Investigation into transcriptomic biomarkers for Gout and Septic Arthritis in humans

Introduction

Gout and Septic Arthritis (SA) are both diseases of the joints, though they differ in their primary pathophysiology. Gout is caused by the accumulation of uric acid there, whereas in SA it is the bacteria that appears. Both as interpreted alike as infection, triggering an immune reaction. Identification and treatment of SA is particularly important, as it can result in death regardless of the further development into sepsis (Gupta et al, 2001; Margaretten et al, 2007)

Clinical presentation of Gout and SA

Clinical symptoms for patients can be deceivingly similar, with varied intensity, between those two diseases. The diagnosis of Gout or SA is thus multifactorial, based on the likelihood ratio of the disease derived from the symptoms, imaging, and laboratory results (Eisenberg et al, 1984; Margaretten et al, 2007).

Both goat and septic arthritis trigger an immune response, leading to fever, pain, swelling, warmth, and limited movement of the affected joint. SA's identification with 50-90% confidence is achieved by confirming the bacteria's presence within the synovial fluid of the joint (Goldenberg, 1998; Weston et al, 1999). In parallel, serum uric acid levels are used to diagnose Gout. However, some laboratory results are non-specific for those diagnoses, notably an elevated neutrophil presence, white blood count, C-reactive protein, and erythrocyte sedimentary rate (Momodu and Savaliya, 2024). Imaging studies such as MRI and radiographs, although also not concluding, contribute to the confidence of SA diagnosis (Momodu and Savaliya, 2024). Hence, recent studies have been directed toward a search for a metabolic or transcriptomic biomarker (Deirmengian et al, 2014; Lee et al, 2024; Schultz et al, 2019). However, there remains a need for an accurate biomarker to distinguish SA from Gout, as well as other arthritis.

Hypothesis and aim of the study

This study aims to uncover the differential genetic profile of gout and septic arthritis in relation to healthy individuals. By analyzing the RNA sequence, we limit our search to the transcriptome. If found, the distinctive gene expression can be used to aid the development of a reliable clinical diagnostic process. The genetic identifier could distinguish between Gout and SA with similar symptomatic presentations in the hospital environment when imaging and laboratory results are insufficient.

Materials and methods

The RNA-sequencing results under analysis were derived from single samples from individuals presenting as healthy (control group), with gout or septic arthritis; based on their disease, they were split into three groups of 9 samples each. The samples were anonymized with their sex provided. There was a close-to-equal ratio between females (n=14) and males (n=13). The differential gene expression was examined with RStudio@ using the code provided. The clinical presentation, illustrated by the neutrophil count, was compared between groups to verify homogeneity and variance within them. Subsequently, due to the small sample, the data was checked visually for normality. The relevant genes were identified from the differential gene expression tables of septic patients and patients with gout against healthy individuals based on the fold change and p-value. Results for specific genes were normalised against respective maximum values to enable comparisons of change in expression between groups and genes. Finally, using correlation and multiple t-tests, the dependence of their expression on sex and neutrophil count was examined.

Results

The neutrophil count is different between groups, but likely not genders

The data for neutrophil count didn't have an equal variance between sample groups and genders (***Figure 1***). The neutrophil count significantly differed between groups (n=14 each, 5.04 +- 2.58, 7.48 +- 3.45, 12.78 +- 2.69, for healthy, Gout and SA, respectively; Kruskal-Wallis: $H = 14.69$, $p < .001$, $df=2$), as presented in ***Figure 1.1***. Although the post-hoc Dunn's test showed no significant difference between Gout and healthy groups ($Z = 1.16$, $p > .05$), the neutrophil count was higher in SA than in Gout ($Z=-2.58$, $p< .05$) and in healthy group ($Z = -3.74$, $p< .001$). Even though there was no significant difference in neutrophil count between genders (n=14 for females, n=13 for males; Kruskal Wallis: Chi-squared = 0.00059, df=1, $p> .05$, ***Figure 1.2***), the distribution of neutrophil counts was clearly visually different between groups for both genders (***Figure 1.3***), possibly reflecting a varied response to the disease.

