

國立臺灣大學公共衛生學院
流行病學與預防醫學研究所



碩士論文

Graduate Institute of Epidemiology and Preventive Medicine

College of Public Health

National Taiwan University

Master Thesis

臺灣類風濕性關節炎全基因體關聯性研究

Genome-Wide Association Study of Rheumatoid

Arthritis in the Taiwanese population

林芳宇

Fang-Yu Lin

指導教授：盧子彬 博士

Advisor: Tzu-Pin Lu, PhD

中華民國 109 年 7 月

July, 2020

致謝

能順利在兩年完成碩士班學業，要特別感謝我的指導老師—盧子彬老師，從入學開始時常關心我的學校生活與課程內容，在基因研究領域也帶領我從陌生到最後能完成論文，討論時間的安排、遺傳統計方法與分析、研究方向與口試準備，老師都給予極大的幫助，使我在研究過程中更能得心應手。另外，要十分感謝奇祐，在研究過程遇到的統計相關困難，不厭其煩與我討論並提供協助，才能順利的完成此篇論文。感謝所上施惟量老師不只教導選修課程內容，也時常關心我的研究進度，幫我加油打氣。在職班同學平常比較少有機會一起在學校，感謝同學豐仁在研究路上常互相鼓勵，分享交流學校消息避免遺漏須完成事項。謝謝我的妹妹倩宇，在研究領域方面是我的前輩，即使人在日本留學中，在我對研究內容困惑或煩惱的時候，總是義不容辭透過通訊軟體給予我實質上的協助與意見，提供我很多學業上的經驗分享，扮演督促我的角色及成為我心靈上的支柱。最後，感謝我的爸媽、兄弟姊妹在我求學過程中的全力支持，讓我能無後顧之憂學習。

林芳宇 謹致

中文摘要



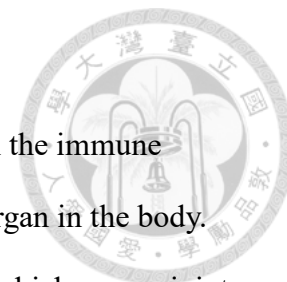
背景：類風濕性關節炎是一種自體免疫疾病，主要特徵是慢性發炎性滑膜炎，而導致關節疼痛、僵硬、腫脹或扭曲，嚴重者可至失能，患者預期生存期比一般人口少了 17 年。類風濕性關節炎與遺傳基因有關，過去研究已經確定了超過 100 個風險位點；然而，雖然該風險位點存在著明顯的族群差異，過去多數相關研究僅以歐洲族群為分析對象。本研究目的為 1) 找出臺灣族群的類風濕性關節炎風險位點；2) 將找出的臺灣族群類風濕性關節炎風險位點，以既有資料庫中的東亞族群與歐洲族群，比較兩者之間的風險位點差異。

方法：本研究樣本取自臺灣人體生物資料庫，共 137 位患者病例組，及對照組 15,785 人，控制類風濕性關節炎的危險因子(年齡、性別、有無類風濕性關節炎家族病史)影響後，使用羅吉斯迴歸模型估計與類風濕性關節炎有關的風險位點。將找出的臺灣族群類風濕性關節炎風險位點，以既有資料庫中的東亞族群與歐洲族群，進行信賴區間估計來找出兩者之間的風險位點差異。

結果：本研究共有 604,202 個單核苷酸多型性位點納入分析，共 36 個單核苷酸多型性位點與類風濕性關節炎顯著相關($p < 0.0001$)。比對東亞族群與歐洲族群的風險位點資訊後，多數風險位點顯示在兩個族群中有明顯差異，除了其中 5 個風險位點。

結論：本研究結果在臺灣族群中發現特定的類風濕性關節炎風險位點，且這些位點多數有明顯的東亞與歐洲族群的差異。我們的結果顯示針對本土樣本探討特有的類風濕性關節炎風險位點是有必要性的，該研究發現將能夠有效應用於當地公共衛生與臨床醫療的發展及治療。

關鍵詞：類風濕性關節炎、全基因體關聯性研究、臺灣



Introduction: Rheumatoid arthritis is an autoimmune disease, with the immune system losing its normal function and attacking normal tissues or organ in the body. The main feature of the disease is chronic inflammatory synovitis, which causes joint pain, stiffness, swelling or even disability. The patients' expected survival time is 17 years shorter than the general population. It is known that most of rheumatoid arthritis is related to genetics and more than 100 risk sites have been identified. However, most of the known risk single nucleotide polymorphisms (SNPs) were identified in the European populations even the difference in risk SNPs between different ethnic groups were emphasized. The purposes of the study were i) to identify the risk SNPs of rheumatoid arthritis in the Taiwanese population; ii) to compare the differences of these risk SNPs between East Asian and European populations.

Methods: A total of 137 rheumatoid arthritis cases and a control group of 15,785 were extracted from the Taiwan Biobank. A logistic regression model was used to evaluate the effect of risk SNPs after controlling for the covariates of rheumatoid arthritis including age, sex, and family history of rheumatoid arthritis. Besides, interval estimates were used to evaluate the difference in risk SNPs between the East Asian and the European populations using the GWAS Catalog database.

Results: A total of 604,202 SNPs were included in the analysis, and 36 SNPs were significantly associated with rheumatoid arthritis ($p < 0.0001$). After comparing the estimates of Asian and the European populations, there were marked differences between the two populations in most of the risk SNPs, with the exception of five out of them.

Conclusions: Our findings identified some specific risk SNPs of rheumatoid arthritis in the Taiwanese population and supported it is essential to identify the risk SNPs of

rheumatoid arthritis using local samples. Our findings could provide new insights into the development of rheumatoid arthritis in public health and clinical medicine.



Keywords: rheumatoid arthritis, GWAS, Taiwan

目錄



致謝.....	II
中文摘要.....	III
英文摘要.....	IV
第一章 前言.....	10
第一節 研究背景.....	10
第二節 研究動機及重要性.....	10
第三節 研究目的.....	11
第二章 文獻回顧.....	12
第一節 類風濕性關節炎.....	12
第二節 類風濕性關節炎族群間差異.....	12
第三節 類風濕性關節炎全基因體關聯性研究.....	13
第三章 方法及材料.....	15
第一節 研究資料.....	15
第二節 品質控管.....	16
第三節 統計分析.....	17
第四章 結果.....	18
第一節 人口學變項分布.....	18
第二節 品質控管.....	18
第三節 全基因體關聯性分析.....	18
第四節 東亞與歐洲族群的風險位點比較.....	19
第五節 驗證過去研究.....	19
第五章 討論.....	21
第一節 主要發現.....	21
第二節 研究限制.....	22

第三節 公共衛生與臨床意義.....	22
參考文獻.....	33
附錄一、2010 年 ACR/EULAR 聯合發表類風濕性關節炎診斷分類標準.....	39



圖目錄

圖一、品質控管流程圖.....	24
圖二、曼哈頓圖.....	25
圖三、此研究不同 p 值篩選出的顯著位點量.....	26



表目錄

表一、人口學變項分布.....	27
表二、36 個顯著單核苷酸多型性位點.....	28
表三、東亞與歐洲族群信賴區間估計(36 個顯著單核苷酸多型性位點).....	29
表四、顯著風險位點與性狀關聯.....	30
表五、59 個未顯著但在 GWAS Catalog 的類風濕性關節炎風險位點.....	31

