Potential Genetic Markers for Susceptibility to Osteoarthritis and Rheumatoid Arthritis

G. Jilani Chaudry, Tho Hua, Jamie L. Myers, Hui Xia, Manuel Y. Caballero, and Sandra Valtier

FINAL REPORT

Date: 8 November 2021

59th Medical Wing
Office of the Chief Scientist
1632 Nellis, BLDG. 5406
JBSA Lackland AFB, TX 78236-7517

DISTRIBUTION A. Approved for public release; distribution is unlimited.

DECLARATION OF INTEREST

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Air Force, Department of Defense, nor the U.S. Government. This work was funded by Project Code Number AC12EM01. Authors are military service members, employees, or contractors of the US Government. This work was prepared as part of their official duties. Title 17 USC §105 provides that 'copyright protection under this title is not available for any work of the US Government.' Title 17 USC §101 defines a US Government work as a work prepared by a military service member, employee, or contractor of the US Government as part of that person's official duties.

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

Qualified requestors may obtain copies of this report from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

STUDY TITLE

Potential Genetic Markers for Susceptibility to Osteoarthritis and Rheumatoid Arthritis

ONEAL.RUBEN.III.1145511212
Digitally signed by ONEAL.RUBEN.III.1145511212
Date: 2021.12.01 12:33:08 -06'00'

Ruben O' Neal III, DAF
Program Analyst
Program
59MDW Office of the Chief Scientist

BRINKLEY.CARLTON.CLARENCE.JR.1048285950
Digitally signed by BRINKLEY.CARLTON.CLARENCE.JR.1048285950
Date: 2021.11.17 16:32:39 -06'00'

Sandra Valtier (Carl. C Brinkley sign on behalf)
PI title and Directorate
59MDW Office of the Chief Scientist

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
8 November 2021	Final Report	July 2015 to November 2021

4. TITLE AND SUBTITLE

Study Title: Potential Genetic Markers for Susceptibility to Osteoarthritis and Rheumatoid Arthritis

5a. CONTRACT NUMBER
FA3016-16-P-0250

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

G. Jilani Chaudry, Tho Hua, Jamie L. Myers, Hui Xia, Manuel Y. Caballero, and Sandra Valtier

5d. PROJECT NUMBER
AC15HF01

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
United States Air Force, 59th Medical Wing (59MDW/ST).
1255 Wilford Hall Loop, Building 4430
Lackland Air Force Base, 78236-9980

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Osteoarthritis (OA) and rheumatoid arthritis (RA) are the two most common arthritic disorders. OA is more frequent than RA, not only in the US general population, but also in the US military. The arthritic disorders along with musculoskeletal (MSK) injuries are the most common reasons for exit from military service. Together, they impose a major healthcare burden not only on the Department of Defense while the service members remain active, but also on the Department of Veterans Affairs once they leave. The fundamental determinants of predisposition to MSK disorders such as OA and RA are genetic, but such factors as nutrition, lifestyle, activity, and physical fitness are also players. Among genetic variations, the most frequent are single nucleotide polymorphisms (SNPs), with nearly 90 million identified in the human population. Most SNPs have no pathologic effects, but some do. The pathogenic SNPs and other genetic variations exert their pathologic effects by changing the activities of genes and their products (RNA, proteins), e.g., up- or down-regulating the activities of genes, or outright disrupting the functions of RNAs or proteins.

Part I of this study aimed to evaluate the relevance of 32 SNPs to OA and RA. These SNPs occur in 21 genes, and were selected from databases and previous reports due to their putative associations with MSK disorders. The study used race and gender mixed cohorts comprising 498 OA, 442 RA, and 499 control (no arthritis) DNA samples purchased from a commercial supplier. Commercially available SNP genotyping reagents were used to determine the frequency of SNP genotypes. The genotype associations with OA or RA were determined using the SAS software. The study revealed that three SNPs show statistically significant associations with OA and ten with RA ($p < 0.05$). Three SNPs in PTPN22 and four in STAT4 showed association with RA in the general population, as well as with OA or RA in certain subgroups within the study cohort. Several other SNPs associated with OA or RA in different subgroups of the study population (race, gender, age group).

Part II focused on transcriptome expression analysis of 117 commercially available gender mixed Caucasian synovial tissue samples, 31 OA, 60 RA, and 26 control (no arthritis). The approach employed Clariom D Human Array Chip (capacity, ~ 540,000 transcripts) and Gene Chip Scanner 3000 7G System. Transcriptome Analysis Console (TAC) software was used to analyze the expression results and graphically visualize the data. This revealed 135,750 transcripts analyzed. Of these, 683 showed > 2-fold higher and 1,610 > 2-fold lower expression in the OA population relative to the control population ($p < 0.05$). The corresponding values for RA were 1,169 and 1,343 ($p < 0.05$). From these up- or down-regulated genes, we selected 23 for further analysis by TaqMan real-time PCR to compare these results with the gene chip analysis results. The three most up-regulated and the three most down-regulated genes in comparison to the control cohort included MXRA5 and DDX3Y. MXRA5 was up-regulated both in OA and RA, 207-fold and 520-fold, respectively. In contrast, DDX3Y was the most down-regulated gene in both disorders; 285-fold in OA and 279-fold in RA. Note, however, that in comparison to the control cohort, an overwhelming majority of transcripts showed < 2-fold expression difference or essentially none.

15. SUBJECT TERMS- SNP, osteoarthritis, rheumatoid arthritis, gene expression, genetic susceptibility

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT:**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	32	

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This project had a two-fold focus: **One**, to evaluate relevance of 32 single nucleotide polymorphisms (SNPs) to osteoarthritis (OA) and rheumatoid arthritis (RA), the two most common inflammatory arthritic disorders. To do this study, DNA samples of OA and RA patients were purchased from a commercial source. The control samples represented subjects without clinical diagnosis of OA or RA. **Two**, to experimentally ascertain differential gene expression profiles of synovial tissues from OA and RA patients in comparison to the control samples (no OA or RA). The synovial tissue samples were purchased from a different company; that is, not the same as for the SNP study samples. Thus, there was no overlap between the two sets of samples. Further, whereas the SNP study population was a race and gender mixed cohort, the gene expression study cohort subjects were all of Caucasian origin. Thus, the two studies were distinct. Therefore, to keep that distinction of the nature and scope of the SNP and gene expression work, the two studies are presented separately in this report, the SNP study as Part I and the gene expression study as Part II.

PART I

An Evaluation of 32 Single Nucleotide Polymorphisms for Association with Osteoarthritis and Rheumatoid arthritis

G. Jilani Chaudry[2,*], Tho Hua[1], Jamie L. Myers[1], Hui Xia[1], Manuel Y. Caballero[1], and Sandra Valtier[1]

[1]Center for Advanced Molecular Detection, [2]Science and Technology, 59 Medical Wing, JBSA-Lackland, TX

[*]**Corresponding Author:** G. Jilani Chaudry
ghulam.j.chaudry.ctr@mail.mil
210-292-7021 (office)
210-381-2894 (cell)

ABSTRACT

Osteoarthritis (OA) and rheumatoid arthritis (RA) are the two most common inflammatory arthritic disorders. OA is more frequent than RA. This holds not only for the general population in the United States, but also the US military population. The arthritic disorders together with musculoskeletal (MSK) injuries constitute the most common reasons for exit from military service. The MSK disorders and injuries impose a major healthcare burden not only on the Department of Defense while the service members remain active, but also on the Department of Veterans Affairs when they leave active duty.

The most fundamental determinants of predisposition to such MSK disorders as OA and RA are genetic. However, such factors as nutrition, lifestyle, activity, and physical fitness are also crucially important in as much as they can positively or negatively affect the disease onset and its subsequent pathologic course and severity. Many different types of human genetic variations exist, the most common type being single nucleotide polymorphisms (SNPs). Close to 90 million SNPs have been found in the human population, and some of them are relevant to diseases.

This study aimed to evaluate the relevance of 32 SNPs, which occur in 21 different genes. We used 1439 commercially available genomic DNA samples; 498 OA, 442 RA, and 499 control (no arthritis). Commercially available SNP genotyping reagents were used to determine the frequency of the SNP genotypes in the three cohorts. We then used the SAS statistical analysis software to evaluate the genotypes' associations with OA and RA.

Three SNPs showed statistically significant association with OA and ten with RA ($p < 0.05$). The p-values were 0.03 or less. In two genes, multiple SNPs showed association with RA: three SNPs in the gene PTPN22 and four in the gene STAT4. The four STAT4 SNPs occur in intron 3, and have physical proximity. Several other SNPs associated with OA or RA in different subgroups of the study population. For example, rs13301537 (ASPN gene) C/C genotype associated with both OA and RA in the Hispanic cohort, but only with RA in the African-American cohort, and it showed no significant association with either disease when the whole study population was considered. The four STAT4 SNPs analyzed associated with RA in the overall population, as well as the White cohort and the under-40 age group.

Some SNPs showed positive associations with OA or RA in the overall population or subgroups, while some others showed negative associations. The odds ratios for those SNPs that showed positive associations ranged from 1.36 to 4.93. The odds ratios for those that showed negative associations with OA or RA ranged from 0.17 to 0.3.

INTRODUCTION

It is estimated that nearly 25 % of the US adult population has clinically diagnosed arthritic disorder of one kind or another. Of the arthritic disorders, osteoarthritis (OA) and rheumatoid arthritis (RA) are the most common, OA being the most prevalent. An estimated 30 million US adults are affected (CDC, 2014). The problem is even more profound in the US Armed Forces personnel due to the high-risk tasks they must perform during deployment and in-garrison. In 2016 the US Armed Forces recorded over 2.1 million medical encounters related to musculoskeletal disease and injury, accounting for 19.2% of all medical visits and resulting in over 15,000 bed days (Hauret, 2010; MSMR, 2014, 2019).

Both OA and RA exhibit greater frequency and prevalence in the US military active-duty personnel and veterans than the general population. According to a 2014 CDC report that summarized findings for 2011-2013, for example, prevalence of arthritis of one form or another among the male veterans was 25 % and the female veterans 31.3 %, both higher than the general population (CDC, 2014). Arthritis is also among the top causes of discharge from the military service. It has been reported that OA is among the major causes for military personnel to become unfit for duty performance, and that posttraumatic OA is the top cause (Tilkeridis, 2005; Jones, 2010). The treatment for both disorders is long-term and expensive, placing a great cost burden on the military and veterans' healthcare systems. The causes for the arthritic disorders are complex, and they span genetic, epigenetic, lifestyle, and environmental factors (Musumeci, 2015; Yarwood, 2014). Precise understanding of these factors and their interactions that determine the onset and the subsequent arthritic pathologies is essential for not only understanding the biology of these disorders, but also for identifying predictive, diagnostic, and prognostic markers. (Ombrello, 2014).