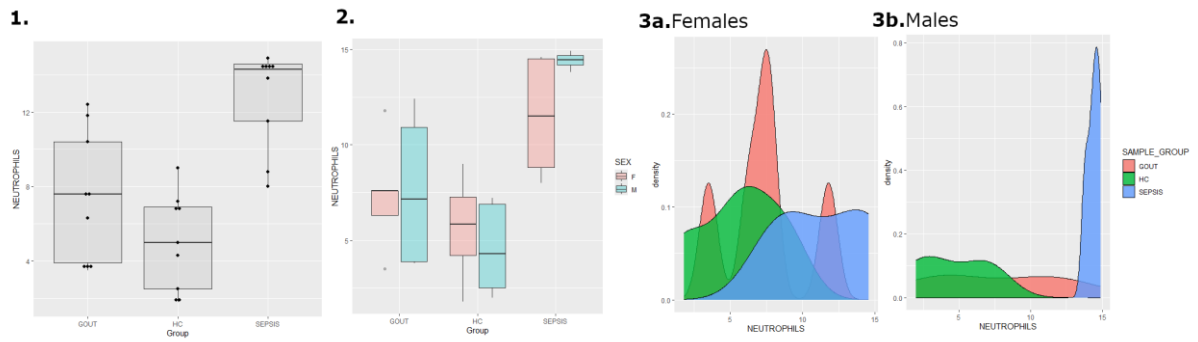


Figure 1. Dependence of the neutrophil counts on disease and sex. ***Figure 1.1.*** *The neutrophil count varies significantly between groups (n=14 each, 5.04 +- 2.58, 7.48 +- 3.45, 12.78 +- 2.69, for healthy, Gout and Septic Arthritis (SA), respectively; Kruskal-Wallis: H = 14.69, p < .001, df=2), being notably higher in Septic Arthritis than under other conditions (Dunn's post hoc test vs. Gout: Z=-2.58, p< .05; vs. healthy: Z = -3.74, p< .001). Data as boxplot with median and individual values marked.* ***Figure 1.2.*** *Even though the variance varied between sexes within groups, the statistical test detected no significant difference between them (Kruskal Wallis: Chi-squared = 0.00059, df=1, p> .05). The spread of the datapoints is the highest in Gout and lowest in Sepsis for males, whereas the opposite trend is visible for females. Data as boxplot with median and individual values marked.* ***Figure 1.3. Density plots of the distribution of the neutrophil counts between groups for females (A) and males (B) separately.*** *The split of the individual data based on sex delineates the differences between genders. In females, the count appears high in Sepsis, low in health, with a varied count in Gout. While the general trend stands true for males, the high count in sepsis has lower probability of occurring in healthy individuals.*

Genes differentially expressed in Gout

A total of 69 genes were expressed differently in Gout than in healthy individuals (provided data, p<.05). After assuming the threshold of the log2fold difference equal to the absolute value of 1, 15 genes remained – 10 upregulated and 5 downregulated in regards to the healthy baseline. As high confidence in the differential expression was crucial for achieving the goal of the study, top 10 genes with the lowest p value were isolated; they comprised the table of the top 10 most significant genes differentially expressed in gout (**Table1**).

Table 1. Genes with top 10 most significant differential expression in Gout vs. in healthy individuals. *The position of genes spanned across 8 chromosomes, including sex chromosome X. Expression of 4 genes at least doubled, with SULT4A1, KLHDC7A and EGFL6 increasing by more than 7 times. The fold difference was downregulated genes was smaller, with all among top 10 decreasing by less than half.*

Gene ID	Symbol	Chromosome	Start position	Stop position	Fold change	p	p.adj
ENSG00000064393	HIPK2	7	139,561,570	139,777,778	1.78	p<.0001	0.006
ENSG00000071909	MYO3B	2	170,178,145	170,655,171	2.55	p<.0001	0.006
ENSG00000118096	IFT46	11	118,544,528	118,572,970	0.75	p<.0001	0.006
ENSG00000130540	SULT4A1	22	43,824,509	43,862,503	9.09	p<.0001	0.008
ENSG00000170385	SLC30A1	1	211,571,568	211,578,742	1.48	p<.0001	0.008