第一章 前言

第一節 研究背景

類風濕性關節炎在自體免疫疾病中有其特殊地位，過去多與免疫、生化學相關的研究，近年則開始整合新的生物學、遺傳學、表觀遺傳學等研究方法(Jutley, Raza, & Buckley, 2015)，是最常見的自體免疫關節炎，人體免疫系統失去應有的作用，無法控制發炎反應，導致關節疼痛、僵硬、腫脹並降低關節的柔韌性，其致病原因尚不明確(American College of Rheumatology, 2019)。除了常見的滑膜關節發炎外，其他全身性包括類風濕性結節、間質性肺疾病、角膜鞏膜炎和各種神經系統疾病的表現通常也很嚴重(Kuo et al., 2013)。類風濕性關節炎患者的生活質量(quality of life)比第二型糖尿病、心肌梗塞、高血壓患者低，且與漸進性失能、全身性併發症、提早死亡及社會經濟成本增加有關(McInnes & Schett, 2011; "Rheumatoid arthritis," 2018)。在疾病發作後的 2-3 年內，20%-35% 的患者因工作能力障礙而離職，在 5-10 年內該百分比可高達 40%，這為患者及家庭、雇主和政府帶來了沉重的負擔(McBurney & Vina, 2012)。

第二節 研究動機及重要性

過去研究指出類風濕性關節炎的盛行率以及活動表徵在不同族群間有所差異，舉例而言，拉丁裔患者的疾病活動指數(disease activity level)較高，非洲裔美國人臨床緩解率較低，兩種族的功能狀態皆比白人差(Greenberg et al., 2013)；而女性及老年人是高風險族群，尤常見於 60-70 歲(Scott, Wolfe, & Huizinga, 2010)，發生率隨年紀增加而增加，其中又以女性發生率高於男性 2-3 倍(Ndosi, Vinall, Hale, Bird, & Hill, 2011)。

根據臺灣的流行病學與死亡率研究指出，女性的類風濕性關節炎發生率與盛行率比男性高，而男性類風濕性關節炎死亡率比女性高。與一般人口相比，類風濕性關節炎患者的死亡率較高，因有較高機會發生合併症，如心血管疾病

(McInnes & Schett, 2011; Scott et al., 2010)、骨質疏鬆性骨折、惡性腫瘤、感染(Kuo et al., 2013)，增加患者死亡率(Jin et al., 2017)，預期生存期比一般人口少了 17 年(Wagan et al., 2016)。

目前，類風濕性關節炎與基因的人口研究仍多集中在西方國家，近年少數研究針對日本、韓國東亞族群發表其類風濕性關節炎風險基因，臺灣亦是東亞族群的一份子但卻尚未有類風濕性關節炎相關基因訊息的研究，故本文旨在以臺灣人口為主要研究對象，找出在臺灣的類風濕性關節炎風險基因，提供亞洲人群的遺傳和表觀遺傳學訊息，為未來的風濕病學研究有所貢獻。

第三節 研究目的

本研究的目的主要有兩項：

- 一、使用臺灣人體生物資料庫比較有無罹患類風濕性關節炎患者的單核苷酸多型性位點(single nucleotide polymorphisms, SNPs)，並從中找出臺灣族群的類風濕性關節炎風險位點；
- 二、將目的一中找出的臺灣族群類風濕性關節炎風險位點，使用資料庫(GWAS Catalog database)中既有的東亞族群與歐洲族群位點資訊，估計其信賴區間後比較族群間是否有顯著差異。

第二章 文獻回顧

第一節 類風濕性關節炎

根據美國風濕病學會(American College of Rheumatology, ACR)與歐洲抗風濕病聯盟(the European League Against Rheumatism, EULAR)於 2010 年聯合發表的「類風濕性關節炎診斷分類標準」，建議針對以下兩種目標族群：1)至少有一個關節患有明確的臨床滑膜炎(腫脹)；2)患者的滑膜炎以其他疾病無法有更好的解釋。再以關節侵犯程度、血清學指標、急性發炎指標、症狀持續時間共四大類累計分數，總分十分，六分以上即可診斷為類風濕性關節炎(附錄一)。

類風濕性關節炎 50-60%與遺傳基因有關(Jalil et al., 2016)，Yarwood 等人甚至指出至少有 2/3 的風險因子跟基因有關。任何年齡都可能發生，但 30 至 50 歲之間是高峰(A. Wasserman, 2018)，女性、有該病家族史為風險因子(Kuo et al., 2017)。如果至少有一個關節明顯腫脹，而又不能用另一種疾病更好地解釋，則應考慮類風濕性關節炎。患有關節炎的患者中，C-反應蛋白(C-reactive protein, CRP)升高、監測全血細胞計數(complete blood count, CBC)，包含白血球計數(White blood cell count, WBC count)，為評估及診斷罹患類風濕性關節炎的重要指標(A. M. Wasserman, 2011)。

第二節 類風濕性關節炎族群間差異

類風濕性關節炎的年發生率和盛行率在不同族群與國家間均有很大差異(Ndosi et al., 2011; Wu et al., 2019)，高盛行率 5.3%的皮馬族(Pima)與 6.8%的齊佩瓦族(Chippewa Native Americans) (Minichiello, Semerano, & Boissier, 2016)，與低盛行率 0.4%的中國與 1.0%的日本(Jin et al., 2017; Yamanaka, Sugiyama, Inoue, Taniguchi, & Momohara, 2014)。而在美國、英國、芬蘭、丹麥和瑞典的年發生率為 0.9%-1.1%；在其他歐洲國家中為 0.4%-0.6%；南美國家/地區的 0.1%-0.5%；亞洲國家為 0.1%-0.3%；非洲為 0%-0.3%。在臺灣，根據 1994 年的一項調查，

農村地區類風濕性關節炎盛行率為 0.3%，城市的盛行率為 0.9%，然而尚無對類風濕性關節炎發生率的估計(Kuo et al., 2013)。

另外，類風濕性關節炎的臨床疾病活動度指標(Clinical Disease Activity Index, CDAI)也存在族群間差異(American College of Rheumatology, 2019)。臨床疾病活動度是評估關節症狀與生活功能的指標，是由四個項目總計分數而成：28 處關節觸痛的數目(TJC28)、28 處關節腫脹的數目(SJC28)、病人自我評估(PaGH)與醫師評估(PrGH)，總分為 0-76，0.0-2.8 為緩解(Remission)，2.9-10.0 為低疾病活動度(Low Activity)，10.1-22.0 為中疾病活動度(Moderate Activity)，22.1-76.0 為高疾病活動度(High Activity)。


這些數據表示類風濕性關節炎的風險因子除了與遺傳有關，在族群之間也具有差異性。

第三節 類風濕性關節炎全基因體關聯性研究

全基因體關聯性研究(genome-wide association study, GWAS)透過分析基因型(通常為單核苷酸多型性位點)與表現型的關聯，來了解複雜性疾病與性狀的變異(Ball, 2013)，目前已成功鑑定出數千種與疾病相關的單核苷酸多型性位點與性狀，有助於了解疾病的致病機轉、尋找疾病相關的生物標記(biomarker)，為疾病診斷與治療提供新的方針。

透過全基因體關聯性統合分析，過去許多研究都曾提出與類風濕性關節炎相關的位點與基因，舉例來說，Stahl et al. (2010)提出與類風濕性關節炎相關的位點(rs934734、rs6859219、rs2623、rs13315591、rs874040、rs3093023、rs10488631)；Kochi, Suzuki, and Yamamoto (2014)則提出類風濕性關節炎的風險基因(PADI4、PTPN22、TNFAIP3、STAT4、CCR6)；Yamamoto, Okada, Suzuki, and Kochi (2015)進一步指出亞洲人口多有 PADI4 基因，而歐洲人口尤其北歐多有 PTPN2 基因，發現存在基因異質性(genetic heterogeneity)，因此應該考慮族群差異。