Genetic variations come in different forms, e.g., single nucleotide changes (deletions, insertions, substitutions), deletions and insertions (small to large), copy number variations, and rearrangements. The simplest and most common type of variations is termed single nucleotide polymorphisms (SNPs). An estimated 85 – 90 million SNPs have already been discovered in the human population (dbSNP; 1000 Genomes Project). Most of these have no disease consequences. Those that do, can range in pathologic effects from conferring susceptibility to one or another disease to almost invariably resulting in a particular disease.

In as much as they confer susceptibility to OA, RA, or both, the genetic factors are of fundamental importance. Once they are definitively determined, the genetic factors for these disorders could be used as biomarkers of arthritic disease predisposition, onset, and pathologic severity. Application and use of such biomarkers would change the clinical practice guidelines in significant ways, undoubtedly to great benefits for the patients in terms of more effective healthcare and treatment approaches.

This study was carried out to assess the association of 32 known SNPs with putative disease associations, including OA and RA. The selected SNPs occur in 21 different genes, and they have already been identified and listed in various databases (see Table 1 footnotes and references). Meta-analyses have also been carried out to evaluate SNP associations with MSK disorders, e.g., OA of the hip (Evangelou, 2014; Jiang, 2012; Wang, 2016). This study did not aim to discover any new SNPs, nor conduct analyses using databases. Our overall approach was first to experimentally determine the prevalence of each of the 32 SNPs in the OA, RA, and control cohorts, and then to conduct statistical analyses to determine each SNP genotype's association with OA, RA, or both.

MATERIALS AND METHODS

SNPs. For this study, we selected 32 SNPs from the databases based on previous reports of positive or negative associations with OA, RA, or other related musculoskeletal conditions, e.g., ligament tears. These SNPs occur in 21 different genes. We did not select the genes *per se*. Rather, the genes included are a consequence of selecting SNPs.

Study Samples. Individual-specific purified genomic DNA samples were purchased from BioServe-REPROCELL (Beltsville, MD). All samples provided were already deidentified, and therefore the only information included in the data sheets was limited demographic information, notably gender, age, and race. The main clinical information relevant to this study was the arthritic disease diagnosis of OA, RA, or no arthritic disease (control samples). We purchased 500 samples for each of the three categories. However, the data presented here is for 499 OA, 442 RA, and 499 control samples; some of the samples could not be analyzed due to low DNA concentration or degraded DNA.

SNP Genotyping. The purified DNA sample concentrations were provided by the supplier. However, we rechecked the concentrations if needed, e.g., when SNP detection failed, and we deemed it necessary to recheck the DNA concentration. Prior to genotyping reactions, the DNA samples were diluted to 2 ng/µL. Genotyping experiments were performed using TaqMan SNP Genotyping Assay reagents (Thermo Fisher Scientific, Waltham, MA). The PCR reaction mixture included TaqMan Genotyping Master Mix, SNP Genotyping Assay mixture, and 10 ng of purified genomic DNA. SNP genotyping was carried out in duplicate in 96-well plates, and each plate also included negative controls in triplicate. For the negative controls, water was used instead of any DNA sample. The real-time PCR instruments used were 7500, ViiA7, or QuantStudio 6. The thermocycling conditions and other parameters for each machine were essentially as recommended by the manufacturer (Thermo Fisher Scientific). The thermocycling protocol was as follows: 95°C for 10 min; 40 cycles of 95°C for 15 sec and 60°C for 1 min. The number of PCR cycles for each assay was selected based on clear separation of genotype clusters, which was determined beforehand by performing pilot assays.

Statistical Analysis. Statistical Analysis Software (SAS Institute, Inc. SAS 9.3. Cary, NC) was used to determine positive or negative association of each SNP with OA or RA. Logistic regression models were performed to generate adjusted odds ratios and the corresponding 95 % confidence intervals.

RESULTS

The genomic DNA samples for this study were purchased from a commercial source, as described in Materials and Methods. No subjects were enrolled for this work. There samples were 498 OA, 442 RA, and 499 control (no arthritic disease). The overall cohort represented both male and female subjects, as well as various self-identified races and ethnicities (Figure 1). In terms of race, the samples consisted of 59.6 % Caucasian, 17.2 % African-American, 11.6 % Hispanic, and 11.6 % other, which included Native Hawaiian, White-Caucasian Native Hawaiian, Native Alaskan, Pacific Islander, Asian, and Japanese. In terms of gender distribution in the total number, the cohort comprised 46.6 % female and 53.4 % male. The self-reported genders were identified only as either female or male. We did not subject the "others" categories to statistical analysis because we considered the numbers in these groups too low for such analysis.

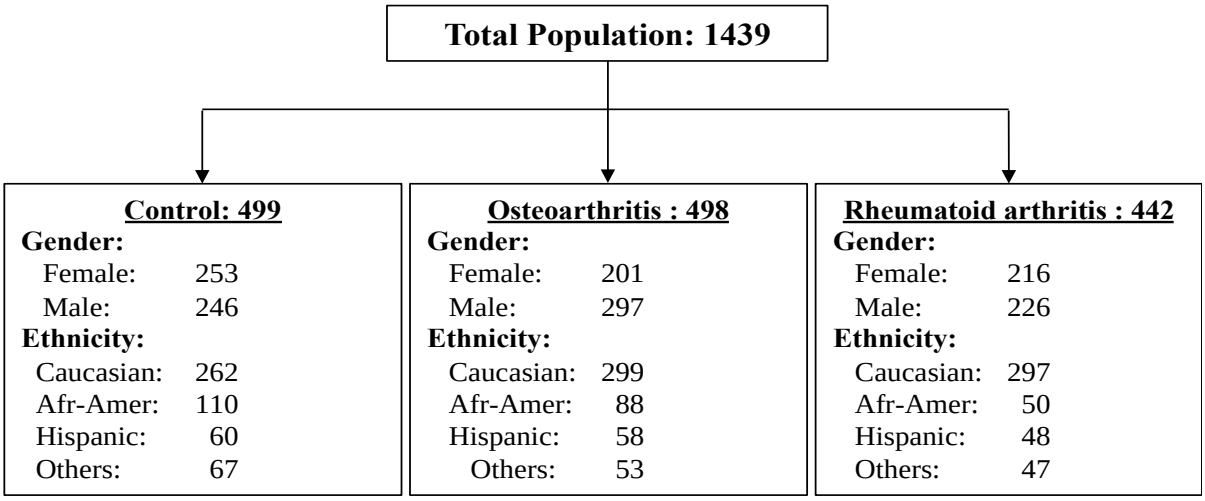


Figure 1. Composition of the overall study population. "Others" included different ethnicities reported by the study subjects, e.g., Native Hawaiian, White-Caucasian Native Hawaiian, Native Alaskan, Pacific Islander, Asian, and Japanese.

The work reported here aimed to determine whether certain SNPs associated with osteoarthritis, rheumatoid arthritis, or both. We picked 32 SNPs previously reported to show positive or negative association with one MSK disorder or another, including OA or RA. The 32 SNPs we analyzed are listed in Table 1, along with their respective genes, the nature of variations, their genomic locations, and any deduced changes in the proteins. As expected, with respect to the chromosomal structures of genes, the SNPs occur in a variety of regions – exons, introns, promoters, and extragenic regions (significantly distal to any known nearby gene). Within mRNA, SNPs occur in the coding region, 5'UTR (untranslated region), or 3'UTR. In principle, a SNP in any of these regions can result in pathologic consequences, and the nature of such consequences depends on the nature of SNP variation.

Table 1. SNPs analyzed in this study.

SNP	Gene	Locus	Chr. strand	Variation in gene + strand; region; position; any amino acid change	ABI assay ID
rs1871054	ADAM12	10q26.2	-	G or A; intron	C_12049599_10
rs13301537	ASPN	9q22.31	-	T or C; intron	C_1899078_10
rs1126499	BGN	Xq28	+	C or T; coding; 540; S180S	C_2617574_1_
rs10153674	CALM2	2p21	-	G or A; 5′ UTR	C_29914416_10
rs1676486	COL11A1	1p21.1	-	T or C; coding; 4486; S1496P	C_8400671_10
rs970547	COL12A1	6q13	-	G or A; coding; 9172; G3058S	C_7580617_10
rs1800012	COL1A1	17q21.33	-	G or T; intron	C_7477170_30
rs12722	COL5A1	9q34.3	+	C or T; 3′ UTR	C_370252_20
rs121912932	COL5A1	9q34.3	+	G or A; coding; 4466; G1489E	C_4735_10
rs61735045	COL5A1	9q34.3	+	G or A; coding;1588; G530S	C_25765014_20
rs2294984	COL9A3	20q13.33	+	G or A; coding; 50; G17E	C_16188902_10
rs61734651	COL9A3	20q13.33	+	C or T; coding; 307; R103W	C_25600748_20
rs142639450	COL9A3	20q13.33	+	G or A; coding; 308; R103Q	C_167193861_10
rs4073	CXCL8	4q13.3	+	A or T; promoter; -352	C_11748116_10
rs225014	DIO2	14q31.1	-	A or G; 3′ UTR	C_15819951_10
rs143383	GDF5	20q11.22	-	C or T; promoter; -275	C_1270479_1_
rs11549465	HIF1A	14q23.2	+	C or T; coding; 1816; P606S	C_25473074_10
rs1800796	IL6	7p15.3	+	G or C; promoter; -636	C_11326893_10
rs2227306	IL8	4q13.3	+	C or T; intron	C_11748169_10
rs679620	MMP3	11q22.2	-	A or G; coding; 133; K45E	C_3047717_1_
rs591058	MMP3	11q22.2	-	A or G; intron	C_785960_1_
rs650108	MMP3	11q22.2	-	C or T; intron	C_785965_10
rs4112788	LCE3D	1q21.3	+	A or G; 3′ UTR	C_31910050_10
rs6920220	RP11-95M15.1	6q23.3	+	G or A; intron	C_29431952_10
rs33996649	PTPN22	1p13.2	-	G or A; coding; 788; R263Q	C_25937239_30
rs2476601	PTPN22	1p13.2	-	C or T; coding; 1858; R620W	C_16021387_20
rs2488457	PTPN22	1p13.2	+	C or G; promoter; -1123	C_16027865_10
rs7574865	STAT4	2q32.2	-	A or C; intron	C_29882391_10
rs10181656	STAT4	2q32.3	-	C or G; intron	C_30530761_10
rs8179673	STAT4	2q32.3	-	G or A; intron	C_11514609-10
rs11889341	STAT4	2q32.2	-	G or A; intron	C_26419582_10
rs1800629	TNFA	6p21.33	+	G or A; promoter; -488	C_7514879_10