ENSG00000179023	KLHDC7A	1	18,480,982	18,486,126	7.60	p<.0001	0.006
ENSG00000198363	ASPH	8	61,500,556	61,714,640	1.61	p<.0001	0.006
ENSG00000198759	EGFL6	X	13,569,605	13,633,575	8.22	p<.0001	0.008
ENSG00000200087	SNORA73B	1	28,508,559	28,508,762	0.32	p<.0001	0.006
ENSG00000205683	DPF3	14	72,619,296	72,894,116	0.58	p<.0001	0.006

In this subset, most genes were upregulated with the mean fold change reaching almost 5 (mean fold change= 4.6, 7 genes), whereas the fold change for downregulated genes was lower – their expression about halved (mean fold change = 0.55, 3 genes). The position of the genes was not isolated to any chromosome and included a sex chromosome X (**Table1**).

The expression values for those genes didn't differ significantly between sexes (multiple t-test with Bonferroni correction, p>.05, ***Figure 2.A***). Five genes also significantly differed () between Sepsis Arthritis and Gout conditions: ENSG00000064393 (F=-6.47, p<.0001), ENSG00000071909 (multiple t-tests with Bonferroni correction, ***Figure 2.B***; F=5.39, p<0.01), ENSG00000118096 (F=-4.2, p<.01), ENSG00000198363 (F=4.63, p<.01), ENSG00000200087 (F=-6.24, p<.001).

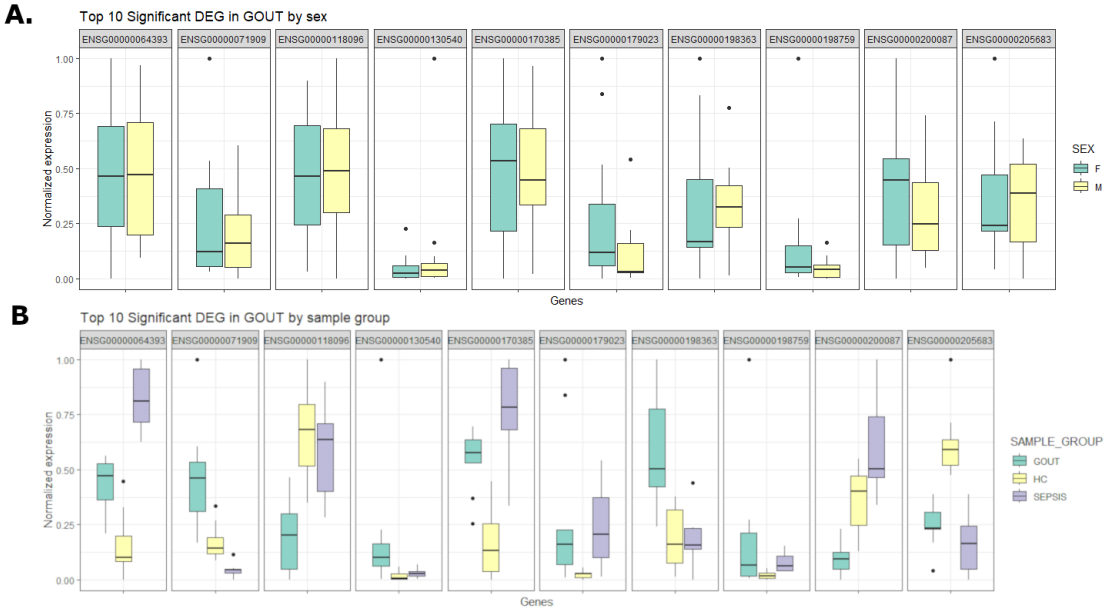


Figure 2. Expression of the top 10 most significant genes in Gout. *The genes in question are expressed similarly between sexes (**Figure 2A**,multiple t tests with Bonferroni correction, p>= .05, data normalised to max value per gene), but very differently in the course of GOUT than in healthy individuals (**Figure 2B**;GOUT vs HC, n=9 per group, p<.05).*

Genes differentially expressed in Sepsis Arthritis

More genes were affected in Sepsis Arthritis than in Gout, with a total of 13 046 genes exceeding the significance threshold. Adapting additional threshold of the log2fold change of absolute value of 1 limited that number to 6 270 genes. Contrary to Gout, a majority of them showed downregulation (4 112 genes), about a third of them was downregulated (2 158

genes). To isolate the genes with the highest confidence in the difference, the results were sorted by the p value to obtain top 10 genes with the most significant difference in expression against healthy controls (***Table 2***).