東亞方面，以日本人口為研究對象，Suzuki et al. (2008)提出，風濕性關節炎



的風險基因也可能同時是其他自體免疫性疾病的風險基因，如 PTPN22、FCRL3、STAT4 與類風濕性關節炎、全身性紅斑性狼瘡(Systemic Lupus Erythematosus, SLE)有關；且不同地理分布人群之間風險變異的異質性，例如，PTPN22 和類風濕性關節炎之間的關聯在歐洲族群顯著，但在東亞族群中並不存在 PTPN22 的風險變異。Freudenberg et al. (2011)則以韓國人口為研究對象，指出與類風濕性關節炎有顯著關聯的風險位點。

過去針對歐洲和東亞人口與類風濕性關節炎的統合分析已經確定了超過 100 個風險基因(Laufer et al., 2019)。Zhu et al. (2016)進行亞洲與歐洲人口的類風濕性關節炎多資料庫的全基因體關聯性研究，發現其中有 20 個顯著風險基因，包括 11 個亞洲與歐洲重疊的風險基因、4 個亞洲特定風險基因、5 個歐洲特定風險基因，這表示了亞洲人和歐洲人有一部分共同承擔類風濕性關節炎的遺傳風險，也再次強調了不同種族之間確實存在差異。

第三章 方法及材料

第一節 研究資料

本研究的目的一樣本取自於臺灣人體生物資料庫 (Taiwan Biobank; https://www.twbiobank.org.tw/new_web/)。臺灣人體生物資料庫於 2012 年正式運作，招募 30-70 歲具臺灣國籍且未診斷出癌症的參與者，長期追蹤參與者的健康變化情形，建立龐大的生物檢體與健康資訊，目標人數為 20 萬名一般民眾及 10 萬名常見疾病患者 (包含心血管疾病、糖尿病、慢性腎臟病等)，提供國內生物醫學領域相關研究使用，並致力於發展精準化醫療。此計畫截至 2020 年 2 月社區民眾收案與醫學中心患者有效收案個數共 130,803 人(臺灣人體生物資料庫, 2020)，具有臺灣本土族群之代表性。

經參與者同意後，在研究過程中需接受訪員進行問卷訪談，該問卷中測量的項目包含基本人口學變項 (如性別、年齡)、個人健康行為、生活環境、飲食狀況與家族疾病史等資訊。

本研究取得臺灣人體生物資料庫所釋出之資料庫，包含 16,000 人數，其中排除未提供是否罹患類風濕性關節炎的資訊 78 人，針對 137 位類風濕性關節炎患者，及未罹患類風濕性關節炎患者 15,785 人進行後續的研究分析。

本研究的二使用 GWAS Catalog(<https://www.ebi.ac.uk/gwas/>)，此資料庫是 2008 年由美國國家人類基因組研究(National Human Genome Research Institute, NHGRI)於 2008 年成立，並於 2010 年開始與歐洲生物資訊研究所(European Bioinformatics Institute, EBI)合作的公開資源，匯集了目前已發表的人類全基因組關聯性研究的位點與結果(Buniello et al., 2019)。截至 2019 年 5 月，該目錄包含 3,989 篇發表、138,312 個性狀變異關聯(European Bioinformatics Institute, 2019)。GWAS Catalog 裡的東亞族群(East Asian)為自我報告或被作者描述為東亞或該地區亞人群之一的個體，例如：日本人、韓國人等；歐洲族群(European)為自我報

告或被作者描述為歐洲人、高加索人、白人或該地區亞人群之一的個人，例如：荷蘭人、英國人(Morales et al., 2018)。



第二節 品質控管

對樣本與單核苷酸多型性變異進行品質控管(Quality Control, QC)，可以提高後續數據分析的準確性。以下針對樣本與單核苷酸多型性變異的品質控管進行更詳細的說明：

一、樣本品質控管：


1. 核查性別(check sex)：識別性別訊息不一致的樣本。
2. 樣本的基因定型率(call rate)：用來排除基因定型缺失率太高的樣本。樣本的基因定型率若缺失率太高，代表此樣本的品質不佳，若樣本之基因定型率低於 95%則將此樣本排除，不進行後續分析。

二、單核苷酸多型性變異品質控管：

1. 變異位點的缺失率(missing rate)：用來排除缺失率太高的變異位點，缺失率太高可能會錯失與疾病相關的位點。若變異位點缺失率 $> 10\%$ ，則將此位點移除，不進行後續分析。
2. 次要等位基因頻率(Minor Allele Frequency, MAF)：用來排除次要等位基因頻率太低的位點。低頻率的單核苷酸多型性往往沒有足夠的統計檢力 (statistical power)，若次要等位基因頻率 $< 1\%$ ，則將此位點移除，不進行後續分析。
3. 哈溫平衡檢定(Hardy-Weinberg equilibrium, HWE)：根據遺傳學法則，當族群在理想狀況下，族群中的等位基因頻率會符合哈溫平衡定律。哈溫平衡檢定以 1.0×10^{-6} 為顯著檢定標準，排除不符合的變異位點。

以上步驟皆使用 PLINK(version1.09)軟體完成(Hannah Meyer, 2019)。在進行品質控管後，剩下 15,922 個樣本及 604,202 個單核苷酸多型性位點進行後續分析。

第三節 統計分析



本研究之病例組及對照組使用羅吉斯迴歸模型(logistic regression model)進行，將年齡、性別、有無家族病史視為共變量(covariate)並於模型中控制，排除類風濕性關節炎的高危險因子影響。全基因體關聯性研究(genome-wide association study, GWAS)一般以 $p < 5 \times 10^{-8}$ 為統計顯著標準，Fadista, Manning, Florez, and Groop (2016)提出應重新評估適合的閾值，雖然有許多校正方式如 Benjamini–Hochberg procedure、Bonferroni correction 等，但這些方法可能過度保守而導致偽陰性增加(Kaler & Purcell, 2019)，p 值過於嚴格可能導致無法檢測到某些與遺傳學中複雜性狀相關的重要基因座(Y. M. Zhang, Jia, & Dunwell, 2019)，因此過去研究亦曾建議放寬 p 值標準(Panagiotou & Ioannidis, 2012)。為了進一步降低篩檢標準的嚴格性，接下來使用 $p < 0.0001$ 為標準，而不進行校正(K. Zhang et al., 2018)，並以曼哈頓圖(Manhattan plot)來呈現不同染色體上的核苷酸變異位點。以上步驟使用 PLINK(version 1.09)軟體及 R 軟體(version 3.5.2)進行。

接著，為了解臺灣族群的類風濕性關節炎單核苷酸多型性風險位點與歐洲族群的差異，將以上步驟得到的單核苷酸多型性風險位點，進一步使用 GWAS Catalog 資料庫，對東亞族群與歐洲族群的鹼基含量分別進行信賴區間估計，比較兩族群之間差異是否顯著。以上步驟使用 R 軟體(version 3.5.2)進行。

第四章 結果

第一節 人口學變項分布

表一為人口學變項分布結果，在性別分布中，病例組女性比例高於男性，在 137 個病例中，女性有 89 位(65%)，男性有 48 位(35%)，與過去研究結果女性發生率及盛行率高於男性一致；對照組男女比例分布平均，女性為 50.1%，男性為 49.9%。在類風濕性關節炎家族病史的部分，病例組中有家族病史比例(10.2%)明顯高於對照組的比例(1.9%)。年齡分布上，相對於對照組在不同年齡層中均勻分布(平均年齡為 49.7 歲)，病例組中的高齡族群的比例稍高(平均年齡為 54.1 歲)。

