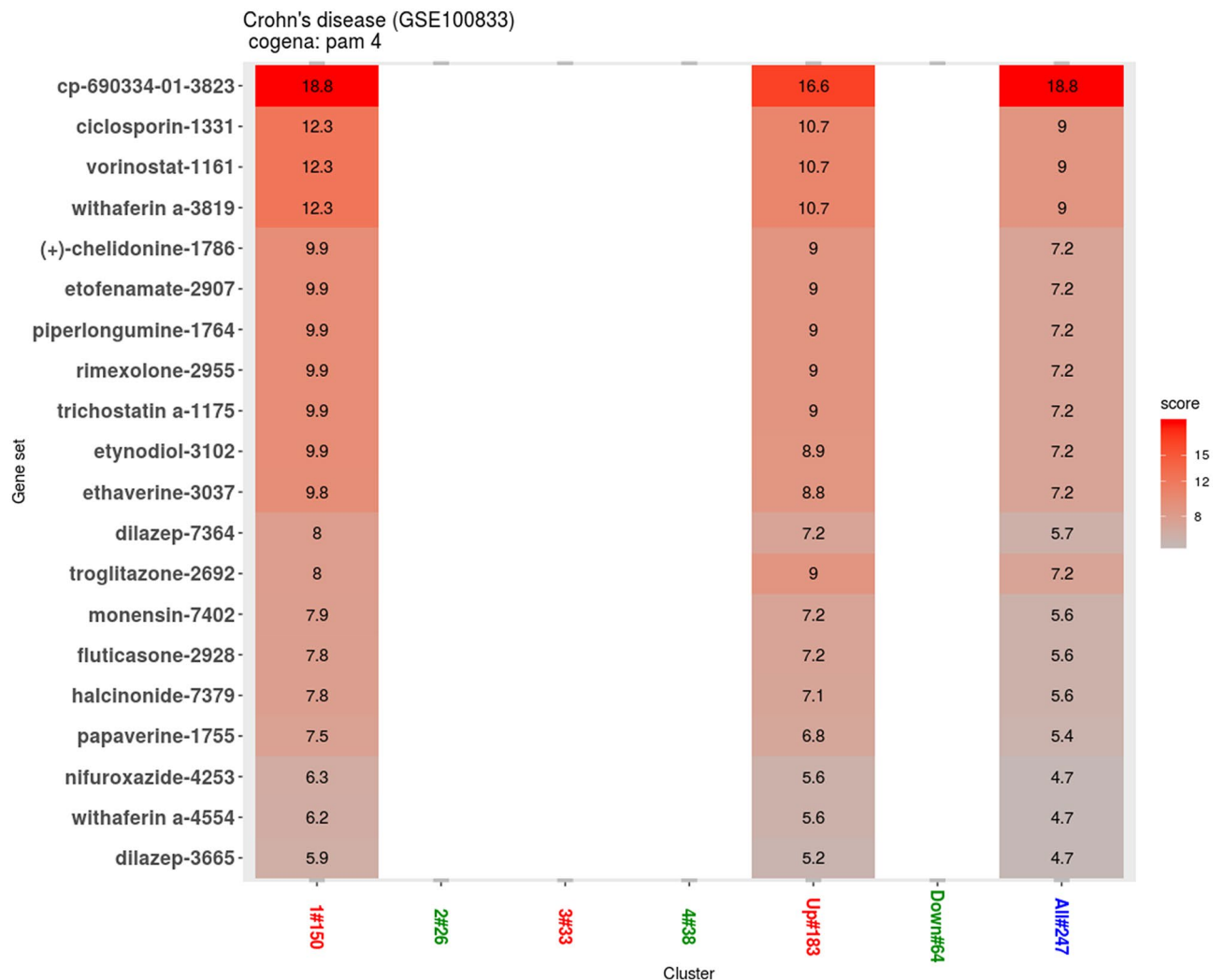


	The number of genes	The number of protein	Actual interactions	Expected interactions	p-value	Ratio
Cluster 1	158	145	443	102	<0.001	4.343
Cluster 2	25	25	0	0	1.00	–
Cluster 3	35	31	19	2	<0.001	9.5
Cluster 4	42	33	3	0	<0.001	–
Up	193	176	508	136	<0.001	3.735
Down	67	58	6	2	<0.001	3
All_DE	260	234	557	178	<0.001	3.129

**Table 1.** Summary of interactions within clusters for GSE100833. STRING interactions are shown for each cluster, up or down-regulated genes and all DEGs, how many genes (gene in cluster), how many proteins (protein in STRING), how many interactions (actual interaction), how many expected interactions (expected interaction), the ratio of actual interactions and expected interactions, together with the p value to get such a number of interactions by chance.



