Introduction

Septic arthritis and gout present themselves in patients with similar physiological symptoms. Gout is primarily caused by hyperuricaemia, the acute build-up of monosodium urate crystals. The combination of excess urate production and a decrease in excretion results in crystal development. (Perez-Ruiz and Dalbeth 2019)The deposition of these crystals and absorption in macrophage inflammasomes (Martinon et al. 2006) can cause acute inflammation in the joints of the patients, primarily in the feet. (Dalbeth et al. 2013)

Septic arthritis present itself in joints as well, however this inflammation is caused by a bacterial infection. The synovial fluid in joints suffers from acute bacteraemia causing the infiltration of inflammatory cells that release cytokines to clear the infection. Cytokines binding to cartilage leads to permanent bone loss. (Goldenberg 1998) *Staphylococcus aureus* is the most common causative pathogen, with methicillin-resistant *S. aureus* cases increasing amongst the elderly. (García-Arias, Balsa and Mola 2011)

In this report, we will examine transcriptomic data from whole blood samples from 27 subjects, where the thirds are healthy controls, Gout patients, and Septic arthritis (SA) patients. The aim of this report is to find differentially expressed genes between these groups to develop a rapid test to differentiate between Gout and SA patients to aid in their recovery.

Methods

* Principal component analysis

This was carried using prcomp() {stats} on the transposed raw expression table to gain a rough insight into where PC1/2 examine relative placement of the sample groups that account for the greatest variation in the data.

* Summary statistics

The groups gene expression and Log2Fold change were filtered for p-values < 0.05 and absolute fold changes greater than 1 for Gout and 2 for SA were carried on the data sets. This created a shortlist of potential candidate genes to accomplish our aims. These genes were further filtered by a coefficient of variation < 0.02 within the group expression counts to avoid outlier-influenced genes.

* T-tests

Unpaired T-tests were carried using the R function stat\_comare\_means() {ggpubr}, the p-values generated can be seen in the boxplot figures 3 -5.

* GLM

Using both the glm() {stats} and logistf() {logistf}, GLMs were generated and model diagnostics were carried out with autoplot() to generate potential predictive models.

Results

*Principal component analysis*

A PCA was carried out on the expression table that contains the gene expression measurements of roughly 30,000 genes from the 27 patients. This was to initially verify if the first two components could distinguish between the sample groups. Figure 1 shows that the SA group is separate from the HC and Gout groups, and the latter groups appear to overlap. This could mean that the underlying pattern in expression shows HC and Gout samples are more similar to each other, where SA sample gene expression is distinct.



*Clinical Sample information*



The information from the study patients shows a balance in the sexes, the neutrophil count was taken from the whole blood samples. These counts can be seen in Figure 2, which shows a great overlap in the counts for HC and Gout and that there is a significant portion of SA patients that have a count above 10. This difference in distribution in neutrophil counts between the two conditions coincides with the acute inflammatory response of a septic infection versus the chronic inflammation of Gout.

*Significant Differential expression analysis*



In Figure 3, significant genes shortlisted by the summary statistics were plotted in boxplots and unpaired t-tests were carried out amongst HC and Gout groups. The significant genes are *PCP2*, *SNHG25, AC010300.1, Z82217.1, SMKR1,* and *SNORA73B.* Several outliers can be seen in certain Gout samples; this is perhaps why they appeared significant in the summary statistics. There is a homogenous mixture of both down and up-regulated genes compared to the HC group. The neutrophil counts are not significantly different compared to the control.



In Figure 4, significant genes shortlisted by the summary statistics were plotted in boxplots and unpaired t-tests were carried out amongst HC and SA groups. All the genes in the figure, apart from the two down-regulated genes *FAR2* and *PECR,* are significantly different. There is a clear separation and upregulation in these genes compared to HC and the neutrophil counts are also significantly increased.



In Figure 5, significant genes shortlisted by the summary statistics were plotted in boxplots and unpaired t-tests were carried out amongst SA and Gout groups. The figure contains genes that were also present in Figure 4; *S100A8, MPZL2, PDZK1IP1, TPBG, FAM110C,* and *C10orf99.*  Neutrophils are also significantly different between both groups, although not to the same extent as in Figure 4.

Discussion

There is a clear distinction in the gene expression between the two disease groups and the healthy controls. This study aimed to ascertain the significantly differentially expressed gene to develop some predictive test to distinguish between these two similar physiological symptoms. A generalized linear model using a binomial link function was used to predict whether certain genes had predictive weight to extricate between the disease groups and a healthy baseline set by the controls. The H0 was that the level of gene expression predicts for the healthy control group. The Halt was that gene expression predicts for either one the disease groups being tested. The function can be seen below:

*Sample Groupi*~ *Bern*(*Gene expressioni*)

Some of the genes showed clear distinctions in gene expression between sample groups with minor variation from their respective means. Yet for many of these clear cases, an error arose whilst generating the GLM. The binomial approximation of the sample group (response) from either one or many gene expressions (explanatory) fail due to data appearing to be quasi-separated. This results in the model output having p-values of 1 and a z-score 0 and the predictive power of the model disappearing. The Hauck-Donner effect could explain why the Wald tests, in this case, are failing as the estimates are trying to be fitted near zero and that “a truly large effect contaminated by the Hauck-Donner effect might be construed as being nonsignificant”. (Yee 2021) This is in part due to low sample sizes (n=9) in each sample group, this issue was discussed in the original paper identifying the effect. (Hauck and Donner 1977) A potential solution instead of using a binary GLM is to use a Likelihood ratio test as outlined in Xing et al. (2012) as they found the Wald test missed out on potentially significant explanatory variables in their GWAS study.