第二節 品質控管

在原始資料中 16,000 人，排除未提供是否罹患類風濕性關節炎的資訊 78 人，共有病例組 137 位、對照組 15,785 位及 646,735 個單核苷酸多型性位點資訊。在圖一品質控管流程圖中，首先進行樣本品質控管，此步驟無樣本被排除；接下來進行單核苷酸多型性品質控管，分別有 21,881 個單核苷酸多型性位點因基因型數據缺失被排除、2,514 個單核苷酸多型性位點因次要等位基因頻率 < 1% 而排除、18,138 個單核苷酸多型性位點因不符合哈溫平衡而排除。經品質控管後剩下 137 位病例組、15,785 位對照組及 604,202 個單核苷酸多型性位點。

第三節 全基因體關聯性分析

以年齡、性別、有無類風濕性關節炎家族病史為共變量，使用羅吉斯回歸模型進行分析，並以 $p < 0.0001$ 為統計標準(K. Zhang et al., 2018)，共有 36 個單核苷酸多型性位點顯著，其中有 26 個位點位於基因上，依序為 rs4708887(PLG)、rs12214168(PLG)、rs661311(MEF2C)、rs12611515(NCKAP5)、rs349135(ARAP3)、rs78849619(TMEM192)、rs10191147(NCKAP5)、rs75895670(LSAMP)、


rs1461957(SRP9)、rs3823055(PLG)、rs10199481(NCKAP5)、rs4682542(ZBTB20)、rs3757017(PLG)、rs13062127(YEATS2)、rs14224(PLG)、rs2295368(PLG)、rs1594307(PLG)、rs17638386(FSTL5)、rs9295131(PLG)、rs1465620(PLG)、rs10511370(LSAMP)、rs783147(PLG)、rs783146(PLG)、rs6677132(SRP9)、rs7601203(TBC1D8)、rs10210192(OLA1)，於表二呈現。

第四節 東亞與歐洲族群的風險位點比較

為進一步了解臺灣人口的類風濕性關節炎相關單核苷酸多型性位點與歐洲人口的差異，將以上 36 個單核苷酸多型性位點分別比較東亞與歐洲族群的信賴區間估計(Confidence Interval Estimate)，得到的信賴區間於表三呈現，可以看出有 5 個位點的信賴區間重疊，東亞族群與歐洲族群的差異不顯著，分別為 rs13062127、rs1582626、rs14224、rs6745548、rs7601203。Affx-20961401 因目前在 GWAS Catalog 資料庫中並無對應的東亞、歐洲族群位點資訊，故無法比較兩族群之差異。其餘 30 個位點在東亞與歐洲族群間有顯著差異。

第五節 驗證過去研究


在本研究得到的 36 個顯著風險位點，並未發現與類風濕性關節炎的過去研究文獻。但在過去文獻中發現 rs14224、rs2295368、rs783147 與脂蛋白相關，脂蛋白濃度過高時會引起發炎反應，易導致動脈粥狀硬化等心血管疾病(Deshmukh et al., 2012; Enkhmaa, Anuurad, & Berglund, 2016; Enkhmaa, Anuurad, Zhang, Tran, & Berglund, 2011; Q. Qi, Workalemahu, Zhang, Hu, & Qi, 2012; Ronald et al., 2011)，雖然與類風濕性關節炎未有直接關聯性，然而可以看出與發炎反應仍有相關，且與類風濕性關節炎患者有較高機會發生心血管合併症吻合，因此這三個位點對於類風濕性關節炎的影響，尤其 rs2295368、rs783147 在臺灣族群與歐洲有顯著差異，仍有待未來研究深入探討。而 rs9295131 在過去文獻中為胎盤早期剝離的候選基因，但研究結果顯示不顯著(Moore et al., 2012)。



進一步分析，26 個位點位於基因上，其中有 16 個位點位於 PLG，過去研究已發表 PLG 與脂蛋白(Lipoprotein)、血液蛋白(Blood protein)等有關(Mack et al., 2017)，血液蛋白中有一大類是抗體，能夠與分子標靶結合，這些結合是人體抵抗感染不可缺少的步驟。如果此機制出了問題，容易罹患自體免疫疾病。另外，有 3 個位點位於 NCKAP5，是一種蛋白質譯碼基因(protein-coding genes)，過去研究已發表與多發性硬化症(Multiple sclerosis)有關(Baranzini et al., 2009)，多發性硬化症也是一種自體免疫疾病，類風濕性關節炎風險基因常也會增加其他自體免疫性疾病的風險(Suzuki et al., 2008)。其他在過去研究已發表的 LSAMP、OLA1 與血液蛋白(Blood protein)有關(Emilsson et al., 2018);YEATS2 與 C-反應蛋白(CRP)有關(Han et al., 2020)、TBC1D8 與白血球計數(WBC count)(Kichaev et al., 2019)、脂蛋白有關(Liu et al., 2017; G. Qi & Chatterjee, 2018)，C-反應蛋白(CRP)及白血球計數(WBC count)皆是臨床上發炎反應的常用指標(Willems, Trompet, Blauw, Westendorp, & de Craen, 2010)。以上與性狀的對應整理於表四。

第五章 討論

第一節 主要發現



本篇為第一篇使用臺灣人口族群針對類風濕性關節炎所做的全基因體關聯性研究。過去文獻指出 $p < 5 \times 10^{-8}$ 對於分析歐洲人群中常見的遺傳變異有效，但在研究不同族群時，因遺傳變異程度也不盡相同，適當的統計顯著性閾值也應該有所調整(Fadista et al., 2016)。過去全基因體關聯性研究也分別使用了不同的統計顯著性閾值：Julià et al. (2008)針對過去兩篇研究使用 $p < 0.0001$ 與 $p < 0.005$ 進行分析；Uh et al. (2012)指出 $p < 0.001$ 的位點與疾病仍有高度相關；Aslibekyan et al. (2012)則使用 $p < 0.0001$ 為統計顯著標準；Sasayama et al. (2013)在一篇分析亞洲族群遺傳變異的文章裡使用 $p < 0.0001$ 為閾值，並指出此閾值有助於縮小候選風險位點的範圍，能包括一定比例的風險位點；Doherty et al. (2018)則因樣本數較少而使用 $p < 5 \times 10^{-7}$ 為閾值。即使這些研究各自選用的統計顯著性閾值有所不同，其篩選後的位點包括已被確立的疾病風險位點，另一部分則提出可能的風險位點供後續研究驗證。在嚴格的統計顯著標準 $p < 5 \times 10^{-8}$ ，常無法檢測到許多與疾病真正相關但微弱的風險位點(Jia, Wang, Meltzer, & Zhao, 2011)，然而，同時也會降低檢測疾病風險位點的統計檢定力(Riancho, 2012)。目前，透過全基因體關聯性研究辨識出的風險位點來預測疾病，多為高特異度 (specificity, 80-90%)、低敏感性(sensitivity, ~40%)，導致高風險者易被誤分類成低風險(Yarwood et al., 2016)。因此本研究嘗試使用 $p < 0.0001$ 的統計標準(圖三)，而非使用一般全基因體關聯性研究的顯著標準($p < 5 \times 10^{-8}$)，提供另一視角來檢視單核苷酸多型性位點與類風濕性關節炎的關聯，並進一步看出東亞與歐洲族群間確實存在差異性。