**Figure 2.** The KEGG pathway analysis results. The enrichment scores are shown based on different clusters, up-regulated, down-regulated and DEGs. In the x axis, the up-regulated clusters are colored red, while down-regulated clusters are coloured green and cluster containing all DEGs is coloured blue. The ranked pathways are shown in the y axis.



**Figure 3.** Drug repositioning results based on cluster 1. Enriched drugs with the instance number are shown on the y axis.

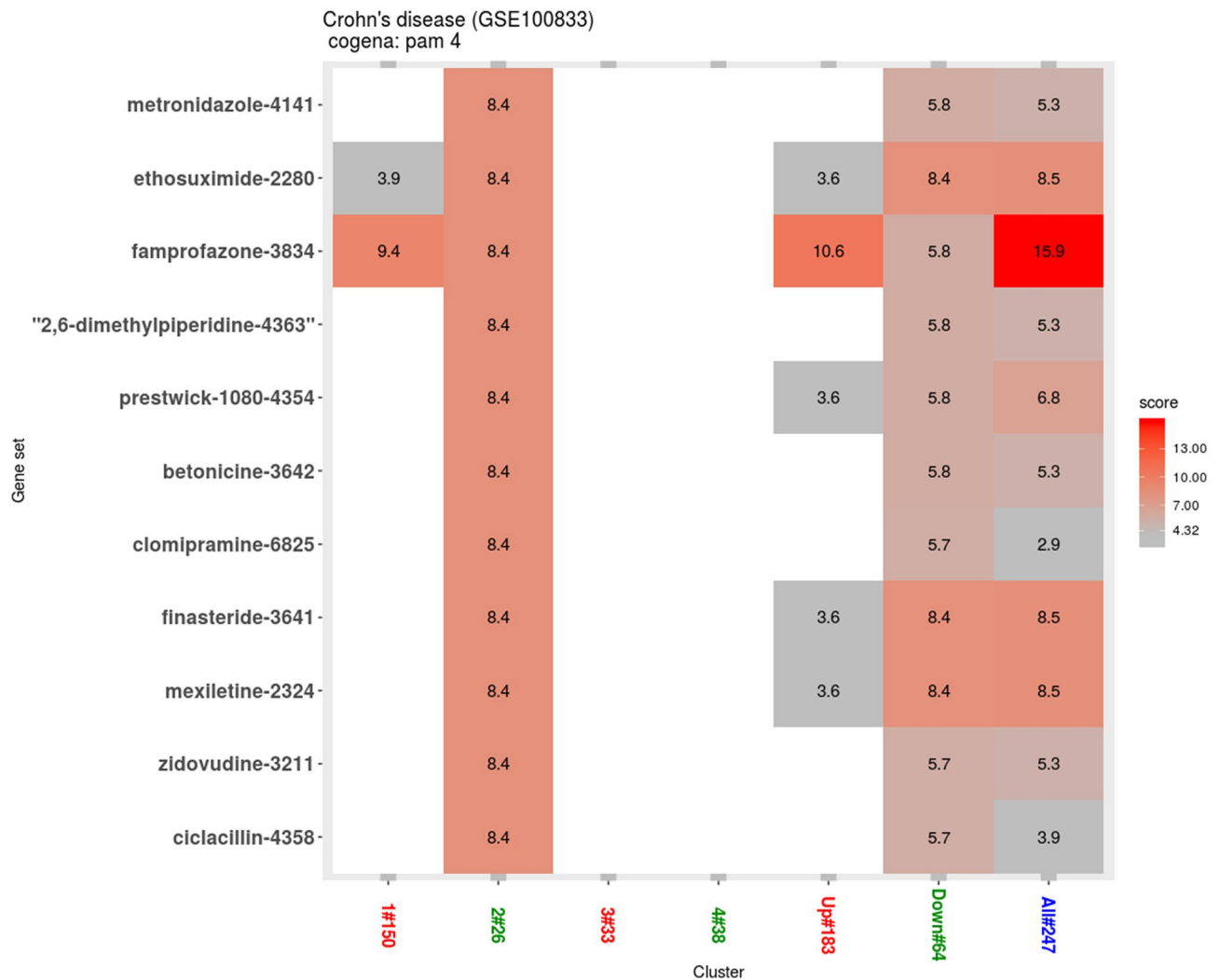
while a small fraction of co-expressed downregulated genes in cluster 4 were enriched in nitrogen metabolism pathways. The upregulated genes in cluster 3 are also observed to be enriched in peroxisome proliferator-activated receptors (PPAR) signaling pathways.