In this report, unpaired T-tests were used to compare the disease groups and to the HC gene expression as seen in figures 3 to 5. The significant genes found in Figure 3 have been linked to Gout patients or similar gene families such as the Long Non-Coding (LNC) genes that will be discussed shortly. Firstly, Transferrin (TF) appears to be upregulated in the gout group. This iron transporter has been shown to be upregulated in other transcriptomic studies, alongside Ferritin and elevated serum iron levels. (Fatima et al. 2018; Yuan and Larsson 2020) Iron and urate are linked biochemically as iron can modulate xanthine oxidase and lead to increased purine degradation. (Ghio et al. 2002) LNCs such as *AC010300.1, Z82217.1, SNORA73B,* and *SNHG25* can be seen to be both down/up regulated compared to HC. Similar LNCs have been shown to be upregulated by Zhong et al. (2018) and linked to innate immunity pathways TNF and NF-kB. A possible explanation is that the LNCs observed could be being expressed to aid in the recruitment of interleukins such as IL-8 discussed in Zhong et al. (2018) These LNCs need to be further characterised as they could be potential markers for gout.

The significant genes found in Figure 4, are all linked to the recruitment and signalling of both the innate and adaptive immune systems. The neutrophil counts are also significantly increased in SA patients, with *S100A8* being expressed in neutrophils and monocytes. (Wang et al. 2018) The gene encodes a calcium-binding protein involved in cytokine secretion and lymphocyte recruitment. *PDZK1IP1* encodes MAP17 which has been shown to activate IL-6, promote dendritic cell differentiation, and be involved in chronic inflammation. (Garcia-Heredia and Carnero, 2017)

*EHF*, *NOD2*, and *HPSE* encode for Ets homologous factor, NOD-like receptor, and Heparanase respectively. All these proteins have been found to be upregulated in inflamed tissue. Furthermore, their interactions between TLR4, TNF-a, IL-1/6, and CXCL1/6 promote the chemotaxis of macrophages and neutrophils, epithelial cell migration and healing in conjunction with *S100A8*, and bacterial peptide sensing. (Flossum et al. 2014; Flossum et al. 2017; Schmidt et al. 2012; Blich et al. 2013; Caruso et al. 2014)

In conclusion, the aim of this report was to find a transcriptomic approach to distinguish these two forms of inflammatory arthritis. As shown in Figure 5, possible diagnostic candidates for SA are genes involved in the production of cytokines/chemokines, and recruitment of innate immunity cells, which could be combined with lymphocyte count data. Possible candidates for Gout diagnosis could be genes involved in Iron homeostasis and these LNCs (although further characterisation is needed). A possible future study with a larger sample size and more patient information will allow us to verify these initial findings and will enable the development of a predictive model to help patient outcomes, especially in severe cases of SA.

References

Blich, M., Golan, A., Arvatz, G., Sebbag, A., Shafat, I., Sabo, E., Cohen-Kaplan, V., Petcherski, S., Avniel-Polak, S., Eitan, A., Hammerman, H., Aronson, D., Axelman, E., Ilan, N., Nussbaum, G. and Vlodavsky, I., (2013). Macrophage activation by heparanase is mediated by TLR-2 and TLR-4 and associates with plaque progression. *Arteriosclerosis, Thrombosis, and Vascular Biology* [online]. **33**(2). [Viewed 30 October 2022]. Available from: doi: 10.1161/atvbaha.112.254961

Caruso, R., Warner, N., Inohara, N. and Núñez, G., (2014). NOD1 and NOD2: signaling, host defense, and inflammatory disease. *Immunity* [online]. **41**(6), 898–908. [Viewed 30 October 2022]. Available from: doi: 10.1016/j.immuni.2014.12.010

Dalbeth, N., Kalluru, R., Aati, O., Horne, A., Doyle, A. J. and McQueen, F. M., (2013). Tendon involvement in the feet of patients with gout: a dual-energy CT study. *Annals of the Rheumatic Diseases* [online]. **72**(9), 1545–1548. [Viewed 30 October 2022]. Available from: doi: 10.1136/annrheumdis-2012-202786

Fatima, T., McKinney, C., Major, T. J., Stamp, L. K., Dalbeth, N., Iverson, C., Merriman, T. R. and Miner, J. N., (2018). The relationship between ferritin and urate levels and risk of gout. *Arthritis Research & Therapy* [online]. **20**(1). [Viewed 30 October 2022]. Available from: doi: 10.1186/s13075-018-1668-y