然而，目前已發表的類風濕性關節炎相關的基因或位點，研究樣本大部分是歐洲族群，在台灣、甚至東亞的風險位點會有什麼變化仍需要更多研究來探討。本研究指出 3 個可能的風險位點(rs14224、rs2295368、rs783147)與 6 個可能的風

險基因(PLG、NCKAP5、LSAMP、OLA1、YEATS2、TBC1D8)，目前與類風濕性關節炎的直接相關性尚未有過去文獻佐證，有待未來研究深入探討。

此外，比對 GWAS Catalog 資料庫裡已發表的類風濕性關節炎風險位點，本研究結果有 59 個位點一致，但結果並不顯著($p < 0.0001$)，整理於表五。其中有 17 個位點的研究樣本來源包含東亞族群，但在本研究中結果不顯著，可能原因有二：1)東亞族群的研究樣本數較小，所以比歐洲族群研究的統計檢定力較低；2)東亞族群的異質性(Yamamoto et al., 2015)。未來能針對這些位點繼續深入探討，提供臺灣族群類風濕性關節炎患者風險評估之參考依據。


第二節 研究限制

本研究有以下研究限制，第一為樣本數不足，全基因體關聯性研究需要更大量的樣本數來提高統計檢定力(statistical power)，然而本研究已是目前從臺灣人體生物資料庫能取得的最大樣本數，期許未來能招募到更多臺灣人口的基因訊息，能有更大量的樣本數做研究分析。第二為樣本的健康問卷為自陳(self-report)資料，樣本自己表示是否罹患類風濕性關節炎疾病，故可能存在患有此疾病者表示沒有罹患此疾病，反之亦然。第三為無法確認因果關係，全基因體關聯性研究為觀察性研究，無法確認研究所發現的單核苷酸多型性位點變異與類風濕性關節炎的因果關係。

第三節 公共衛生與臨床意義

本研究進行的類風濕性關節炎全基因體關聯性研究提出了臺灣族群可能特有的變異位點。目前已知的類風濕性關節炎相關位點多集中在歐洲族群樣本的研究，即使在不同族群之間共享風險位點，不同族群間仍存在差異(Laufer et al., 2019)，這也支持需要對臺灣人口類風濕性關節炎風險位點進行更深入的分析。

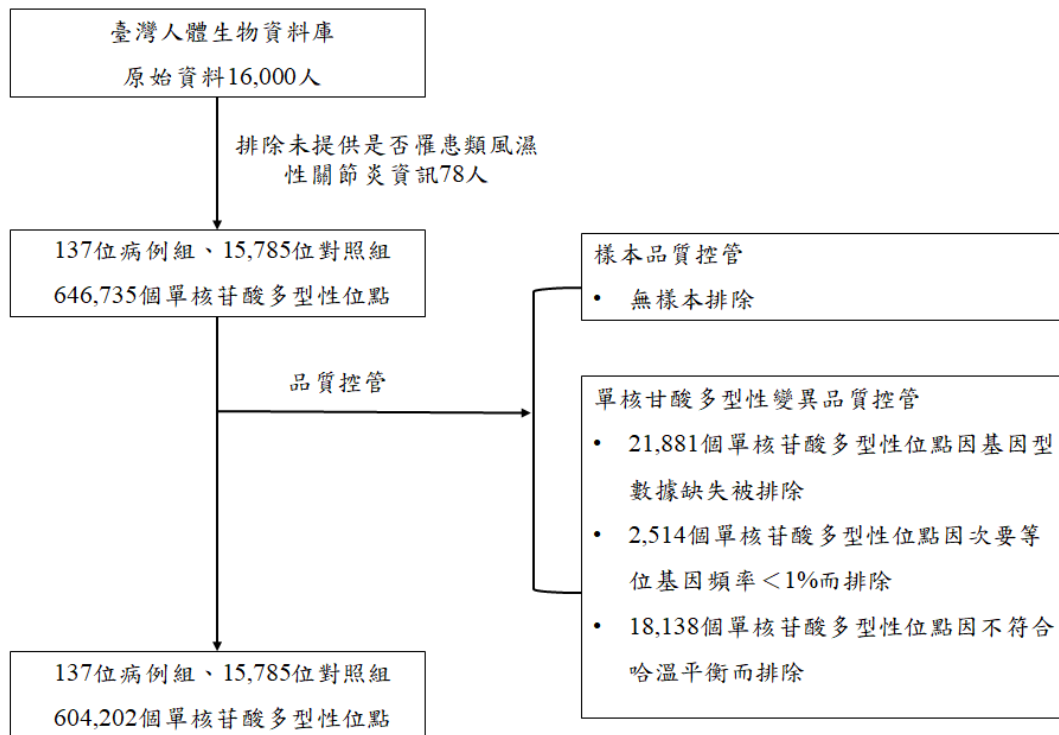
透過全基因體關聯性研究發現的遺傳變異與疾病之間的顯著關聯有助於：1)對疾病易感性的洞察：了解新的致病基因和機轉，用於識別疾病高危險族群；2)



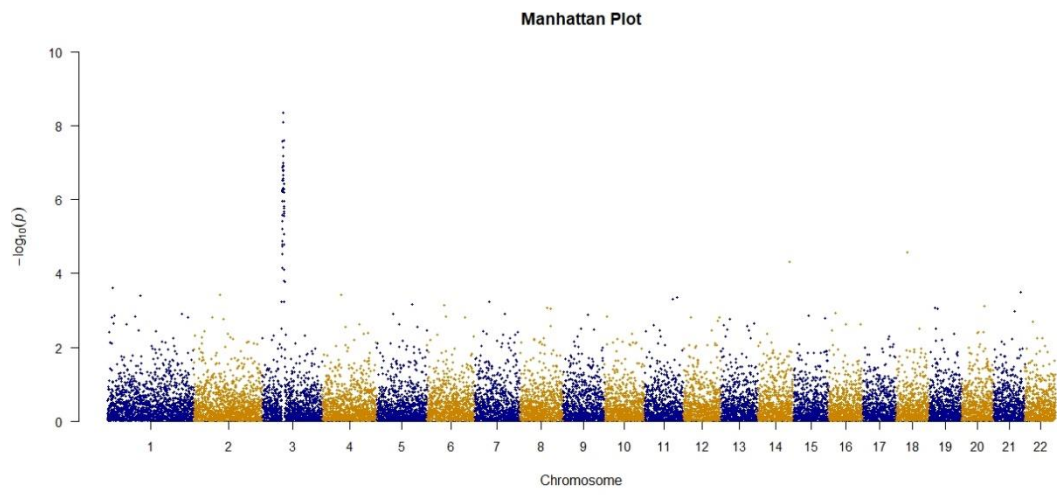
臨床護理：新藥靶標和疾病生物標誌物的鑑定，可以提高藥物開發的成功率，減少開發新藥物的時間和成本；3)精準醫療：根據個體基因型預測疾病風險，及早發現、預防或優化治療方案，提高患者預後(Tam et al., 2019; Tang, 2019; Viatte & Barton, 2017)。儘管目前已經確定了多種基因在預測患者治療狀態中的效果，但未完全闡明其中一些基因在類風濕性關節炎發病機轉中的作用，因此類風濕性關節炎的大型遺傳學關聯性研究仍十分重要(Karami, Aslani, Jamshidi, Garshasbi, & Mahmoudi, 2019)。