**Drug repositioning analysis.** The results show a high enrichment score in cluster 1 for CP-690334-01 drug profile which is not available for its information (Fig. 3). Following that, cyclosporine, vorinostat and withaferin drug profiles shows the same and second highest enrichment profiles in cluster 1. (Fig. 3). The drug-response associated genes are shown in Supplementary Fig. S3. CMap drug signatures matching cluster 2 were also investigated (Fig. 4) and 11 drug candidates with same enrichment score of the DEGs were selected as follows: metronidazole, ethosuximide, famprofazone, 2,6-dimethylpiperidine, prestwick, betonicine, clomipramine, finasteride, mexiletine, zidovudine, and ciclacilline.

Cluster 3 does not show any significant drug profile enrichments, despite of its highly enriched for PPAR signaling pathway. We also observed several candidate drug enrichment profiles (pf-00539758-00, clotrimazole, etanidazole, exisulind, lovastatin, myosmine, pentamidine, prochlorperazine, and sodium phenylbutyrate) in cluster 4, which was enriched in nitrogen metabolism pathway (Fig. 5). However, the ranks of the drugs are relatively lower compared with ranks of drugs in cluster 1 and 2 (Fig. 5).

## Discussion

To the best of our knowledge, this was the first study to predict the drug candidate compounds for anti-TNF refractory CD using CMap-based compound target prediction. Three biologically related groups of pathways, which have an established role in CD, are highlighted in clusters 1 shown in Fig. 2. The chemokine signaling pathway and cytokine–cytokine receptor interaction pathway play important roles in the pathogenesis of CD<sup>8</sup>. Toll-like receptors, as sensors of gut microbiota, play a critical role in maintaining the gut's homeostasis and