Notes:

1. **Database: https://www.ncbi.nlm.nih.gov/snp/; https://www.ncbi.nlm.nih.gov/variation/view/; http://useast.ensembl.org/Homo_sapiens/Info/Index?db=core; https://genome.ucsc.edu.**
2. Chromosomal strand on which each gene is located is indicated as + or – in column 4.
3. SNP positions are shown only for those variations that are in promoter or coding regions. For example, rs1126499 in BGN is at the third position of codon 180, which results in no amino acid change (serine at 180 remains). In contrast, rs1676486 in COL11A1 is at the first position of codon 1496, resulting in replacement of serine with proline (S1496P).
4. The chromosomal position numbers for SNPs are not shown for the reason that they may change upon updating of sequences in the databases. Further, different databases may have different chromosomal position numbers for the same SNP, depending on annotations and archiving dates.

Table 2. SNPs and their genotypes that associate with OA or RA in the overall study population.

Disease	SNP ID	Gene	Genotype	Control No. (%)	Case No. (%)	Odds Ratio (95% CI)	Association	p
OA	rs2294984	COL9A3	**A/G** vs. G/G	96 (19.2)	127 (25.5)	1.36 (1.00 – 1.84)	Positive	**0.049**
	rs1800796	IL6	**C/G** vs. G/G	93 (18.6)	65 (13.1)	0.67 (0.47 – 0.96)	Negative	**0.03**
	rs1800629	TNFA	**A/G** vs. G/G	105 (21.0)	135 (27.1)	1.39 (1.03 – 1.87)	Positive	**0.03**
RA	rs10153674	CALM2	**A/A** vs. G/G	19 (3.8)	4 (0.9)	0.24 (0.08 – 0.73)	Negative	**0.01**
	rs2294984	COL9A3	**A/G** vs. G/G	96 (19.2)	108 (24.4)	1.43 (1.03 – 1.99)	Positive	**0.03**
	rs225014	DIO2	**G/A** vs. A/A	253 (50.7)	194 (43.9)	0.70 (0.52 – 0.93)	Negative	**0.02**
	rs2476601	PTPN22	**T/C** vs. C/C	67 (13.4)	91 (20.6)	1.59 (1.10 – 2.30)	Positive	**0.01**
	rs2488457	PTPN22	**C/C** vs. G/G	24 (4.8)	33 (7.5)	2.02 (1.13 – 3.59)	Positive	**0.02**
	rs2476601	PTPN22	**T/T** vs. C/C	4 (0.8)	12 (2.7)	3.77 (1.16 – 12.32)	Negative	**0.03**
	rs7574865	STAT4	**A/A** vs. C/C	25 (5.0)	44 (10.0)	2.46 (1.42 – 4.24)	Positive	**0.001**
	rs10181656	STAT4	**C/C** vs. G/G	25 (5.0)	45 (10.2)	2.58 (1.49 – 4.44)	Positive	**0.0007**
	rs11889341	STAT4	**A/A** vs. G/G	24 (4.8)	44 (10.0)	2.49 (1.44 – 4.31)	Positive	**0.001**
	rs8179673	STAT4	**G/G** vs A/A	25 (5.0)	45 (10.2)	2.58 (1.49 – 4.44)	Positive	**0.0007**

Notes. Table 2 shows only the SNPs that have significant associations ($p < 0.05$) with OA or RA in the overall study population. In the Genotype column, the first of the two genotypes shown for each SNP is the one that shows association when compared with the second one. In the Control and Case columns, the numbers and percentages shown are for the genotypes that associate with OA or RA. Based on $p < 0.05$ as an index of statistical association, the A/G of rs2294984 shows association with both OA and RA. However, the p-value of 0.049 for its association with OA is borderline, and with a larger cohort size, it may change either way.

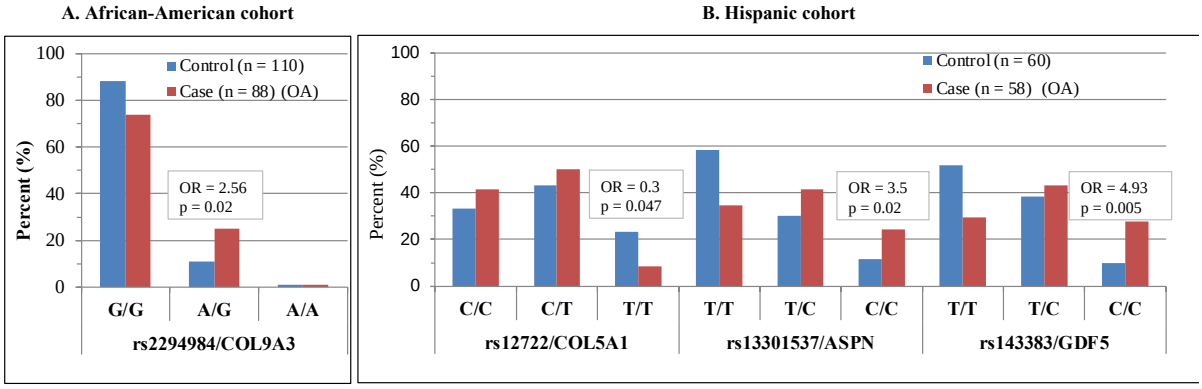


Figure 2. Frequency of occurrence and association with osteoarthritis (OA) of certain SNPs within the African-American (**A**) and Hispanic (**B**) cohorts. The odds ratio (OR) and p-values are shown for only those genotypes that have statistically significant association with OA.

In the African-American cohort, the A/G genotype of rs2294984 (COL9A3) showed 2.56-fold higher risk of OA in comparison to G/G (Figure 2). This SNP also associated with OA in the overall study population (Table 2) and the under-40 age group (not shown). In the Hispanic cohort, the T/T genotype of rs12722 (COL5A1) showed negative association with OA (OR =

0.3), i.e., 70 % less risk of OA in comparison to C/C, the most frequent genotype (Figure 2). In contrast, the C/C genotype of both rs13301537 (ASPN) and rs143383 (GDF5) showed positive association with OA in the Hispanic cohort, with odds ratios of 3.5 and 4.93, respectively (Figure 2). Further, the GDF5 SNP also showed association with OA in the White and African-American cohorts (not shown), as well as with RA in the Hispanic cohort (Figure 7). It should be noted that although these three SNPs showed association with OA in the Hispanic population (Figure 2), none of them showed association with OA or RA when the whole race and gender mixed population was considered (Table 2).

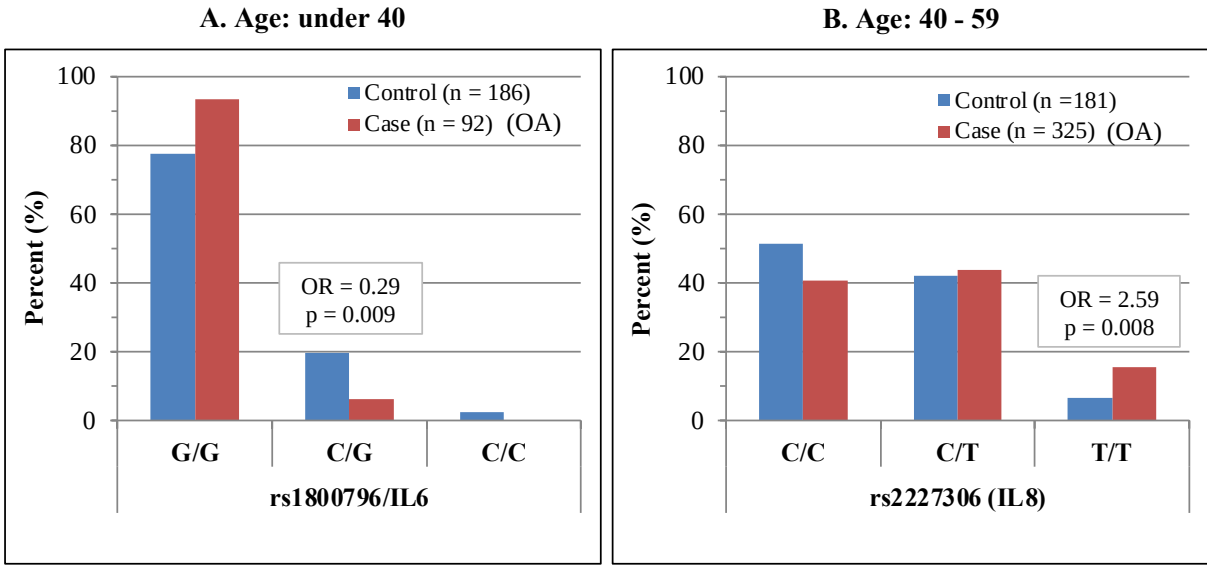


Figure 3. Two SNPs showed association with OA in two different age groups.

For the under-40 age group in the mixed study population, the C/G genotype of rs1800796 (IL6) showed 71 % less risk of OA in comparison to G/G (OR = 0.29). This SNP genotype also showed association in the whole study population, with p-value of 0.03 and OR of 0.67 (Table 2). For the 40 - 59 age group, T/T genotype of rs2227306 (IL8) showed positive association with OA, with OR of 2.59 in comparison to the homozygous C/C genotype, the most frequent in the control and OA subjects in this age bracket. However, in contrast to the IL6 SNP mentioned here (Figure 3), the IL8 SNP rs2227306 showed no association with OA or RA in the overall study population (Table 2).

The CALM2 SNP rs10153674 A/A genotype showed negative association with RA in the overall mixed population, the male cohort, and the 40 – 59 age group (Table 2; Figure 4). The odds ratios for these groups for the homozygous A/A genotype of this SNP were 0.24, 0.18, and 0.21, respectively, suggesting the genotype is protective against RA.