類風濕性關節炎是一高度異質性的疾病所以不容易預測，而異質性有部分是因為遺傳因素(Yamamoto et al., 2015)，未來能利用東亞或其他地區資料對本研究結果進行比對或驗證(Validation)，進一步探討多基因組合、添加其他風險因素(例如環境因素、生物標誌物和臨床預測因子)能提高預測模型的功能，應用於臨床診斷與醫療發展(Yarwood et al., 2016)。類風濕性關節炎對社會經濟層面的影響足以構成重大的公共衛生議題(Minichiello et al., 2016)，辨識出更多風險位點變異對疾病的影響，可能是未來預防類風濕性關節炎的關鍵。

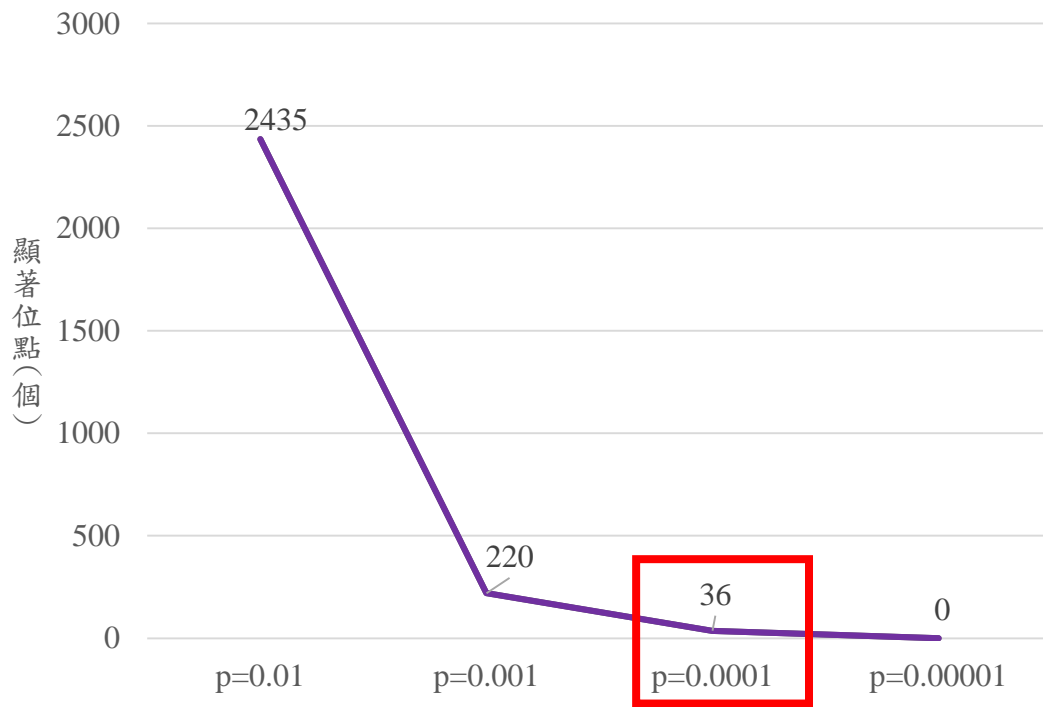
圖一、品質控管流程圖



圖二、曼哈頓圖



圖三、本研究不同 p 值篩選出的顯著位點量



表一、人口學變項分布

	病例組(n=137)		對照組(n=15,785)	
	N	%	N	%
性別				
男性	48	35.0	7883	49.9
女性	89	65.0	7902	50.1
類風濕性關節炎 家族病史				
有	14	10.2	304	1.9
無	123	89.8	15481	98.1
年齡(歲)				
<30	0	0.0	0	0.0
30-40	14	10.2	3828	24.3
40-50	32	23.4	3941	25.0
50-60	35	25.5	3956	25.1
60-70	53	38.7	3910	24.8
≥70	3	2.2	150	1.0
平均值±標準差	54.12±10.64		49.67±11.35	



表二、36 個顯著單核苷酸多型性位點

SNP	Gene	CHR	BP	OR	P value
rs4708887	PLG	6	161206387	0.5489	1.02E-05
rs12214168	PLG	6	161195801	0.5518	1.13E-05
rs661311	MEF2C	5	88034867	2.349	1.59E-05
Affx-20961401	—*	3	114245776	2.046	1.65E-05
rs12611515	NCKAP5	2	133541575	0.5713	1.90E-05
rs349135	ARAP3	5	141066161	1.692	2.30E-05
rs78849619	TMEM192	4	166038634	1.782	2.60E-05
rs10191147	NCKAP5	2	133542545	0.5773	2.63E-05
rs75895670	LSAMP	3	117334083	2.039	2.83E-05
rs1461957	SRP9	1	225965832	1.946	2.94E-05
rs76488461	intergenic	6	156973759	1.986	3.10E-05
rs3823055	PLG	6	161138917	1.671	3.28E-05
rs10199481	NCKAP5	2	133577528	0.5852	3.62E-05
rs2378013	intergenic	1	218951464	0.5215	3.70E-05
rs4682542	ZBTB20	3	114247125	1.999	4.13E-05
rs3757017	PLG	6	161158734	0.5949	4.13E-05
rs13062127	YEATS2	3	183503634	1.909	4.46E-05
rs1582626	intergenic	2	186199224	1.66	4.98E-05
rs14224	PLG	6	161137779	1.648	4.98E-05
rs17026435	intergenic	4	150683768	1.811	5.25E-05
rs2295368	PLG	6	161139502	1.645	5.32E-05
rs1594307	PLG	6	161141184	1.652	5.36E-05
rs17638386	FSTL5	4	162596227	2.007	5.40E-05
rs6745548	intergenic	2	186187392	1.647	5.77E-05
rs9295131	PLG	6	161136294	1.64	5.84E-05
rs1465620	PLG	6	161136485	1.637	6.20E-05
rs10511370	LSAMP	3	117325977	1.989	6.47E-05
rs850908	intergenic	2	186220642	1.657	6.78E-05
rs783147	PLG	6	161137990	1.629	7.77E-05
rs783146	PLG	6	161138848	1.631	7.92E-05
rs6677132	SRP9	1	225967597	1.709	8.19E-05
rs7601203	TBC1D8	2	101727601	1.687	8.62E-05
rs80200347	intergenic	4	12677786	2.302	8.76E-05
rs10210192	OLA1	2	174924658	1.638	8.94E-05
rs2649086	intergenic	1	5782769	1.62	8.97E-05
rs61294929	intergenic	6	157038510	2.007	9.33E-05

SNP: Single Nucleotide Polymorphism; CHR: Chromosome; BP: base pair

*Affx-20961401 目前無對應的基因

表三、東亞與歐洲族群信賴區間估計(36 個顯著單核苷酸多型性位點)