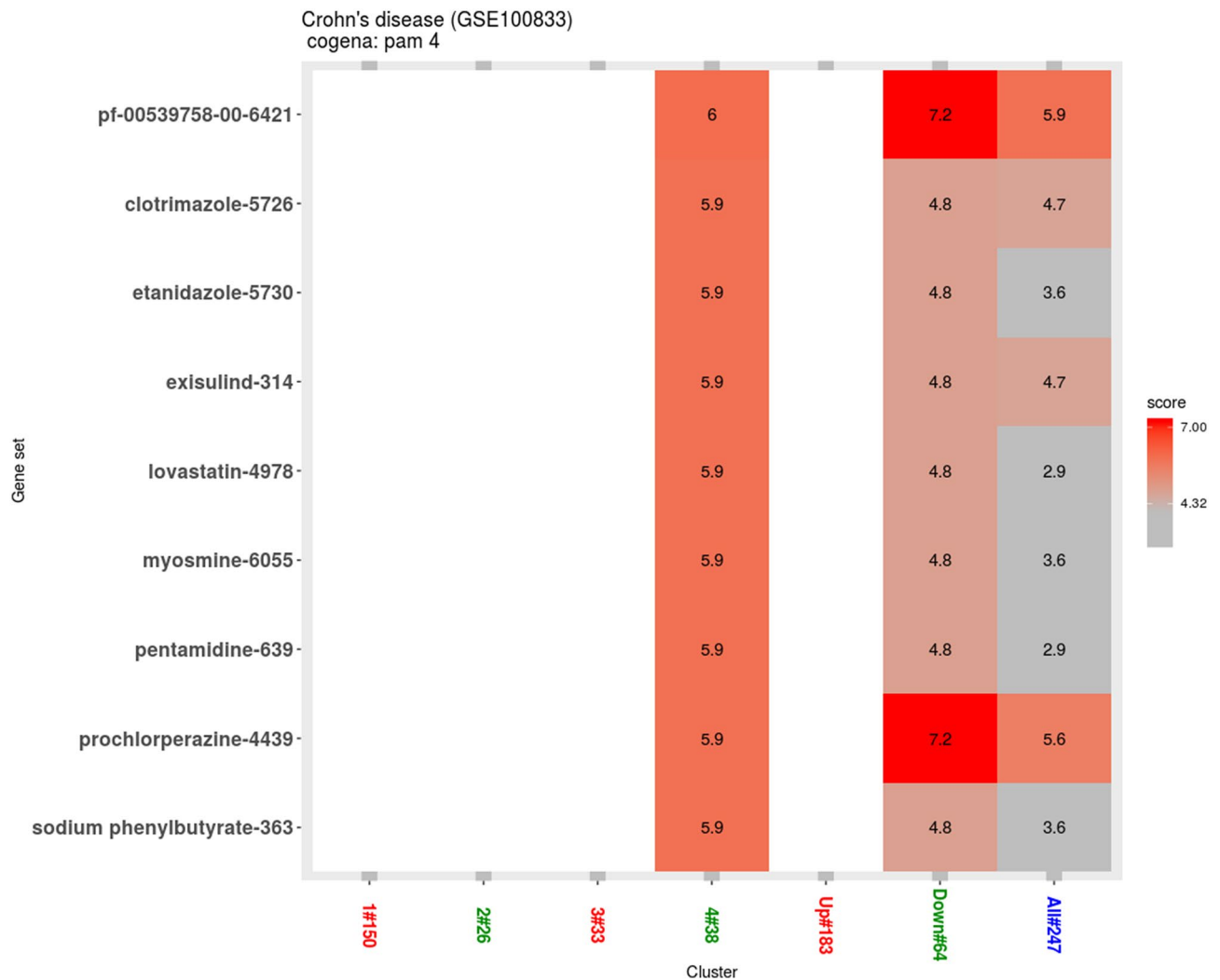
**Figure 4.** Drug repositioning results based on cluster 2. Enriched drugs with the instance number are shown on the y axis.

controlling the immune responses, thereby inciting CD<sup>9,10</sup>. They have been targeted by biologic drugs for the treatment of CD with some success<sup>11–13</sup>.

We focus our repositioning analysis on cluster 1. Cluster 1, which is highly enriched for chemokine and cytokine receptor interaction, identifies relatively many drug profile enrichments. Also, it is notable that the pathways represented in cluster 1 are now the focus of biologic therapies in CD, perhaps in response to the lack of efficacy of small molecule therapies in these pathways.

The information of CP-690334-01, ranked 1st, is not available, however the overrepresentation in anti-TNF refractory CD patients could give us a clear vision when applied to murine model and real patient. An FDA-approved drug for CD, cyclosporin, is captured and ranked 2nd. The data with respect to the use of cyclosporin in CD are very limited. However, recent Australian prospective study shows that a combination strategy of tacrolimus or cyclosporin with vedolizumab was safe and effective when the induction therapy was introduced to CD patients who had failed vedolizumab monotherapy<sup>14</sup>. This suggests that salvage therapy using cyclosporin may be a strategy to induce remission in to anti-TNF refractory CD patients.

Next, several histone deacetylase inhibitors (HDACi) were identified as potential candidate drugs. Vorinostat are approved for treating certain haematological malignancies, however, recent evidence also illustrates they are modulators of the immune system<sup>15–17</sup>. Therefore, there is emerging evidence that HDACi could have utility in the treatment of rheumatoid arthritis<sup>18–20</sup>, multiple sclerosis<sup>21</sup>, systemic lupus erythematosus<sup>22,23</sup>, and airway hyperresponsiveness<sup>24</sup>, sharing some clinical similarity with CD. There is solid evidence that in various colitis models in particular, HDACi exert an anti-inflammatory capacity. Several preclinical studies from experimental colitis models reported that the HDACi ameliorated macroscopic and histologic inflammation, in addition to reduced IFN- $\gamma$  production accompanied by increased production of IL-10<sup>25–27</sup>. Rather, one naturally occurring HDACi, butyrate, decreased TNF production and proinflammatory cytokine mRNA expression by intestinal biopsies and isolated lamina propria cells from CD patients<sup>28</sup>. Trichostatin A, another HDACi, induce energy in Th cells in vitro, as characterized by an inhibition of proliferation accompanied by a reduction of IL-2 production



**Figure 5.** Drug repositioning results based on cluster 4. Enriched drugs with the instance number are shown on the y axis.

in Th1 cell cultures<sup>29</sup>. These results that have identified the anti-inflammatory effect of HDACi also support a therapeutic role for CD.