Table 2. Genes with top 10 most significant differential expression in SA vs. in healthy individuals. *The position of genes spanned across 9 chromosomes. All genes with the most significant result were upregulated when compared to the healthy individuals. Expression of all but two genes (TMEM165, TGM1) was increased at least 20 times.*

Gene ID	Symbol	Chromosome	Start position	Stop position	Fold change	p	p.adj
ENSG00000092295	TGM1	14	24,249,114	24,264,432	6.86	p<0.0001	p<0.0001
ENSG00000115919	KYNU	2	142,877,498	143,055,832	33.32	p<0.0001	p<0.0001
ENSG00000124102	PI3	20	45,174,876	45,176,544	440.65	p<0.0001	p<0.0001
ENSG00000134827	TCN1	11	59,852,800	59,866,575	332.03	p<0.0001	p<0.0001
ENSG00000134851	TMEM165	4	55,395,957	55,453,397	2.89	p<0.0001	p<0.0001
ENSG00000135114	OASL	12	121,019,111	121,039,242	51.07	p<0.0001	p<0.0001
ENSG00000136688	IL36G	2	112,973,203	112,985,665	26.69	p<0.0001	p<0.0001
ENSG00000140519	RHCG	15	89,471,398	89,496,613	48.29	p<0.0001	p<0.0001
ENSG00000188373	C10orf99	10	84,173,738	84,185,294	30.86	p<0.0001	p<0.0001
ENSG00000198074	AKR1B10	7	134,527,592	134,541,408	86.47	p<0.0001	p<0.0001

All of those genes were upregulated, with median fold change equal 40.8. Two genes, ENSG00000124102 and ENSG00000134827, showed the highest difference of 441 and 332 times increase, respectively. The expression of those genes per group and sex is shown in **Figure 3.** No genes showed a significant difference between genders (multiple t tests with Bonferroni correction, p>= .05). However, the result for individuals with Septic Arthritis was significantly higher than for other groups (SA vs HC: result provided, p<.05; SA vs Gout: multiple t tests with Bonferroni correction, p<.05), with every recording exceeding those of other groups.

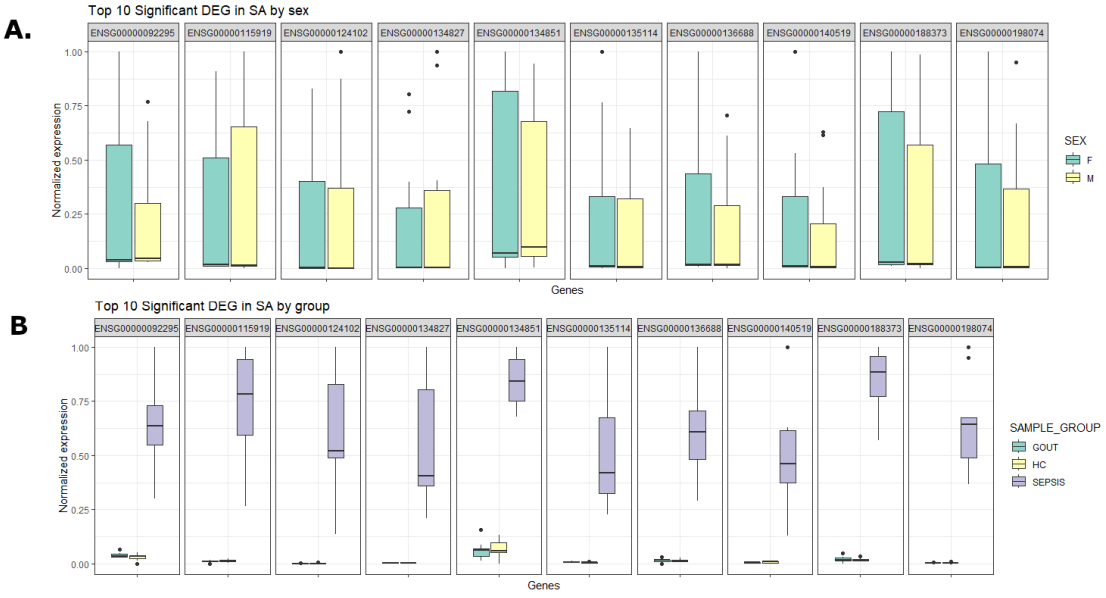


Figure 3. Expression of the top 10 most significant genes in SA. *The genes in question are expressed similarly between sexes with similar variance (**Figure 3A**, multiple t tests with Bonferroni correction, p>= .05), but very differently in the course of Sepsis Arthritis than in healthy individuals (**Figure 3B**; SA vs HC, n=9 per group, p<.05, data normalised to max value per gene). For those genes, the expression in healthy individuals or those with Gout is negligible in comparison to the SA group.*