SNP	Gene	Alleles	MAF	東亞族群位點信賴區間	歐洲族群位點信賴區間
rs4708887	PLG	A/G	0.47	(0.385, 0.446)	(0.617, 0.676)
rs12214168	PLG	C/T	0.41	(0.387, 0.448)	(0.623, 0.682)
rs661311	MEF2C	T/C	0.29	(0.051, 0.082)	(0.194, 0.245)
Affx-20961401	—*	—*	—*	—*	—*
rs12611515	NCKAP5	C/T	0.37	(0.472, 0.534)	(0.677, 0.733)
rs349135	ARAP3	A/G	0.46	(0.641, 0.699)	(0.441, 0.503)
rs78849619	TMEM192	T/C	0.04	(0.823, 0.868)	(0.969, 0.987)
rs10191147	NCKAP5	G/A	0.35	(0.475, 0.537)	(0.718, 0.772)
rs75895670	LSAMP	G/T	0.03	(0.925, 0.954)	(0.988, 0.998)
rs1461957	SRP9	G/T	0.12	(0.868, 0.907)	(0.783, 0.832)
rs76488461	intergenic	A/G	0.03	(0.878, 0.916)	(1.000, 0.000)
rs3823055	PLG	A/G	0.41	(0.513, 0.574)	(0.281, 0.338)
rs10199481	NCKAP5	T/C	0.41	(0.454, 0.516)	(0.644, 0.702)
rs2378013	intergenic	G/A	0.16	(0.664, 0.721)	(0.900, 0.935)
rs4682542	ZBTB20	T/C	0.10	(0.901, 0.935)	(0.954, 0.977)
rs3757017	PLG	T/C	0.39	(0.451, 0.513)	(0.315, 0.373)
rs13062127	YEATS2	C/T	0.04	(0.891, 0.927)	(0.923, 0.953)
rs1582626	intergenic	C/T	0.32	(0.674, 0.731)	(0.665, 0.722)
rs14224	PLG	T/C	0.45	(0.513, 0.574)	(0.572, 0.633)
rs17026435	intergenic	A/G	0.11	(0.811, 0.857)	(0.963, 0.983)
rs2295368	PLG	G/A	0.35	(0.514, 0.576)	(0.580, 0.641)
rs1594307	PLG	C/T	0.40	(0.483, 0.545)	(0.281, 0.338)
rs17638386	FSTL5	A/G	0.03	(0.883, 0.920)	(1.000, 0.000)
rs6745548	intergenic	T/C	0.32	(0.669, 0.726)	(0.654, 0.713)
rs9295131	PLG	G/A	0.49	(0.518, 0.579)	(0.289, 0.347)
rs1465620	PLG	G/A	0.49	(0.518, 0.579)	(0.289, 0.347)
rs10511370	LSAMP	C/T	0.20	(0.930, 0.959)	(0.832, 0.876)
rs850908	intergenic	C/A	0.28	(0.231, 0.285)	(0.175, 0.225)
rs783147	PLG	G/A	0.40	(0.444, 0.506)	(0.523, 0.584)
rs783146	PLG	C/G	0.28	(0.444, 0.506)	(0.825, 0.869)
rs6677132	SRP9	A/C	0.38	(0.773, 0.822)	(0.626, 0.684)
rs7601203	TBC1D8	C/T	0.16	(0.737, 0.789)	(0.785, 0.833)
rs80200347	intergenic	T/G	0.01	(0.945, 0.970)	(0.987, 0.998)
rs10210192	OLA1	A/G	0.45	(0.473, 0.535)	(0.767, 0.817)
rs2649086	intergenic	T/C	0.32	(0.370, 0.430)	(0.248, 0.303)
rs61294929	intergenic	C/T	0.06	(0.900, 0.934)	(0.980, 0.994)

SNP: Single Nucleotide Polymorphism

*Affx-20961401 目前無對應的基因

表四、顯著風險位點與性狀關聯

SNP	Traits
rs14224	脂蛋白(Lipoprotein)
rs2295368	脂蛋白(Lipoprotein)
rs783147	脂蛋白(Lipoprotein)
Gene	Traits
PLG	脂蛋白(Lipoprotein)、血液蛋白(Blood protein)
NCKAP5	多發性硬化症(Multiple sclerosis)
LSAMP	血液蛋白(Blood protein)
OLA1	血液蛋白(Blood protein)
YEATS2	C-反應蛋白(C-reactive protein, CRP)
TBC1D8	脂蛋白(Lipoprotein)、白血球計數(WBC count)

SNP: Single Nucleotide Polymorphism



表五、59 個未顯著但在 GWAS Catalog 的類風濕性關節炎風險位點

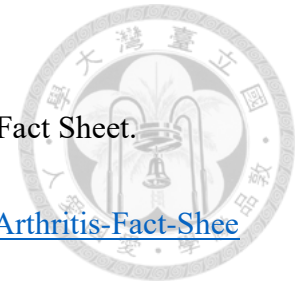
SNP	P value	研究樣本來源*
rs3890745	0.03	歐洲(3,393 cases, 12,460 controls; 3,929 cases, 5,807 controls)
rs11900673	0.03519	日本(4,074 cases, 16,891 controls; 5,277 cases, 21,684 controls)
rs7731626	0.04598	歐洲(14,361 cases, 42,923 controls; 3,775 cases, 5,801 controls) 東亞(4,873 cases, 17,642 controls; 6,871 cases, 6,392 controls)
rs840016	0.08078	歐洲(5,539 cases, 20,169 controls; 6,768 cases, 8,806 controls)
rs2240335	0.09951	韓國(801 cases, 757 controls; 718 cases, 719 controls)
rs16867384	0.1306	歐洲(4,418 cases, 3,300 controls)
rs2867461	0.1429	日本(4,074 cases, 16,891 controls; 5,277 cases, 21,684 controls)
rs7579944	0.1655	歐洲(4,418 cases, 3,300 controls)
rs9296015	0.1939	日本(1,247 cases, 1,486 controls; 4,168 cases, 5,003 controls)
rs1875463	0.201	歐洲(13,838 cases, 33,742 controls)
rs7765379	0.2062	韓國(801 cases, 757 controls; 718 cases, 719 controls)
rs10174238	0.2255	歐洲(4,595 cases, 19,704 controls)
rs660895	0.2388	歐洲(3,297 cases, 15,870 controls)
rs1854853	0.2402	漢族(952 cases, 943 controls; 2,132 cases, 2,553 controls)
rs10484565	0.2875	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs9784858	0.3047	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs9275334	0.3253	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs3998159	0.3396	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs802791	0.3529	歐洲(4,595 cases, 19,704 controls)
rs3957146	0.3548	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs7454108	0.3556	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs4921283	0.3652	歐洲(3,911 cases, 10,398 controls)
rs9271348	0.3944	歐洲(1,148 cases, 6,008 controls; 774 cases, 1,079 controls)
rs13015080	0.4009	日本(31 mild cases, 282 moderate cases, 131 severe cases)
rs3806624	0.4053	南美(916 cases, 1,392 controls) 東亞(4,873 cases, 17,641 controls) 歐洲(19,234 cases, 61,654 controls)
rs3129890	0.4262	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs6457620	0.4265	歐洲(3,393 cases, 12,460 controls; 3,929 cases, 5,807 controls)
rs744600	0.428	歐洲(4,595 cases, 19,704 controls)
rs1571878	0.432	歐洲(9,585 cases, 33,742 controls)
rs7574865	0.4387	歐洲(5,539 cases, 20,169 controls; 6,768 cases, 8,806 controls)
rs6596147	0.4438	歐洲(868 cases, 1,194 controls)
rs11937061	0.448	歐洲(1,157 cases; 526 cases)

表五、59 個未顯著但在 GWAS Catalog 的類風濕性關節炎風險位點(續)