Timely, phase I/II clinical trial is currently recruiting the patients to study the safety and efficacy of vorinostat in treating patients with moderate-to-severe CD who are not controlled by standard maintenance therapy (Clinical Trial ID, NCT03167437). Therefore, providing the efficacy, we expect that novel drug would be available soon.

Corticosteroid, such as withaferin A, rimexolone, etynodiol, fluticasone and halcinonide, is effective treatments for CD, but its systemic use has been widely replaced in the clinic by biologics, due to a shared adverse event risk.

Some alkaloids were also selected as potential therapeutic drugs. Chelidonium has several known therapeutic effects, which include mild analgesic, antimicrobial, and oncostatic effects, and it acts as a sedative in the central nervous system<sup>30</sup>. Further, previous Korean study suggest the chelidonium ameliorates colon injury and inhibits the increase of inflammatory mediators, such as IL-6 and TNF- $\alpha$ , and oxidative damage in murine colitis model<sup>31</sup>. Harris et al.<sup>32</sup> found that papaverine, another alkaloid, attenuated the disease activity index, in large part by significantly reducing rectal bleeding in DSS-induced colitis model. And there are few studies reporting on the anti-inflammatory activity of ethaverine, the ethyl analogue of papaverine. These results provide novel insights into the pharmacological actions of alkaloid and its potential use for the treatment of intestinal inflammation.

Previous study suggest that piperlongumine, a natural product isolated from the fruit, might modulate reactive oxygen species production under inflammatory conditions for rheumatoid arthritis<sup>33</sup>. The vast majority of literature studies have confirmed beneficial effects of PPAR agonists, such as troglitazone, on macroscopic and histopathological features of colitis<sup>34–36</sup>. However, supporting evidence for therapeutic potential is still limited in remained compounds of cluster 1, compounds enriched for PPAR signaling pathways (cluster 3), and compounds revealed enrichment of nitrogen metabolism pathway (cluster 4).

Several limitations of our study merit discussion. First, biological entities are nonlinear systems showing 'chaotic behaviour'. As such, there is no relation between the magnitude of the input and the magnitude of the output, with even the most minuscule differences between initial conditions rapidly translating into major

differences in the output. And second, there is potentially selection bias from single dataset, which would make the study unrepresentative of the entire anti-TNF refractory CD patients and may restrict the generalizability of our results. Therefore, our computational method would not be able to fully model the complexity of biological systems in CD patients with various confounding factors. Further experiments would therefore need to be performed in multiple, genetically diverse patient samples to confirm this.

In conclusion, among the top compounds predicted to be therapeutic for anti-TNF refractory CD patients by our approach were cyclosporine, a calcineurin inhibitors known to treat IBD, and vorinostat, a HDACi previously described to demonstrate positive results in numerous murine studies. Further, chelidonium and piperlongumine are also identified as potential candidates. These drugs could lead to in-human clinical trials and rapidly and relatively inexpensively offer several new treatments for patients with anti-TNF refractory CD. Since the results are based on in silico analysis, further in-depth studies are necessary to add to the validity of these results.

## Methods

**Data sources and processing.** Transcriptomic profiles dataset GSE100833 of anti-TNF refractory CD patients was downloaded from NCBI's Gene Expression Omnibus database (GEO datasets). A total of 327 patients consisting of non-inflamed and inflamed colonic tissues were profiled based on Affymetrix Human Genome U133 Plus 2.0 microarrays platform<sup>37</sup>. Each gene expression profile was normalized using robust multi-array average (RMA)<sup>38,39</sup> and non-expressed and non-informative genes were filtered using the MetaDE package<sup>40</sup>. The limma package<sup>41</sup> was used to identify differentially expressed genes with the thresholds of FDR less than 0.05 and absolute logFC more than 1. Co-expression analysis was performed using the coExp function. We have constructed a drug repositioning and drug mechanism-of-action (MoA) discovery pipeline based on the cogena framework using a pathway gene set and the CMap gene set<sup>42</sup>.