Subset of genes uniquely altered in Gout, while none in SA

Among significant genes detected that passed the log2fold threshold, 47 genes overlapped between the diseases – 26 downregulated in both, 15 upregulated in both, 4 upregulated in Gout and downregulated in SA, 2 downregulated in Gout and upregulated in SA. No genes were significantly different in Sepsis only, but 22 such genes were detected for Gout (**Table 3**). In accordance with previous trends, most of those genes are upregulated. Five genes have a magnitude of the fold change of at least 2. Three genes are uniquely altered in Gout and have obtained one of the lowest p values when compared to the healthy controls: ENSG00000118096, ENSG00000198363, ENSG00000130540.

Table 3. Genes with a significant change in expression for Gout, but not for Sepsis Arthritis. *A total of 22 genes were detected to change their expression uniquely for Gout (log2fold >1, p< .05), with a slight majority of them being upregulated (13 genes). The position of genes spanned across 16 chromosomes, including sex chromosome X. Expression of 5 genes at least doubled, whereas the average fold change for the downregulated genes was smaller – equal to the 30% reduction.*

GENE_ID	symbol	chromosome	start	stop	Fold change	p	p.adj.
ENSG00000130540	SULT4A1	22	43,824,509	43,862,503	9.09	0.00000	0.008
ENSG00000091513	TF	3	133,745,956	133,796,640	4.54	0.00004	0.028
ENSG00000214708	AC116407.1	17	32,141,226	32,143,135	3.52	0.00004	0.028
ENSG00000279652	Z82217.1	22	35,992,321	36,000,469	2.40	0.00014	0.046
ENSG00000102359	SRPX2	X	100,644,166	100,675,788	2.00	0.00010	0.040
ENSG00000150938	CRIM1	2	36,355,926	36,551,135	1.67	0.00003	0.028
ENSG00000198363	ASPH	8	61,500,556	61,714,640	1.61	0.00000	0.006
ENSG00000163291	PAQR3	4	78,887,127	78,939,438	1.58	0.00003	0.028
ENSG00000138593	SECISBP2L	15	48,988,476	49,046,563	1.43	0.00005	0.028
ENSG00000101236	RNF24	20	3,927,309	4,015,582	1.37	0.00018	0.050
ENSG00000138032	PPM1B	2	44,167,969	44,244,384	1.33	0.00006	0.029
ENSG00000111897	SERINC1	6	122,443,354	122,471,822	1.32	0.00018	0.050
ENSG00000112941	TENT4A	5	6,713,007	6,757,048	1.32	0.00011	0.040
ENSG00000086589	RBM22	5	150,690,794	150,701,107	0.84	0.00006	0.029
ENSG00000139546	TARBP2	12	53,500,921	53,506,431	0.78	0.00006	0.030
ENSG00000118096	IFT46	11	118,544,528	118,572,970	0.75	0.00000	0.006
ENSG00000129465	RIPK3	14	24,336,021	24,340,045	0.73	0.00005	0.028
ENSG00000149196	HIKESHI	11	86,302,211	86,345,931	0.72	0.00004	0.028
ENSG00000136161	RCBTB2	13	48,488,959	48,533,256	0.72	0.00010	0.040
ENSG00000255717	SNHG1	11	62,851,988	62,855,914	0.66	0.00015	0.049
ENSG00000130701	RBBP8NL	20	62,410,237	62,427,533	0.61	0.00017	0.050

ENSG00000240204	SMKR1	7	129,502,479	129,512,932	0.49	0.00013	0.044

None of those genes showed a variation between sex (multiple t-test with Bonferroni correction, p>.05), but all differed in Gout but not in SA vs. healthy individuals. All but two of those - ENSG00000130540, ENSG00000091513 – have contrasting expression between Gout and Sepsis (multiple t-test with Bonferroni correction, p<.05).