ASPN rs13301537 showed fundamentally different RA association results for two subgroups in the study population. Whereas its heterozygous genotype, T/C, negatively associated with RA

in the African-American cohort (OR = 0.19; 81 % less risk), the same genotype showed positive association with RA in the Hispanic cohort (OR = 4.05), suggesting nearly 4 times greater risk of RA when compared to the homozygous genotype, T/T.

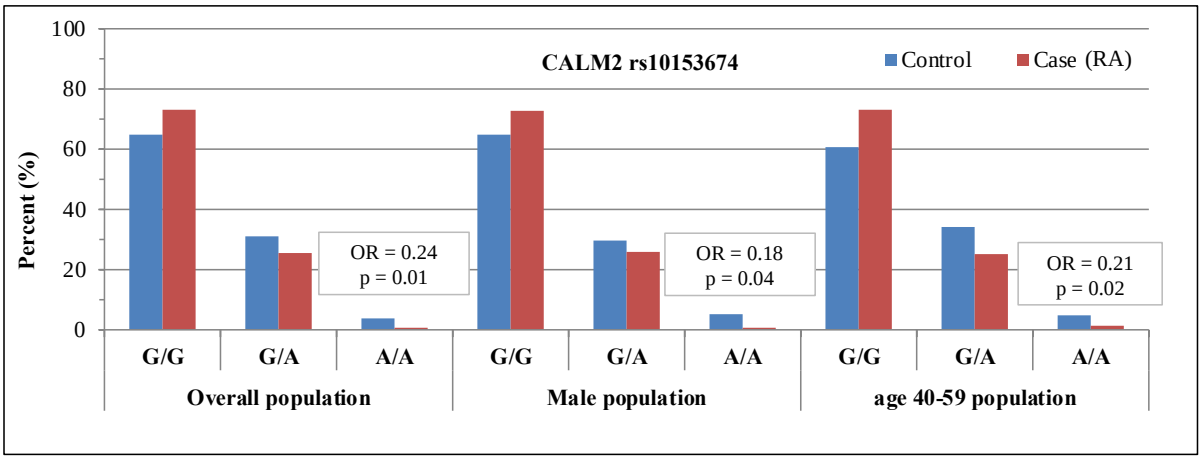


Figure 4. Negative association of CALM2 rs10153674 homozygous A/A genotype with rheumatoid arthritis (RA).

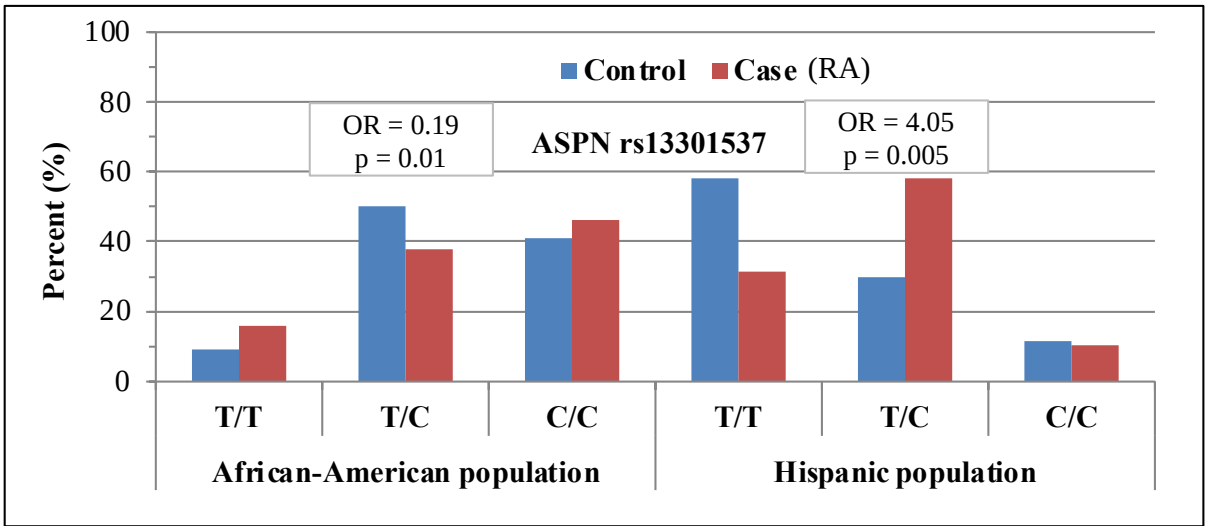


Figure 5. The SNP rs13301537 shows different associations with rheumatoid arthritis in the African-American and Hispanic cohorts within the study population.

The results in Figure 6 (A, B, C, D) show that the least frequent homozygous genotype of each of the four STAT4 SNPs show positive association with RA. The associations held for the overall study population, the Caucasian cohort, and the age under 40 cohort. The SNPs are rs10181656, rs11889341, rs7574865, and rs8179673, and their respective genotypes that associate with RA are C/C, A/A, A/A, and G/G. All four SNPs are located in proximity to one another in intron 3 of STAT4 gene, and are therefore likely haplotypes.

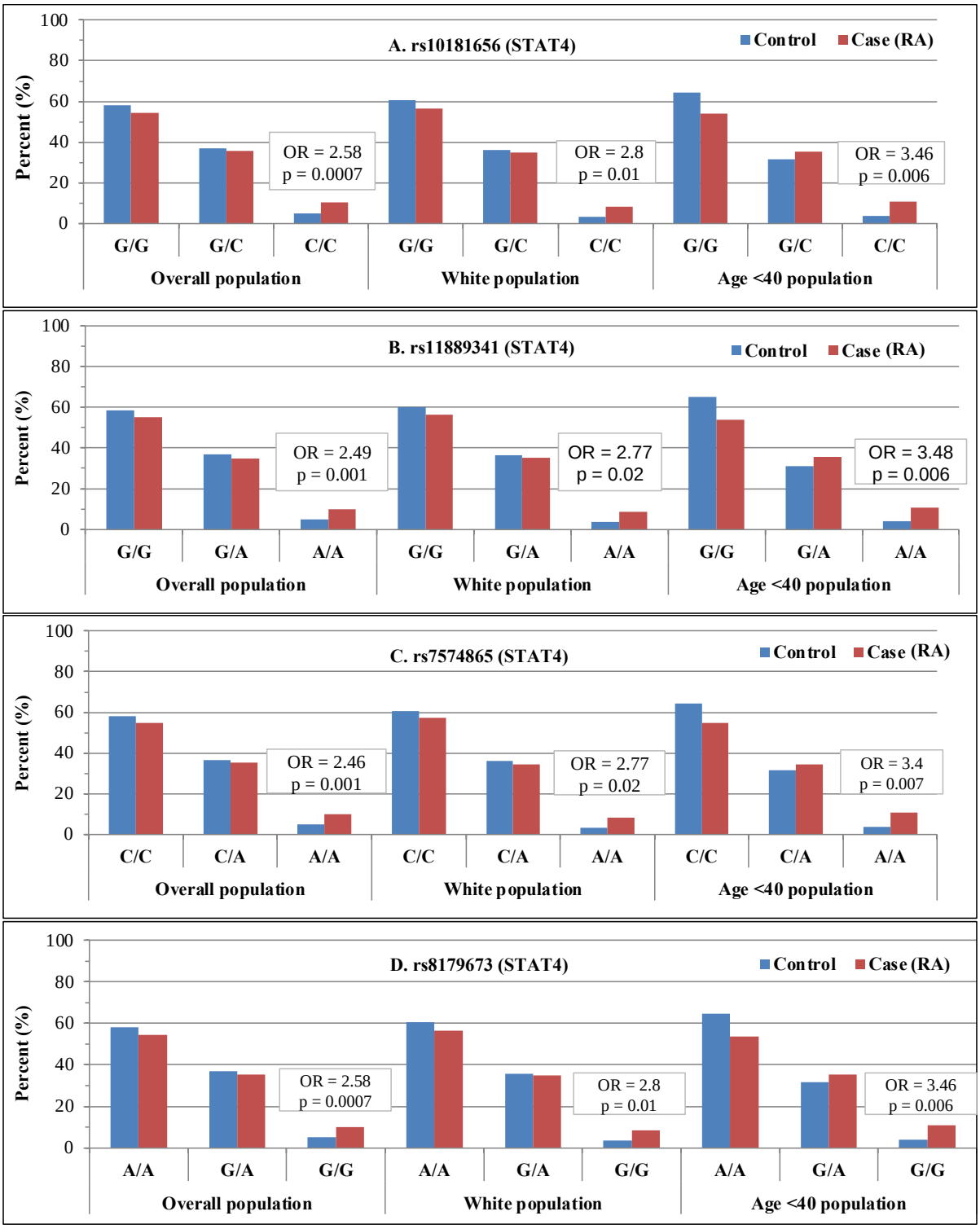


Figure 6 A – D. Four SNPs located in intron 3 of STAT4 associate with RA in the overall study population, White cohort, and <40 racially mixed age group.

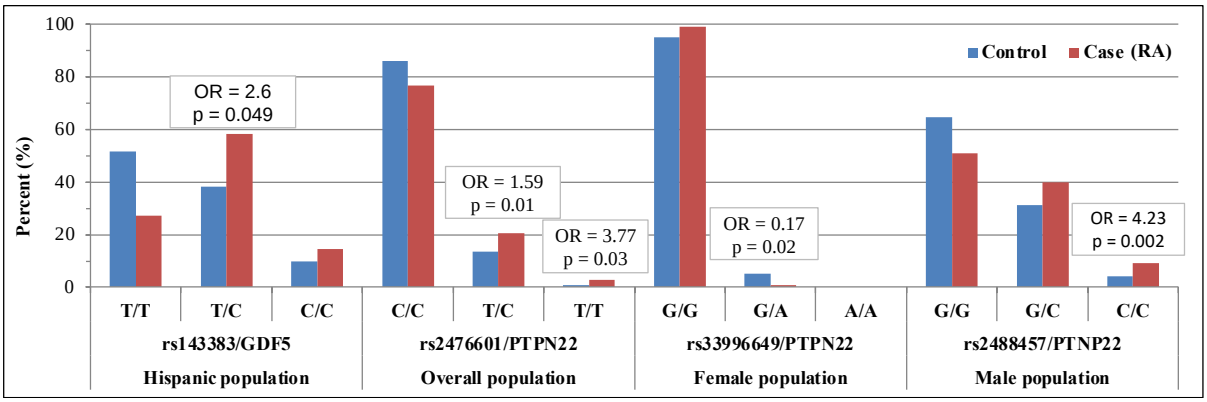


Figure 7. Four SNPs that show association with RA in the overall study population and the Hispanic, male, or female cohorts.