**Consensus clustering.** We used partitioning around medoids (PAM) clustering based on Euclidean distance, in order to identify the characteristics of gene expression profiles<sup>43</sup>. For the choice of cluster number, ConsensusClusterPlus R-package<sup>44</sup> was used to identify clusters using 1,000 iterations (reps), 80% sample resampling (pltem) from 2 to 20 clusters (k). The distance matrix was set to Pearson correlation (distance) and linkage function was set as wald. D (innerLinkage) and average (finalLinkage). In order to select optimal cluster number k, we calculated the empirical cumulative distribution (CDF) and the proportional area change under CDF ( $\Delta(k)$ ). According to the  $\Delta(k)$  vs k plot, the k where  $\Delta(k)$  started to approach zero was optimal. We also plotted the heatmap of consensus matrix at k to observe whether boundaries of each cluster were sharp. Considering the results of the  $\Delta(k)$  vs k plot and the heatmap, we determined the optimal cluster numbers.

**Connectivity map query.** To find compounds that have similar gene expression patterns, we identified the most significantly ( $Z < 0.001$ ) up- and down-regulated probes based on log ratio of gene expression of adjacent normal tissue, compared with the inflammatory lesions in anti-TNF refractory CD patients. After removal of duplicates, we ended up with 193 up-regulated genes and 65 down-regulated genes (Supplementary Table S2) that we submitted simultaneously for our CMap query (build02; [www.broadinstitute.org/cmap/](http://www.broadinstitute.org/cmap/)). Each signature was queried against the CMap using the gene set enrichment analysis algorithm<sup>45,46</sup>. By inputting a gene-expression profile of interest and querying it against the CMap data, a list of ranked CMap drugs is obtained (Supplementary Table S3).

**Protein–protein interactions and pathway analysis.** Following co-expression analysis, we specifically investigated the extent of protein–protein interactions among genes in clusters based on a protein–protein interaction database, STRING (Table 1)<sup>47</sup>. The expected interaction and p value are calculated based on a random background model that preserves the degree distribution of the input proteins<sup>47,48</sup>, implemented via the get-summary function in the STRINGdb package. The pathway analysis was performed using the clEnrich function based on KEGG gene sets and visualized with Enrichr program<sup>49,50</sup>. For network analysis, a novel web-based tool, OmicsNet (<https://www.omicsnet.ca>), was used<sup>51</sup>.

**Drug repositioning based on co-expression analysis.** Each of the co-expressed gene clusters were subjected to drug repositioning analysis based on CMap gene sets to identify the likelihood of drug MOA in identified pathways. The statistical analyses were performed using R version 3.5.1<sup>52</sup>. All significant thresholds were set at a two-sided p-value of 0.05.

Received: 7 September 2019; Accepted: 12 June 2020

Published online: 01 July 2020

## References

1. Strober, W., Fuss, I. & Mannon, P. The fundamental basis of inflammatory bowel disease. *J. Clin. Investig.* **117**, 514–521 (2007).
2. Hanauer, S. B. *et al.* Maintenance infliximab for Crohn's disease: The ACCENT I randomised trial. *Lancet* **359**, 1541–1549 (2002).
3. Rutgeerts, P. *et al.* Comparison of scheduled and episodic treatment strategies of infliximab in Crohn's disease. *Gastroenterology* **126**, 402–413 (2004).
4. Sandborn, W. J. *et al.* Adalimumab induction therapy for Crohn disease previously treated with infliximab: A randomized trial. *Ann. Intern. Med.* **146**, 829–838 (2007).
5. Lamb, C. A., O'Byrne, S., Keir, M. E. & Butcher, E. C. Gut-selective integrin-targeted therapies for inflammatory bowel disease. *J. Crohn's Colitis* **12**, S653–S668 (2018).
6. Sandborn, W. J. *et al.* Natalizumab induction and maintenance therapy for Crohn's disease. *N. Engl. J. Med.* **353**, 1912–1925 (2005).
7. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