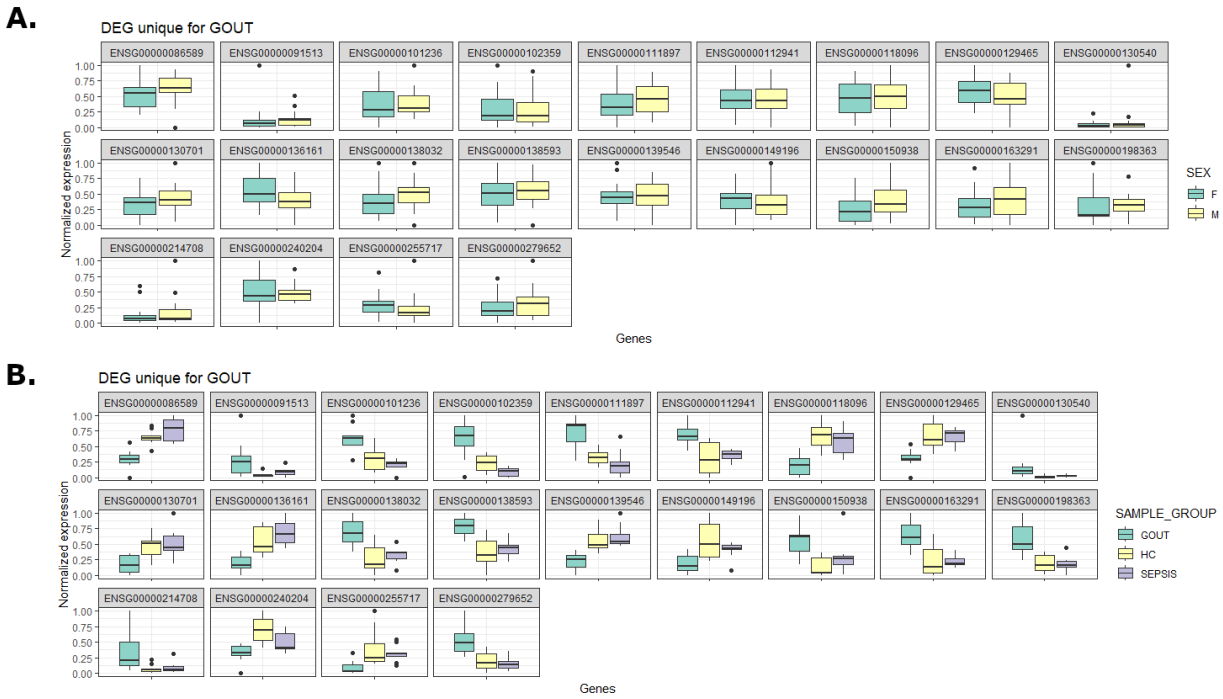


Figure 2. Expression of the genes significantly altered in GOUT, but not in Sepsis Arthritis, in comparison to the healthy individuals. *The genes in question are expressed similarly between sexes with similar variance (**Figure 2A**,multiple t tests with Bonferroni correction, p>= .05). The genes are altered significantly in comparison to the healthy individuals (p<.05, data provided for assessment) with detailed p value and fold change displayed in **Table 2**. In all but two cases (ENSG00000130540, ENSG00000091513; p>=.05), the normalised expression in individuals with Gout is significantly different than in those with Sepsis Arthritis (multiple t tests with Bonferroni correction, p<.05) but very differently in the course of Sepsis Arthritis than in healthy individuals (**Figure 2B**; SA vs HC, n=9 per group, p<.05, data normalised to max value per gene). For those genes, the expression in healthy individuals or those with Gout is negligible in comparison to the SA group. DEG – differential gene expression, HC – healthy,*

Discussion

In accordance with the aims, the report identified a number of genes with the potential to be used diagnostically.

As many as 22 genes have been only affected in Gout, while not in SA; among those, the magnitude of change was at least equal to 2 for five genes: SULT4A1, TF, SRPX2, and two novel constructs of AC116407.1 and Z82217.1. The statistical significance of SULT4A1 was one of the

highest, however, its expression was also altered in SA. The same was true for TF. Hence, it is the genes for SRPX2, AC116407.1, and Z82217.1 that require further investigation towards their diagnostic potential in Gout.