Fossum, S. L., Mutolo, M. J., Tugores, A., Ghosh, S., Randell, S. H., Jones, L. C., Leir, S.-H. and Harris, A., (2017). Ets homologous factor (EHF) has critical roles in epithelial dysfunction in airway disease. *Journal of Biological Chemistry* [online]. **292**(26), 10938–10949. [Viewed 30 October 2022]. Available from: doi: 10.1074/jbc.m117.775304

Fossum, S. L., Mutolo, M. J., Yang, R., Dang, H., O'Neal, W. K., Knowles, M. R., Leir, S.-H. and Harris, A., (2014). Ets homologous factor regulates pathways controlling response to injury in airway epithelial cells. *Nucleic Acids Research* [online]. **42**(22), 13588–13598. [Viewed 30 October 2022]. Available from: doi: 10.1093/nar/gku1146

García-Arias, M., Balsa, A. and Mola, E. M., (2011). Septic arthritis. *Best Practice & Research Clinical Rheumatology* [online]. **25**(3), 407–421. [Viewed 30 October 2022]. Available from: doi: 10.1016/j.berh.2011.02.001

García-Heredia, J. M. and Carnero, A., (2017). The cargo protein MAP17 (PDZK1IP1) regulates the immune microenvironment. *Oncotarget* [online]. **8**(58), 98580–98597. [Viewed 30 October 2022]. Available from: doi: 10.18632/oncotarget.21651

Ghio, A. J., Kennedy, T. P., Stonehuerner, J., Carter, J. D., Skinner, K. A., Parks, D. A. and Hoidal, J. R., (2002). Iron regulates xanthine oxidase activity in the lung. *American Journal of Physiology-Lung Cellular and Molecular Physiology* [online]. **283**(3), L563—L572. [Viewed 30 October 2022]. Available from: doi: 10.1152/ajplung.00413.2000

Goldenberg, D. L., (1998). Septic arthritis. *The Lancet* [online]. **351**(9097), 197–202. [Viewed 30 October 2022]. Available from: doi: 10.1016/s0140-6736(97)09522-6

Hauck, W. W. and Donner, A., (1977). Wald's test as applied to hypotheses in logit analysis. *Journal of the American Statistical Association* [online]. **72**(360a), 851–853. [Viewed 30 October 2022]. Available from: doi: 10.1080/01621459.1977.10479969

Martinon, F., Pétrilli, V., Mayor, A., Tardivel, A. and Tschopp, J., (2006). Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature* [online]. **440**(7081), 237–241. [Viewed 30 October 2022]. Available from: doi: 10.1038/nature04516

Perez-Ruiz, F. and Dalbeth, N., (2019). Gout. *Rheumatic Disease Clinics of North America* [online]. **45**(4), 583–591. [Viewed 30 October 2022]. Available from: doi: 10.1016/j.rdc.2019.08.001

Schmidt, E. P., Yang, Y., Janssen, W. J., Gandjeva, A., Perez, M. J., Barthel, L., Zemans, R. L., Bowman, J. C., Koyanagi, D. E., Yunt, Z. X., Smith, L. P., Cheng, S. S., Overdier, K. H., Thompson, K. R., Geraci, M. W., Douglas, I. S., Pearse, D. B. and Tuder, R. M., (2012). The pulmonary endothelial glycocalyx regulates neutrophil adhesion and lung injury during experimental sepsis. *Nature Medicine* [online]. **18**(8), 1217–1223. [Viewed 30 October 2022]. Available from: doi: 10.1038/nm.2843

Wang, S., Song, R., Wang, Z., Jing, Z., Wang, S. and Ma, J., (2018). S100A8/A9 in inflammation. *Frontiers in Immunology* [online]. **9**. [Viewed 30 October 2022]. Available from: doi: 10.3389/fimmu.2018.01298

Xing, G., Lin, C.-Y., Wooding, S. P. and Xing, C., (2012). Blindly using wald's test can miss rare disease-causal variants in case-control association studies. *Annals of Human Genetics* [online]. **76**(2), 168–177. [Viewed 30 October 2022]. Available from: doi: 10.1111/j.1469-1809.2011.00700.x

Yee, T. W., (2021). On the hauck–donner effect in wald tests: detection, tipping points, and parameter space characterization. *Journal of the American Statistical Association* [online]. 1–12. [Viewed 30 October 2022]. Available from: doi: 10.1080/01621459.2021.1886936

Yuan, S. and Larsson, S., (2020). Causal associations of iron status with gout and rheumatoid arthritis, but not with inflammatory bowel disease. *Clinical Nutrition* [online]. **39**(10), 3119–3124. [Viewed 30 October 2022]. Available from: doi: 10.1016/j.clnu.2020.01.019

Zhong, X., Peng, Y., Liao, H., Yao, C., Li, J., Yang, Q., He, Y., Qing, Y., Guo, X. and Zhou, J., (2019). Aberrant expression of long non‑coding RNAs in peripheral blood mononuclear cells isolated from patients with gouty arthritis. *Experimental and Therapeutic Medicine* [online]. [Viewed 30 October 2022]. Available from: doi: 10.3892/etm.2019.7816

Appendix