8. Danese, S. New therapies for inflammatory bowel disease: From the bench to the bedside. *Gut* **61**, 918–932 (2012).
9. Kordjazy, N. *et al.* Role of toll-like receptors in inflammatory bowel disease. *Pharmacol. Res.* **129**, 204–215 (2018).
10. Lu, Y., Li, X., Liu, S., Zhang, Y. & Zhang, D. Toll-like receptors and inflammatory bowel disease. *Front. Immunol.* **9**, 72 (2018).
11. Moschen, A. R., Tilg, H. & Raine, T. IL-12, IL-23 and IL-17 in IBD: Immunobiology and therapeutic targeting. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 185–196 (2019).
12. Laroui, H. *et al.* Dextran sodium sulfate (DSS) induces colitis in mice by forming nano-lipocomplexes with medium-chain-length fatty acids in the colon. *PLoS ONE* **7**, e32084 (2012).
13. Esposito, G. *et al.* Palmitoylethanolamide improves colon inflammation through an enteric glia/toll like receptor 4-dependent PPAR-alpha activation. *Gut* **63**, 1300–1312 (2014).
14. Christensen, B. *et al.* Safety and efficacy of combination treatment with calcineurin inhibitors and vedolizumab in patients with refractory inflammatory bowel disease. *Clin. Gastroenterol. Hepatol.* **17**, 486–493 (2019).
15. Reilly, C. M., Regna, N. & Mishra, N. HDAC inhibition in lupus models. *Mol. Med.* **17**, 417–425 (2011).
16. Hancock, W. W., Akimova, T., Beier, U. H., Liu, Y. & Wang, L. HDAC inhibitor therapy in autoimmunity and transplantation. *Ann. Rheum. Dis.* **71**(Suppl 2), i46–54 (2012).
17. Waibel, M. *et al.* Manipulation of B-cell responses with histone deacetylase inhibitors. *Nat. Commun.* **6**, 6838 (2015).
18. Choo, Q. Y., Ho, P. C. & Lin, H. S. Histone deacetylase inhibitors: New hope for rheumatoid arthritis?. *Curr. Pharm. Des.* **14**, 803–820 (2008).
19. Young, D. A. *et al.* Histone deacetylase inhibitors modulate metalloproteinase gene expression in chondrocytes and block cartilage resorption. *Arthritis Res. Ther.* **7**, R503–512 (2005).
20. Chung, Y. L., Lee, M. Y., Wang, A. J. & Yao, L. F. A therapeutic strategy uses histone deacetylase inhibitors to modulate the expression of genes involved in the pathogenesis of rheumatoid arthritis. *Mol. Ther.* **8**, 707–717 (2003).
21. Gray, S. G. & Dangond, F. Rationale for the use of histone deacetylase inhibitors as a dual therapeutic modality in multiple sclerosis. *Epigenetics* **1**, 67–75 (2006).
22. Sun, X. Y., Shi, J., Han, L., Su, Y. & Li, Z. G. Anti-histones antibodies in systemic lupus erythematosus: Prevalence and frequency in neuropsychiatric lupus. *J. Clin. Lab. Anal.* **22**, 271–277 (2008).
23. Li, N. *et al.* HDAC inhibitor reduces cytokine storm and facilitates induction of chimerism that reverses lupus in anti-CD3 conditioning regimen. *Proc. Natl. Acad. Sci. USA.* **105**, 4796–4801 (2008).
24. el Assem, S. *et al.* Effects of a selection of histone deacetylase inhibitors on mast cell activation and airway and colonic smooth muscle contraction. *Int. Immunopharmacol.* **8**, 1793–1801 (2008).
25. Tao, R. *et al.* Deacetylase inhibition promotes the generation and function of regulatory T cells. *Nat. Med.* **13**, 1299–1307 (2007).
26. Glauben, R. *et al.* Histone deacetylases: Novel targets for prevention of colitis-associated cancer in mice. *Gut* **57**, 613–622 (2008).
27. Glauben, R. *et al.* Histone hyperacetylation is associated with amelioration of experimental colitis in mice. *J. Immunol.* **176**, 5015–5022 (2006).
28. Segain, J. P. *et al.* Butyrate inhibits inflammatory responses through NFkappaB inhibition: Implications for Crohn's disease. *Gut* **47**, 397–403 (2000).
29. Edens, R. E., Dagtas, S. & Gilbert, K. M. Histone deacetylase inhibitors induce antigen specific anergy in lymphocytes: A comparative study. *Int. Immunopharmacol.* **6**, 1673–1681 (2006).
30. Colombo, M. L. & Bosisio, E. Pharmacological activities of *Chelidonium majus* L. (Papaveraceae). *Pharmacol. Res.* **33**, 127–134 (1996).