Although no genes were only affected in individuals with SA, the magnitude of change for the top 10 most significant genes compared to other groups indicates their diagnostic potential. Especially noteworthy are genes PI3 and TCN1, whose expression reached the highest values in SA. The sample for further study should also include genes with contrary direction of expression change, which were not elaborated on within this report.

The results of the study are not exhaustive of the subject, nor free from limitations. The most important hindrance was the time constraints and the magnitude of data, that requires a continued effort to uncover more patterns within. Particularly, genes of interest outlined by VanItallie (2010) should be investigated. Secondly, the effect of more variables should be examined on the aforementioned genes, such as the age of participants. The current sample size of the study didn't allow for further subdivision of data within the diseases eg. Into genders. Accounting for the individual differences, the future studies should include the genetic expression prior to the diseased state for paired-comparison when possible. Nonetheless, the report lays a groundwork for future investigative efforts to find diagnostic biomarkers for the diagnosis of SA.

References

Deirmengian, C. *et al.* (2014a) “Diagnosing Periprosthetic Joint Infection: Has the Era of the Biomarker Arrived?,” *Clinical Orthopaedics and Related Research*, 472(11), pp. 3254–3262. Available at: https://doi.org/10.1007/s11999-014-3543-8.

Eisenberg, J.M. *et al.* (1984) *Usefulness of Synovial Fluid Analysis in the Evaluation of Joint Effusions Use of Threshold Analysis and Likelihood Ratios to Assess a Diagnostic Test*.

Goldenberg, D.L. (1998) “Septic arthritis,” *The Lancet*, 351(9097), pp. 197–202. Available at: https://doi.org/https://doi.org/10.1016/S0140-6736(97)09522-6.

Gupta, M.N., Sturrock, R.D. and Field, M. (2001) “A prospective 2-year study of 75 patients with adult-onset septic arthritis,” *Rheumatology*, 40(1), pp. 24–30. Available at: https://doi.org/10.1093/rheumatology/40.1.24.

Lee, B.H. *et al.* (2024) “Tryptophanyl tRNA synthetase is an alternative synovial biomarker for diagnosis of septic arthritis in knee joint,” *Knee Surgery and Related Research*, 36(1). Available at: https://doi.org/10.1186/s43019-024-00229-2.

Margaretten, M.E. *et al.* (2007) “Does This Adult Patient Have Septic Arthritis?,” *JAMA*, 297(13), pp. 1478–1488. Available at: https://doi.org/10.1001/jama.297.13.1478.

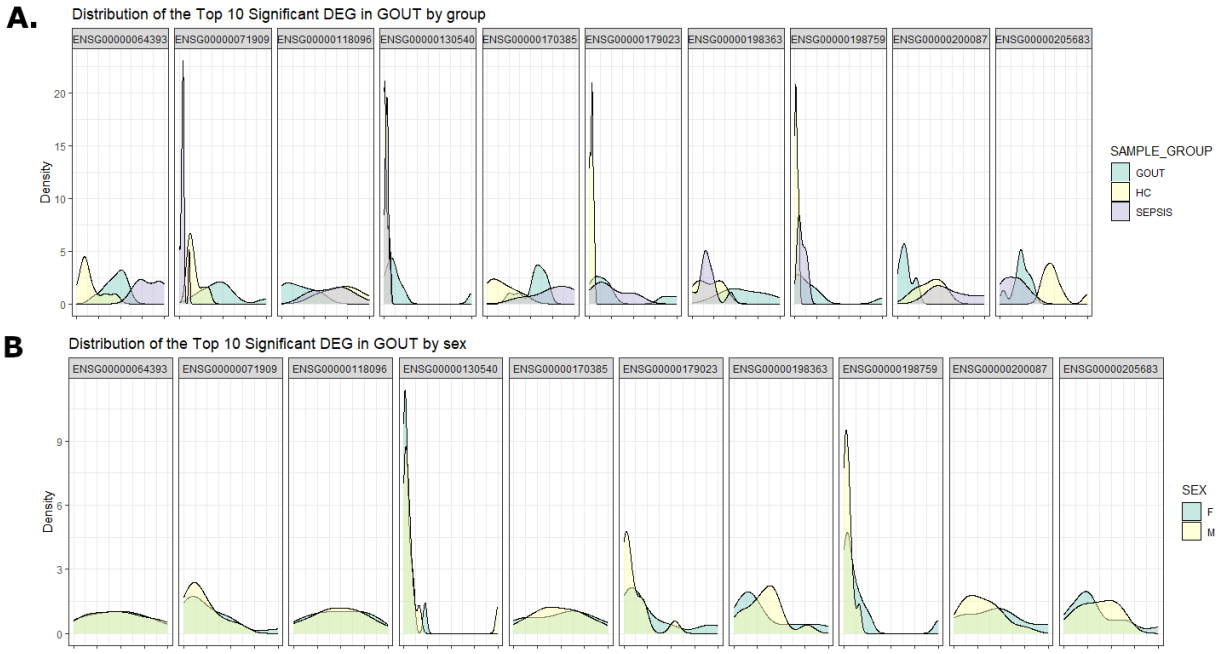
Momodu II, Savaliya V. Septic Arthritis. [Updated 2023 Jul 3]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK538176/