The single GDF5 SNP and three PTPN22 SNPs evaluated in this study showed statistically significant association with RA. The minor allele homozygous genotype, C/C, of GDF5 SNP (rs143383) associated with RA in the Hispanic cohort. The minor allele homozygous genotypes of the three PTPN22 SNPs T/T, A/A, and C/C associated with RA in the overall study population, the overall female cohort, and the overall male cohort, respectively (Figure 7).

SUMMARY AND DISCUSSION

In this gender and race mixed study, we separately evaluated the statistical association of 32 SNPs with osteoarthritis (OA) and rheumatoid arthritis (RA). The SNPs we analyzed are listed in Table 1, along with their genes, variations, and other relevant attributes. The objective of this study was not *de novo* discovery of any SNPs. Rather, the 32 SNPs we analyzed had been reported before, and are listed in various databases, e.g., the NCBI dbSNP and Variation Viewer. They have also been reported in various publications in terms of putative associations with OA, RA, or other musculoskeletal (MSK) disorders or injuries, e.g., ligament tears, and some with other diseases as well. For various SNPs, the reported associations have been positive, negative, or neutral for OA, RA, and various other MSK disorders.

Our overall findings have several main points: **1)** Only some of the 32 SNPs we analyzed show statistically significant associations with OA, RA, or both; most do not. **2)** The association patterns are quite varied; some SNPs show associations in the overall population, some only in one or another subgroup within the total study population, some associate with only OA, some with only RA, and some with both in different groups. **3)** We think such differential results are expected when different SNPs occurring in different genes with a wide spectrum of functions are analyzed in a study population that has subjects from different genetic lineages (race, ethnicity, gender). Some of the findings are further discussed below.

In the whole study population, 3 SNPs associate with OA and 10 with RA, and only one of these SNPs, rs2294984 A/G (COL9A3), associates with both OA and RA, albeit with a marginal p-value of 0.049 for OA (Table 2). The same SNP also shows association with OA in the African-American cohort, with strong odds ratio of 2.56 (Figure 2). However, it shows no association with OA in the Hispanic or White cohort. In contrast, the C/C genotype of both rs13301537 (ASPN) and rs143383 (GDF5) show strong positive association with OA in the Hispanic population, with odds ratios of 3.5 and 4.93, respectively.

In the Hispanic cohort, rs12722 T/T (COL5A1) shows negative association with OA (OR = 0.3), i.e., 70 % less risk of OA in comparison to C/C, the most frequent genotype (Figure 2). This SNP has been reported to show positive association with musculoskeletal injury severity, although not predisposition to injury (Massidda, 2015). COL5A1 rs61735045 polymorphism is in the coding region, and the variation results in Gly530Ser replacement in the protein. The variation has been associated with Ehlers-Danlos Syndrome and other MSK disorders (Giunta, 2000 & 2002). Our study did not reveal its association with OA or RA.

The age group association analysis revealed that rs1800796 C/G (IL6) associates with OA in the under-40 mixed cohort, with odds ratio of 0.29, i.e., 71% less risk of OA in comparison to the G/G (Figure 3). It has been reported that rs1800796 G/C genotype negatively associates with hip and knee OA in an elderly cohort (Fernandez, 2015). In contrast, rs1800796 shows positive

association with RA in an Egyptian cohort (Amr, 2016). For the 40 - 59 age group, the T/T genotype of rs2227306 (IL8) shows positive association with OA in comparison to its C/C genotype, the odds ratio being 2.59.

The CALM2 rs10153674 least frequent homozygous genotype, A/A, showed negative association with RA in the overall study population, the male cohort, and the 40 – 59 age group, with odds ratios of 0.24, 0.18, and 0.21, respectively (Figure 4). Although we found no association between this SNP and OA in our study, two SNPs in intron 2 of this gene have been reported to show statistically significant association with OA in a Japanese population study (Mototani, 2010).

As shown in Figure 5, the heterozygous genotype, T/C, of rs13301537 (ASPN) gave intriguing results; whereas it showed strong negative association with RA in the African-American cohort (OR = 0.19), the same genotype showed strong positive association with RA in the Hispanic cohort (OR = 4.05). Further, the least frequent genotype of this SNP in the former cohort is T/T, whereas in the latter cohort it is C/C.

The least frequent genotypes of all four STAT4 SNPs analyzed showed positive association with RA in the overall population, male cohort, and the below-40 age group (Figure 6 A, B, C, D). All four SNPs are located in proximity to one another in intron 3 of the gene, and it is therefore likely these SNPs are haplotypes. However, from this limited study, it is not feasible to reliably determine which of the four SNPs shows primary association with RA. It should be noted that STAT4 rs7574865 heterozygous genotype, C/A, has been reported to associate with reduced risk of RA in an Asian population (Jing, 2014). However, we found no association between this heterozygous genotype and OA or RA. Rather, our study revealed that the rs7574865 homozygous genotype, A/A, positively associates with RA in the whole population, White cohort, and the under-40 age group (Figure 6, C). These findings are not necessarily inconsistent, as Jing et al. (2014) studied an Asian cohort, whereas we had too small a number of the Asian subjects in our study to allow a reliable analysis, and therefore opted not to do it.

It was reported early on that COL1A1, which encodes collagen type I alpha-1, harbors a polymorphism, designated rs1800012, that predisposes to a number of different skeletal disorders and injuries, e.g., hip osteoarthritis, osteoporotic fracture, ligament tear, and tendon damage, apparently by resulting in low bone density (Mann, 2001; Lian, 2005; Khoschnau, 2008; Maffulli, 2013). However, our study did not find any association between this SNP and OA or RA.

MMP3, a matrix proteinase, primarily degrades non-collagen matrix proteins, and its activity has been associated with degradation of cartilage. In arthritis, it shows elevated expression (Burrage, 2006; Raleigh, 2009). We thought any of the MMP3 SNPs listed in Table 1 might associate with OA or RA. But we found no such association.

All three PTPN22 SNPs show association with RA in three different cohorts: rs2476601 and rs2488457 positively associate with RA in the overall population and male cohort, respectively, and rs33996649 associates negatively with RA in the female cohort (Figure 7).

A number of points regarding this study are noted below.

1. The sample sizes for each arthritic condition and the control cohort were limited. Therefore we cannot make any broad generalizations of results in terms of applicability to much larger sample sizes across a given or various mixed populations.

2. We note that statistically significant associations do not *per se* suggest cause and effect relationships. However, any genetic variation that reproducibly associates with a disease in a mixed population or a specific cohort, (e.g., gender, race) has the potential to be a diagnostic or predictive marker, regardless of whether it is a player in pathogenesis.

3. No mechanistic hints of how a given SNP may bring about any arthritic disorders can be discerned from studies of this nature and limited size. We think suggestions of possible mechanisms would be speculative, and therefore we have avoided doing that.

4. In as much as any defined genotype that causes a disease would not be a common occurrence, one would think that if a heterozygous genotype carrying the major and minor allele associates with a disease, then the minor allele homozygous genotype would do so even more strongly. However, we did not see this phenomenon uniformly. One example is COL9A3 rs2294984. The heterozygous genotype, A/G, of this SNP positively associates with RA and OA (Table 2), and with African-American cohort (Figure 2), but the homozygous minor allele genotype, A/A, does not. We think this is likely due to the vary small numbers that represent the A/A genotype (for example, see Figure 2), i.e., not enough to reliably reveal any association.

PART II

Transcriptome Expression Analysis of Osteoarthritis and Rheumatoid Arthritis Synovial Tissues

G. Jilani Chaudry[2,*], Tho Hua[1], Jamie L. Myers[1], Hui Xia[1], Manuel Y. Caballero[1], and Sandra Valtier[1]

[1]Center for Advanced Molecular Detection, [2]Science and Technology, 59 Medical Wing, JBSA-Lackland, TX

[*]**Corresponding Author:** G. Jilani Chaudry
ghulam.j.chaudry.ctr@mail.mil
210-292-7021 (office)
210-381-2894 (cell)

ABSTRACT

Osteoarthritis (OA) and rheumatoid arthritis (RA) occur markedly more frequently among the US military service members than in the general population, with the highest frequency being in the Army, followed by the Air Force, Marine Corps, and Navy. Genetic predisposition to both arthritic diseases is well-established. Thus, individual genetic constitution (genotypes, polymorphisms), gene expression, and protein expression profiles are important determinants of disease onset, severity, and prognosis. The current study assessed gene expression differences in synovial tissues obtained from OA and RA patients. The control samples were from subjects who had no arthritic disease. We analyzed whole transcriptome expression using RNA isolated from the synovial tissues to elucidate expression differences between the normal (control), OA, and RA cohorts. To do that, we purchased 117 commercially available fresh-frozen Caucasian synovial tissue samples. Of these, 26 were control samples (no musculoskeletal disease), 31 OA, and 60 RA. For the expression analysis, we used the Clariom D Human Array Chip and the GeneChip Scanner 3000 7G System. The expression data were analyzed with the Transcriptome Analysis Console (TAC) software. Expression levels of certain genes of interest were then further confirmed by more precise quantitative real-time PCR.

The combined Clariom D and TAC software revealed analysis of 135,750 transcripts. Overall, we found that 85 transcripts showed $\geq$ 5-fold expression in the OA samples in comparison to the control samples, while 191 transcripts showed $\leq$ 5-fold expression. For RA, 187 transcripts showed $\geq$ 5-fold expression, while 152 transcripts showed $\leq$ 5-fold expression. We note here that although the focus of this study was gene expression differences between normal and OA and RA samples, at least some of the genetic and gene expression differences relevant to these disorders likely also play crucial roles in injury susceptibility, as well as subsequent tissue repair. This was a limited pilot study. Similar studies with much greater number of cohort samples would need to be done to discern expression patterns that may signify predictive, prognostic, or pathologic value of expression patterns in OA or RA cohorts relative to the control cohort.