31. Kim, D. S. *et al.* The therapeutic effect of chelidonic acid on ulcerative colitis. *Biol. Pharm. Bull.* **35**, 666–671 (2012).
32. Harris, N. R., Specian, R. D., Carter, P. R. & Morgan, G. A. Contrasting effects of pseudoephedrine and papaverine in dextran sodium sulfate-induced colitis. *Inflamm. Bowel Dis.* **14**, 318–323 (2008).
33. Xu, S. *et al.* Piperlongumine inhibits the proliferation, migration and invasion of fibroblast-like synoviocytes from patients with rheumatoid arthritis. *Inflamm. Res.* **67**, 233–243 (2018).
34. Sanchez-Hidalgo, M., Martin, A. R., Villegas, I. & de la Lastra, C. A. Rosiglitazone, a PPARgamma ligand, modulates signal transduction pathways during the development of acute TNBS-induced colitis in rats. *Eur. J. Pharmacol.* **562**, 247–258 (2007).
35. Sanchez-Hidalgo, M., Martin, A. R., Villegas, I. & Alarcon De La Lastra, C. Rosiglitazone, an agonist of peroxisome proliferator-activated receptor gamma, reduces chronic colonic inflammation in rats. *Biochem. Pharmacol.* **69**, 1733–1744 (2005).
36. Celinski, K. *et al.* Comparison of the anti-inflammatory and therapeutic actions of PPAR-gamma agonists rosiglitazone and troglitazone in experimental colitis. *J. Physiol. Pharmacol.* **63**, 631–640 (2012).
37. Nair, R. P. *et al.* Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat. Genet.* **41**, 199–204 (2009).
38. Irizarry, R. A. *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**, e15 (2003).
39. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).
40. Wang, X. *et al.* An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection. *Bioinformatics* **28**, 2534–2536 (2012).
41. Smyth, G. K., Ritchie, M., Thorne, N. & Wettenhall, J. LIMMA: Linear models for microarray data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Statistics for Biology and Health.* (2005).
42. Jia, Z. *et al.* Cogenia, a novel tool for co-expressed gene-set enrichment analysis, applied to drug repositioning and drug mode of action discovery. *BMC Genom.* **17**, 414 (2016).
43. Rousseeuw, P. J. & Kaufman, L. *Finding Groups in Data* (Wiley Online Library, Hoboken, 1990).
44. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).
45. Lamb, J. *et al.* The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
46. Lamb, J. The Connectivity Map: A new tool for biomedical research. *Nat. Rev. Cancer* **7**, 54–60 (2007).
47. Szklarczyk, D. *et al.* STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–452 (2015).
48. Pradines, J. R., Farutin, V., Rowley, S. & Dancik, V. Analyzing protein lists with large networks: Edge-count probabilities in random graphs with given expected degrees. *J. Comput. Biol.* **12**, 113–128 (2005).
49. Kuleshov, M. V. *et al.* Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–97 (2016).
50. Chen, E. Y. *et al.* Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **14**, 128 (2013).
51. Zhou, G. & Xia, J. OmicsNet: A web-based tool for creation and visual analysis of biological networks in 3D space. *Nucleic Acids Res.* **46**, W514–w522 (2018).
52. Venables, W. & Ripley, D. An Introduction to R. <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>. (2020).



## Acknowledgements

This research was supported by the Basic Science Research Program of the National Research Foundation of Korea (NRF), which is funded by the Korean Ministry of Science, ICT and Future Planning [Grant number NRF-2019R1C1C1003524].

## Author contributions

M.S.K. planned the study. M.S.K., H.H.L. and J.M.C. designed the study and drafted the manuscript. M.S.K., H.P.S., J.W.J., J.Y.Y. and H.H.L. conducted data collection, analysis, and interpretation. All authors revised the manuscript critically for important intellectual content and approved the final version.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-67801-0>.

**Correspondence** and requests for materials should be addressed to M.S.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020