Schultz, B.J. *et al.* (2019a) “Pilot study of a novel serum mRNA gene panel for diagnosis of acute septic arthritis,” *World Journal of Orthopedics*, 10(12), pp. 424–433. Available at: https://doi.org/10.5312/wjo.v10.i12.424.

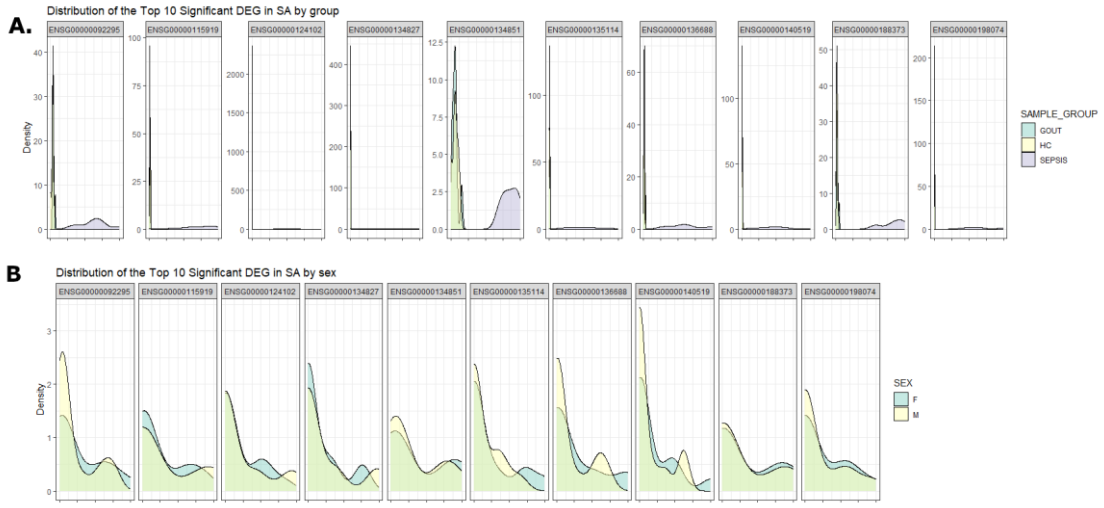
VanItallie, T.B. (2010) “Gout: epitome of painful arthritis,” *Metabolism - Clinical and Experimental*, 59, pp. S32–S36. Available at: https://doi.org/10.1016/j.metabol.2010.07.009.

Weston, V.C. *et al.* (1999) “Clinical features and outcome of septic arthritis in a single UK Health District 1982-1991,” *Annals of the Rheumatic Diseases*, 58(4), pp. 214–219. Available at: https://doi.org/10.1136/ard.58.4.214.

Appendix



Supplementary Figure 1. *Distribution of the expression of the top10 most significantly differentially expressed genes (DEG) in Gout, divided by group (A) or sex (B).*



Supplementary Figure 2. *Distribution of the expression of the top10 most significantly differentially expressed genes (DEG) in SA, divided by group (A) or sex (B).*