INTRODUCTION

Musculoskeletal diseases such as osteoarthritis (OA) and rheumatoid arthritis (RA) are prevalent in the general population as well as in the US military (CDC, 2014). But the frequency among the service members is markedly greater than in the general population, with the Army showing the highest occurrence, followed by Air Force, Marine Corps, and Navy (MSMR, 2014, 2019). Moreover, nearly 33 % of the veteran population suffers from these disorders. Musculoskeletal injuries and disorders are among the leading medical reasons for discharge from military service. Of these reasons for discharge, post-traumatic OA is the top reason, with 94.4 % of the cases attributed to combat injury. Among the active-duty personnel, the under-20 age group exhibits about 26 % higher OA rate than the corresponding group in the general population. Clearly, a research focus on arthritic diseases in the military, as well as MSK injuries, is of paramount importance.

Extrinsic factors such as physical activity, fitness, nutrition, and environmental factors play important roles in musculoskeletal injuries and disorders. But the most fundamental determinants of susceptibility or predisposition to MSK diseases, like many other diseases, are genetic. Part I of this study evaluated 32 SNPs for statistical associations with OA and RA in a gender and race mixed study population, as well as in various subgroups within the study cohort (race, gender, age group). This part (Part II) focuses on elucidating gene expression differences in synovial tissues from OA, RA, and control subjects. As stated earlier, the genetic variations exert their pathologic effects, if any, by altering the expression levels of genes, or disrupting the functions of gene products, i.e., RNAs and proteins. Thus, elucidating differential gene expression in disease and control samples is an indispensable approach to identifying expression patterns that associate with disease. Both RNA expression and protein expression studies have been reported on OA or RA, e.g., on synovial tissue samples, yielding valuable information on potential candidate biomarkers (Lorenz, 2003; Balakrishnan, 2014; Burska, 2014).

As for the SNP study (Part I), the samples for gene expression work reported in Part II here were purchased from a commercial source, but a different one. There are two important differences between the SNP study subjects and the synovial tissue gene expression study subjects: **One**, whereas the SNP study had comparatively much larger number of subjects in each category (OA, RA, control), the gene expression analysis study had much smaller numbers in the same categories. **Two**, whereas the SNP study had a race and gender mixed cohort, the gene expression study had all Caucasian samples, but was still gender mixed.

MATERIALS AND METHODS

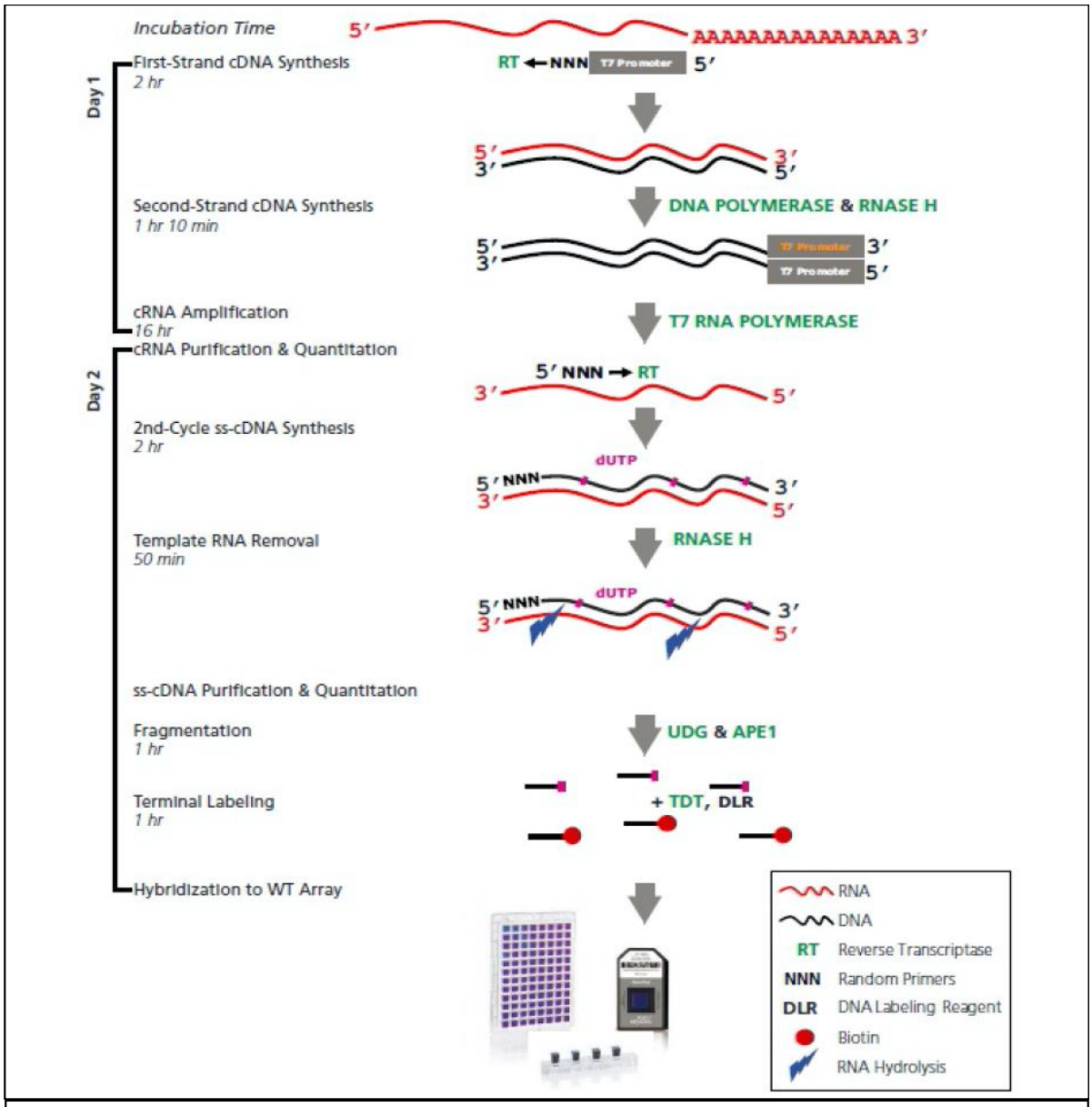
We purchased 117 commercially available fresh-frozen human synovial tissue samples of Caucasian origin (ProteoGenex, Inglewood, CA). Of these, 26 were control samples (no musculoskeletal disease associated; 10 female and 16 male), 31 OA samples (19 female and 12 male), and 60 RA samples (49 female and 11 male). Total RNA was isolated from the tissues by the TriZol method, as directed by the supplier (ThermoFisher Scientific, Waltham, MA). For the whole transcriptome expression analysis, we used the Clariom D Human Array Chip and the GeneChip Scanner 3000 7G System (Affymetrix, Santa Clara, CA). Each Clariom D array has the capacity to analyze about 540,000 transcripts.

Transcriptome Analysis Console (TAC) software was used to analyze the expression results and graphically visualize the data. The average transcript levels in the disease and control samples were compared to calculate the quantitative differences in expression. Student t-test was performed to determine age difference significance between the OA vs. control and RA vs. control groups.

For TaqMan real-time PCR validation of the microarray results, we selected 23 genes of interest that were either up-regulated or down-regulated in the OA and/or RA group as compared to the control group. A custom TaqMan Array Card (TAC) with our 23 selected gene targets was pre-loaded in a 384-well microfluidic card, with GAPDH as an expression normalization control.

RESULTS

This work aimed to elucidate differential gene expression profiles of synovial tissue samples from subjects with OA, RA, or no arthritis (control group). Total RNA isolated from samples was used to evaluate whole transcriptome expressions as described in Materials and Methods, and schematically shown in Figure 1 below.



Adapted from GeneChip™ Whole Transcript (WT) Expression Array User Guide

Figure 1. Schematic depiction of the method for whole transcriptome expression analysis.

Altogether, there were 117 samples: 26 control (no arthritis; 10 female and 16 male); 31 OA (19 female and 12 male), and 60 RA (49 female and 11 male). The samples were not race mixed; all were of Caucasian origin. The analyses were carried out as directed by the suppliers (see Materials and Methods; Figure 1).

The OA vs. Control transcriptome expression analysis revealed 2293 transcripts that showed >2-fold difference of expression in the OA cohort in comparison to the control cohort. Of these, 683 transcripts were up-regulated and 1610 down-regulated. For RA vs. Control, 2512 transcripts were differentially expressed, showing >2-fold difference, with 1169 up-regulated and 1343 down-regulated. The cut-off expression filter of >2-fold difference in expression excluded most of the transcripts, i.e., they had <2-fold difference in comparison to the control cohort. The proportions of various categories of transcripts revealed are presented in Figure 2 A and B.

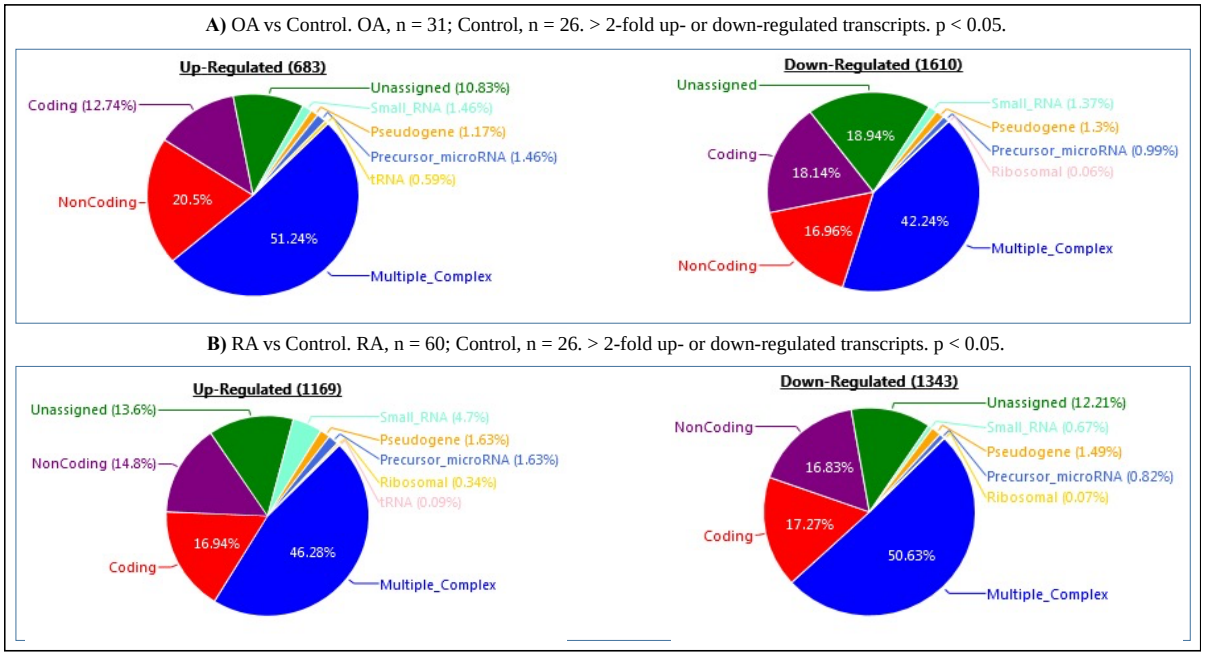


Figure 2. Categories and proportions of transcripts revealed by whole transcriptome analysis. The charts are based on transcripts that show > 2-fold up- or down-regulation, with $p < 0.05$. **A)** OA vs Control. OA, n = 31; Control, n = 26. **B)** RA vs Control. RA, n = 60; Control, n = 26.

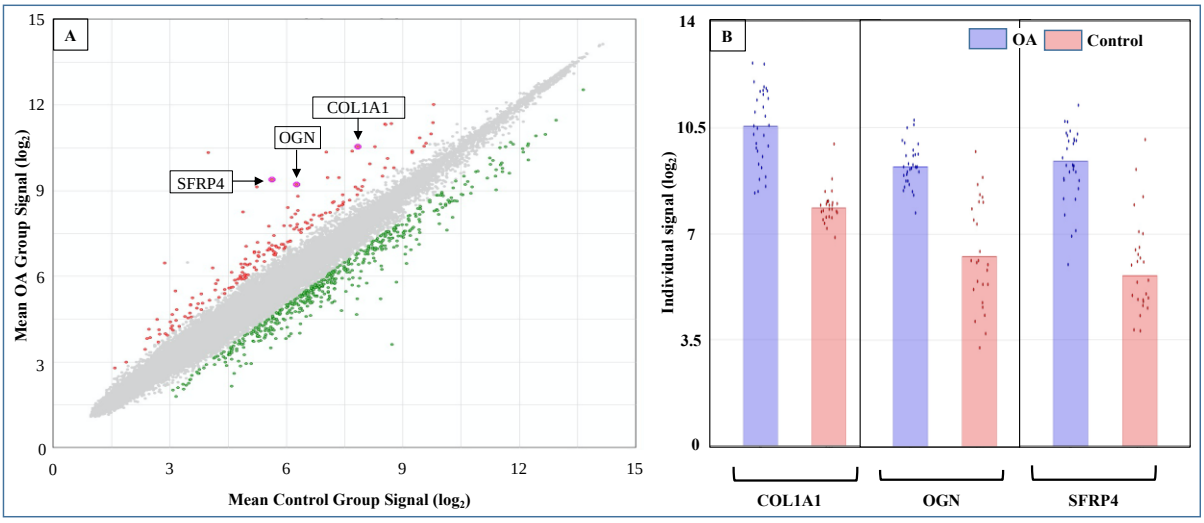


Figure 3. A) Scatter plot of differential transcriptome expression in OA vs. Control using Transcriptome Analysis Console (TAC) Software with RMA setting. **B)** Differential expression of three genes as examples; OGN (osteoglycin), COL1A1 (collagen, type I, alpha 1), and SFRP4 (secreted frizzled-related protein 4).

The scatter plot **(Figure 3, A)** shows transcripts with > 2-fold higher expression (red dots), those with > 2-fold lower expression (green dots), and those with < 2-fold expression (grey) differences in comparison to the control cohort. Each dot represents a unique transcript. For the three example genes **(Figure3, B)**, each dot represents an individual OA sample (blue bars) or control sample (red bars). The heights of bars represent the mean of expression signal in the respective categories.

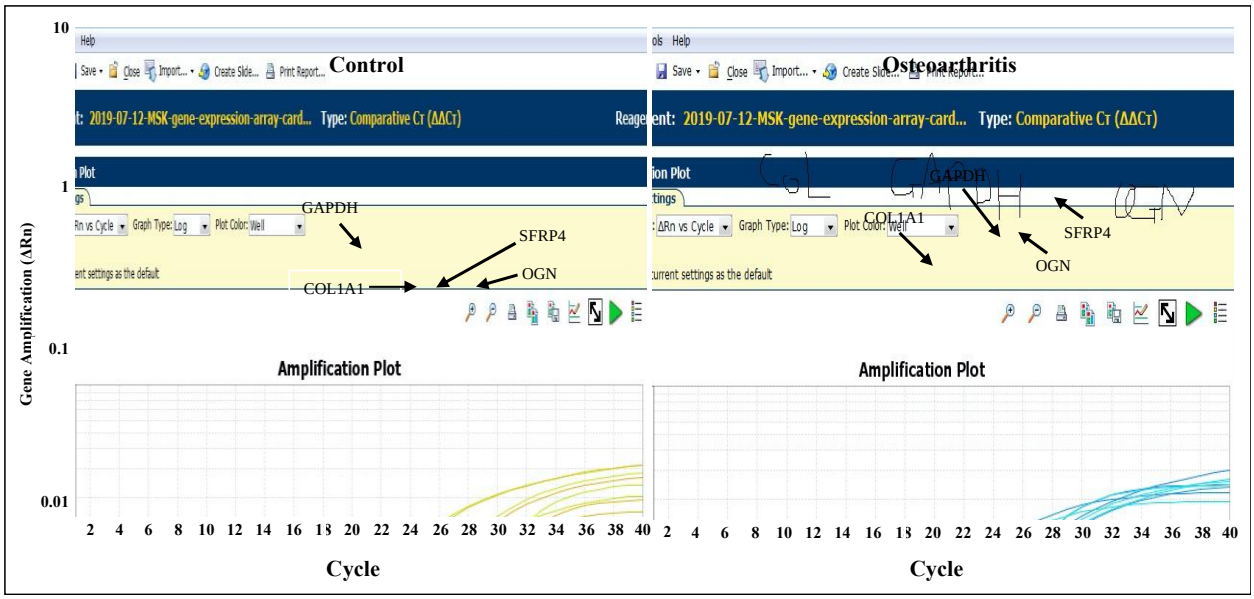


Figure 4. Graphic representation of quantitative real-time PCR amplification of three example genes and GAPDH gene in control and OA samples.

Table 1. Quantitative OA vs. control expression comparison of the 23 genes analyzed by using Clariom D Human gene chip and TaqMan RT-PCR Array method.

Gene	TaqMan Array Card	TAC 4.0 RMA	Gene	TaqMan Array Card	TAC 4.0 RMA
COL1A1	34.31	8.83	COL14A1	1.98	3.41
COL3A1	11.06	8.96	TGFBI	1.85	***
SFRP4	9.90	11.75	BMP4	1.20	***
MXRA5	9.18	23.31	GAPDH	1.00	***
COL1A2	9.18	5.64	CTGF	-1.23	***
SFRP2	6.45	5.04	MIR199A2	-1.25	4.66
OGN	5.00	8.64	ADIPOQ	-1.3	-5.7
ASPN	4.62	7.25	PLIN1	-1.56	-4.81
SCRG1	3.78	15.28	CFD	-2.70	-2.18
IGK	3.05	6.74	DDX3Y	-3.03	***
TGFB3	2.94	2.58	TGFBR3	-5.0	-4.42
COL5A1	2.56	3.85	PLIN2	-14.28	-6.94

TaqMan Array Card data are shown as fold differences in expression relative to GAPDH level in OA vs. Control. The TAC 4.0 RMA reflects GeneChip expression analysis results, which are generally meant to reveal transcripts that are up- or down-regulated. *** The difference did not meet a threshold parameter. Positive numbers signify higher expression levels and the negative numbers lower. The TaqMan Array Card data are generally considered more precise and reliable. The analysis shown here only includes 14 OA samples and 15 control samples. **COL1A1**, collagen, type I, alpha 1. **COL3A1**, collagen, type III, alpha 1. **SFRP4,** secreted frizzled-related protein. **MXRA5,** matrix remodeling associated 5. **COL1A2**, collagen, type I, alpha 2. **SFRP2**, secreted frizzled-related protein 2. **OGN,** Osteoglycin. **ASPN**, asporin. **SCRG1**, stimulator of chondrogenesis 1. **IGK**, immunoglobulin kappa locus. **TGFB3**, transforming growth factor beta 3. **COL5A1**, collagen, type V, alpha 1. **COL14A1**, collagen, type XIV, alpha 1. **TGFBI**, transforming growth factor, beta-induced. **BMP4**, bone morphogenetic protein 4. **CTGF**, connective tissue growth factor. **MIR199A2**, microRNA 199a-2. **ADIPOQ**, adiponectin, C1Q, and collagen domain containing. **PLIN1**, perilipin 1. **CFD**, complement factor D (adipsin). **DDX3Y**, DEAD (Asp-Glu-Ala-Asp) box helicase 3, Y-linked. **TGFBR3**, transforming growth factor beta receptor III. **PLIN2**, perilipin 2.

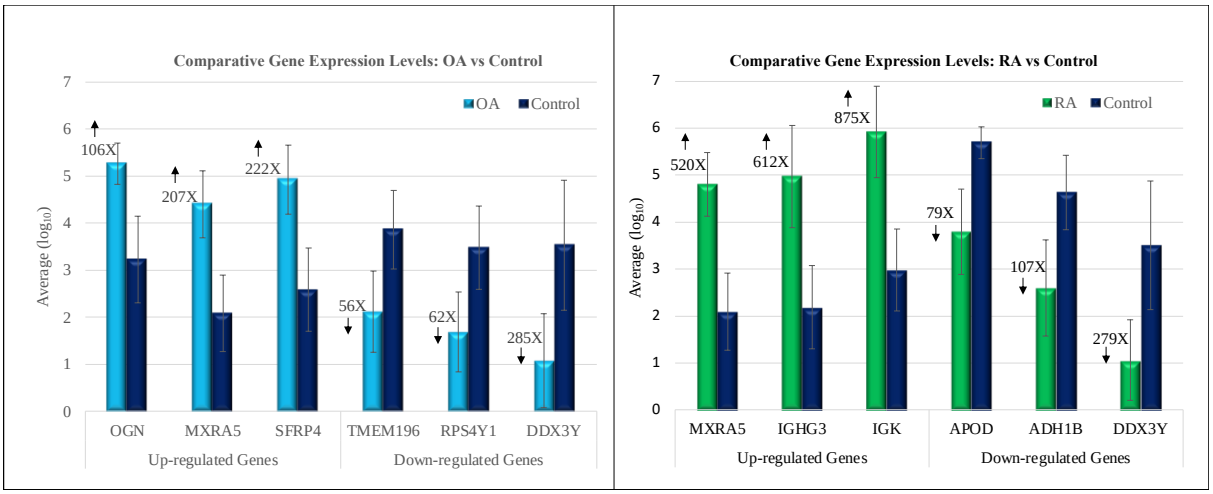


Figure 5. Top 3 up- and top 3 down-regulated transcripts in the OA (left panel) and RA (right panel) samples relative to the control samples.

The gene expression levels are noted as fold differences relative to the control levels. The up and down arrows indicate up- and down-regulation, respectively. **OGN,** Osteoglycin. **MXRA5,** matrix remodeling associated 5. **SFRP4,** secreted frizzled-related protein 4. **TMEM196,** transmembrane protein 196. **RPS4Y1,** ribosomal protein S4 Y-linked 1. **DDX3Y,** DEAD-box helicase 3 Y-linked. **IGHG3,** immunoglobulin heavy constant gamma 3. **IGK,** immunoglobulin kappa locus. **APOD,** apolipoprotein D. **ADH1B,** alcohol dehydrogenase 1B, beta polypeptide.

SUMMARY AND DISCUSSION

This study assessed whole transcriptome expression profile of 117 Caucasian synovial tissue samples representing OA (n = 31), RA (n = 60), and control (n = 26; no arthritis). The analysis used total RNA isolated from the samples. There were two levels of analysis. In the first, we used the gene chip method for whole transcriptome expression in each sample. For this approach we used the Clariom D Human Array Chip, which has the capacity to analyze about 540,000 transcripts. The chip analysis reported a total of 135,750 transcripts analyzed. Of these, 683 showed > 2-fold higher and 1,610 > 2-fold lower expression in the OA population relative to the control population ($p < 0.05$). The corresponding values for RA were 1,169 and 1,343 ($p < 0.05$).

In the second approach we analyzed a select number of genes (Table 1) by TaqMan RT-PCR. This work involved 15 control samples and 14 OA samples (Figure 5). The data correlate reasonably well with the gene chip analysis data. Mainly, the transcripts that show up- or down-regulation in the gene chip data also show up- or down-regulation in the RT-PCR data. The rest of the samples were not analyzed.

A previous study employed a similar approach to assess gene expression in synovial tissue cells and any associations with RA (Galligan, 2007). That study identified a number of genes differentially expressed in RA. We note, however, that the authors used the Affymetrix gene chip U133A, which has a capacity of 14,500 transcripts, and they also propagated the synovial tissue fibroblasts *in vitro* before subjecting them gene expression analysis. The study reported a number of genes differentially expressed in RA. Surprisingly, the major genes they listed have no overlap with the major up- or down-regulated genes we describe here.

We underscore the fact that this was a limited pilot study. To discern gene expression patterns for broad generalizations in terms of applicability, such studies must have much larger numbers of samples.

REFERENCES

1. Armed Forces Health Surveillance Center. Brief report: Morbidity burden to attributable to illnesses and injuries in deployed (per Theater Medical Data Store TMDS) compared to nondeployed (per Defense Medical Surveillance System DMSS) settings, active component, U.S. Armed Forces. MSMR, 26 (5): 2–51. 2019.

2. Amr K, El-Awady R, and Raslan H. Assessment of the −174G/C (rs1800795) and −572G/C (rs1800796) Interleukin 6 Gene Polymorphisms in Egyptian Patients with Rheumatoid Arthritis. Maced. J. Med. Sci., 4 (4):5 74-577. 2016.

3. Balakrishnan L, Nirujogi RS, Ahmad S, et al. Proteomic analysis of human osteoarthritis synovial fluid. Clinical Proteomics, 11, article 6. 2014.

4. Burrage PS, Mix KS, and Brinckerhoff CE. Matrix Metalloproteinases: Role In Arthritis. Front. Biosci., 11; 529-543. 2006.

5. Burska AN, Roget K, Blits M, et al. Gene expression analysis in RA: towards personalized medicine. Pharmacogenomics J., 14 (2): 93-106. 2014.

6. Evangelou E, Kerkhof HJ, Styrkarsdottir U, et al. A meta-analysis of genome-wide association studies identifies novel variants associated with osteoarthritis of the hip. Ann. Rheum. Dis., 73 (12): 2130-2136. 2014.

7. Fernandes MTP, Fernandes KBP, Marquez AS, et al. Association of interleukin-6 gene polymorphism (rs1800796) with severity and functional status of osteoarthritis in elderly individuals. Cytokine, 75; 316–320. 2015.

8. Galligan CL, Baig E, Bykerk V, et al. Distinctive gene expression signatures in rheumatoid arthritis synovial tissue fibroblast cells: correlates with disease activity. Genes and Immunity, 8; 480–491. 2007.

9. Giunta C, Nuytinck L, Raghunath M, et al. Homozygous Gly530Ser substitution in COL5A1 causes mild classical Ehlers-Danlos syndrome. Am. J. Med. Genet., 109 (4): 284-290. 2002.

10. Giunta C and Steinmann B. Compound heterozygosity for a disease-causing G1489E (corrected) and disease-modifying G530S substitution in COL5A1 of a patient with the classical type of Ehlers-Danlos syndrome: an explanation of intrafamilial variability? Am. J. Med. Genet., 90 (1): 72-79. 2000.

11. Hauret KG, Jones BH, Bullock SH, et al. Musculoskeletal injuries: description of an under-recognized injury problem among military personnel. Am. J. Prev. Med., 38 (1 Suppl): S61-70. 2010.

12. https://www.cdc.gov/arthritis/data_statistics/national-statistics.html.

13. https://www.ncbi.nlm.nih.gov/snp/.

14. https://www.ncbi.nlm.nih.gov/variation/view/.

15. http://useast.ensembl.org/Homo_sapiens/Info/Index?db=core.

16. https://genome.ucsc.edu.

17. Jiang X, Zhou Z, Zhang Y, et al. An updated meta-analysis of the signal transducer and activator of transcription 4 (STAT4) rs7574865 G/T polymorphism and rheumatoid arthritis risk in an Asian population. Scand. J. Rheumatology, 43 (6); 477-480. 2014.

18. Jiang Y, Zhang R, Zheng J, et al. Meta-analysis of 125 rheumatoid arthritis-related single nucleotide polymorphisms studied in the past two decades. PLoS One, 7 (12); e51571. 2012.

19. Jones, BH, Canham-Chervak, M, Canada, S, et al. Medical surveillance of injuries in the U.S. Military descriptive epidemiology and recommendations for improvement. Am. J. Prev. Med., 38 (Suppl. 1): S42–S60, 2010.

20. Khoschnau S, Melhus H, Jacobson A, et al. Type I collagen alpha1 Sp1 polymorphism and the risk of cruciate ligament ruptures or shoulder dislocations. Am. J. Sports. 36: 2432-2436. 2008.

21. Lian K, Zmuda JM, Nevitt MC, et al. Type I collagen alpha1 Sp1 transcription factor binding site polymorphism is associated with reduced risk of hip osteoarthritis defined by severe joint space narrowing in elderly women. Arthritis Rheum., 52: 1431-1436. 2005.

22. Lorenz P, Ruschpler P, Koczan D, et al. From transcriptome to proteome: Differentially expressed proteins identified in synovial tissue of patients suffering from rheumatoid arthritis and osteoarthritis by an initial screen with a panel of 791 antibodies. Proteomics, 3; 991-1002. 2003.

23. Maffulli N, Margiotti K, Longo UG, et al. The genetics of sports injuries and athletic performance. Muscles Ligaments Tendons J., 3 (3): 173-189. 2013.

24. Mann V, Hobson EE, Li B, et al. A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality. J. Clin. Invest., 107: 899-907. 2001.

25. Massidda M, Bachis V, Corrias L, et al. Influence of the COL5A1 rs12722 on musculoskeletal injuries in professional soccer players. J. Sports Med. Phys., 55 (11): 1348-1353. 2015.

26. Mototani H, Iida A, Nakamura Y, and Ikegawa S. Identification of sequence polymorphisms in CALM2 and analysis of association with hip osteoarthritis in a Japanese population. J. Bone and Min. Metab., 28; 547–553. 2010.

27. Musumeci G, Aiello FC, Szychlinska MA, et al. Osteoarthritis in the XXIst century: risk factors and behaviours that influence disease onset and progression. Int. J. Mol. Sci.,16 (3): 6093-6112. 2015.

28. Ombrello MJ, Sikora KA, and Kastner DL. Genetics, genomics, and their relevance to pathology and therapy. Best Pract. Res. Clin. Rheumatol., 28 (2): 175-189. 2014.

29. Raleigh SM, van der Merwe L, Ribbans WJ, et al. Variants within the MMP3 gene are associated with Achilles' tendinopathy: possible interaction with the COL5A1 gene. Br. J. Sports Med., 43 (7): 514-20. 2009.

30. Tilkeridis C BT, Garantziotis S, and Stratakis CA. Association of a COL1A1 polymorphism with lumbar disc disease in young military recruits. J. Med. Genet., 42: e44. 2005.

31. Wang T, Liang Y, Li H, et al. Single nucleotide polymorphisms and osteoarthritis: An overview and a meta-analysis. Medicine, 95 (7); e2811, 1-13. 2016.

32. Yarwood A, Huizinga TWJ, and Worthington J. The genetics of rheumatoid arthritis: risk and protection in different stages of the evolution of RA. Rheumatology, 53; 1-11. 2014.

SYMBOLS, ABBREVIATIONS, AND ACRONYMS

CI	Confidence Interval
DNA	Deoxyribonucleic Acid
MSK	Musculoskeletal
PCR	Polymerase Chain Reaction
OA	Osteoarthritis
OR	Odds Ratio
RA	Rheumatoid Arthritis
RNA	Ribonucleic Acid
RT-PCR	Real-Time Polymerase Chain Reaction
SNP	Single Nucleotide Polymorphism
TAC	TaqMan Array Card. Transcriptome Analysis Console.
UTR	Untranslated Region