SNP	P value	研究樣本來源*
rs657075	0.4596	南美(916 cases, 1,392 controls) 東亞(4,873 cases, 17,641 controls) 歐洲(19,234 cases, 61,654 controls)
rs6457617	0.5032	歐洲(400 cases, 400 controls; 410 cases, 394 controls)
rs6931277	0.5039	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs12525220	0.5298	漢族(952 cases, 943 controls; 2,132 cases, 2,553 controls)
rs12529514	0.5341	日本(4,074 cases, 16,891 controls; 5,277 cases, 21,684 controls)
rs934734	0.5474	歐洲(5,539 cases, 20,169 controls; 6,768 cases, 8,806 controls)
rs17604670	0.5702	歐洲(868 cases, 1,194 controls)
rs11571302	0.5763	歐洲(9,585 cases, 3,742 controls)
rs6568431	0.6031	歐洲(3,911 cases, 10,398 controls; 13,641 cases, 36,509 controls)
rs615672	0.6035	歐洲(1,860 cases, 2,938 controls)
rs3093024	0.655	日本(2,303 cases, 3,380 controls; 4,768 cases, 17,359 controls)
rs3093023	0.6666	歐洲(5,539 cases, 20,169 controls; 6,768 cases, 8,806 controls)
rs231735	0.6706	歐洲(2,418 cases, 4,504 controls; 2,604 cases, 2,882 controls)
rs2233424	0.684	南美(916 cases, 1,392 controls) 東亞(4,873 cases, 17,641 controls) 歐洲(19,234 cases, 61,654 controls)
rs6910071	0.7024	歐洲(5,539 cases, 20,169 controls; 6,768 cases, 8,806 controls)
rs9268839	0.7596	歐洲(14,361 cases, 42,923 controls; 3,775 cases, 5,801 controls) 東亞(4,873 cases, 17,642 controls; 6,871 cases, 6,392 controls)
rs3087243	0.7731	歐洲(5,539 cases, 20,169 controls; 6,768 cases, 8,806 controls)
rs3731714	0.7763	歐洲(3,921 cases, 11,232 controls)
rs2561477	0.7891	南美(916 cases, 1,392 controls) 東亞(4,873 cases, 17,641 controls) 歐洲(19,234 cases, 61,654 controls)
rs9268557	0.8286	歐洲(833 cases, 1,740 controls; 608 cases, 908 controls)
rs3763309	0.8424	歐洲(3,034 cases, 5,271 controls; 4,726 cases, 2,625 controls)
rs10499194	0.9447	歐洲(397 cases, 1,211 controls; 875 cases, 832 controls)
rs6448119	0.9448	歐洲(1,921 cases, 1,079 controls; 887 cases, 1,218 controls)
rs26232	0.9556	歐洲(5,539 cases, 20,169 controls; 6,768 cases, 8,806 controls)
rs3816587	0.9575	歐洲(1,860 cases, 2,938 controls)
rs2062583	0.9664	韓國(801 cases, 757 controls; 718 cases, 719 controls)
rs13192471	0.9883	日本(2,303 cases, 3,380 controls; 4,768 cases, 17,359 controls)

*Retrieved from GWAS Catalog (<https://www.ebi.ac.uk/gwas/>)

參考文獻



- American College of Rheumatology. (2019). Rheumatoid Arthritis Fact Sheet.
Retrieved from
<https://www.rheumatology.org/Portals/0/Files/Rheumatoid-Arthritis-Fact-Sheet.pdf?ver=2019-06-20-145158-000>
- Aslibekyan, S., Goodarzi, M. O., Frazier-Wood, A. C., Yan, X., Irvin, M. R., Kim, E., . . . Arnett, D. K. (2012). Variants identified in a GWAS meta-analysis for blood lipids are associated with the lipid response to fenofibrate. *PLoS ONE*, 7(10), e48663. doi:10.1371/journal.pone.0048663
- Ball, R. D. (2013). Designing a GWAS: power, sample size, and data structure. *Methods Mol Biol*, 1019, 37-98. doi:10.1007/978-1-62703-447-0_3
- Baranzini, S. E., Wang, J., Gibson, R. A., Galwey, N., Naegelin, Y., Barkhof, F., Oksenberg, J. R. (2009). Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet*, 18(4), 767-778. doi:10.1093/hmg/ddn388
- Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*, 47(D1), D1005-d1012. doi:10.1093/nar/gky1120
- Deshmukh, H. A., Colhoun, H. M., Johnson, T., McKeigue, P. M., Betteridge, D. J., Durrington, P. N., Hitman, G. A. (2012). Genome-wide association study of genetic determinants of LDL-c response to atorvastatin therapy: importance of Lp(a). *J Lipid Res*, 53(5), 1000-1011. doi:10.1194/jlr.P021113
- Doherty, A., Smith-Byrne, K., Ferreira, T., Holmes, M. V., Holmes, C., Pulit, S. L., & Lindgren, C. M. (2018). GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nat Commun*, 9(1), 5257. doi:10.1038/s41467-018-07743-4
- Emilsson, V., Ilkov, M., Lamb, J. R., Finkel, N., Gudmundsson, E. F., Pitts, R., Gudnason, V. (2018). Co-regulatory networks of human serum proteins link genetics to disease. *Science*, 361(6404), 769-773. doi:10.1126/science.aaq1327
- Enkhmaa, B., Anuurad, E., & Berglund, L. (2016). Lipoprotein (a): impact by ethnicity and environmental and medical conditions. *J Lipid Res*, 57(7), 1111-1125. doi:10.1194/jlr.R051904
- Enkhmaa, B., Anuurad, E., Zhang, W., Tran, T., & Berglund, L. (2011). Lipoprotein(a): genotype-phenotype relationship and impact on atherogenic risk. *Metab Syndr Relat Disord*, 9(6), 411-418. doi:10.1089/met.2011.0026

- European Bioinformatics Institute. (2019). What is the GWAS Catalog? Retrieved from <https://www.ebi.ac.uk/training/online/course/gwas-catalog-exploring-snp-trait-associations-2019/what-gwas-catalog>
- Fadista, J., Manning, A. K., Florez, J. C., & Groop, L. (2016). The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. *Eur J Hum Genet*, 24(8), 1202-1205. doi:10.1038/ejhg.2015.269
- Freudenberg, J., Lee, H. S., Han, B. G., Shin, H. D., Kang, Y. M., Sung, Y. K., Bae, S. C. (2011). Genome-wide association study of rheumatoid arthritis in Koreans: population-specific loci as well as overlap with European susceptibility loci. *Arthritis Rheum*, 63(4), 884-893. doi:10.1002/art.30235
- Greenberg, J. D., Spruill, T. M., Shan, Y., Reed, G., Kremer, J. M., Potter, J., Harrold, L. R. (2013). Racial and ethnic disparities in disease activity in patients with rheumatoid arthritis. *Am J Med*, 126(12), 1089-1098. doi:10.1016/j.amjmed.2013.09.002
- Han, X., Ong, J. S., An, J., Hewitt, A. W., Gharahkhani, P., & MacGregor, S. (2020). Using Mendelian randomization to evaluate the causal relationship between serum C-reactive protein levels and age-related macular degeneration. *Eur J Epidemiol*, 35(2), 139-146. doi:10.1007/s10654-019-00598-z
- Hannah Meyer. (2019). Genotype quality control with plinkQC. Retrieved from <https://cran.r-project.org/web/packages/plinkQC/vignettes/plinkQC.pdf>
- Jalil, S. F., Arshad, M., Bhatti, A., Ahmad, J., Akbar, F., Ali, S., & John, P. (2016). Rheumatoid arthritis: What have we learned about the causing factors? *Pak J Pharm Sci*, 29(2), 629-645.
- Jia, P., Wang, L., Meltzer, H. Y., & Zhao, Z. (2011). Pathway-based analysis of GWAS datasets: effective but caution required. *Int J Neuropsychopharmacol*, 14(4), 567-572. doi:10.1017/s1461145710001446
- Jin, S., Li, M., Fang, Y., Li, Q., Liu, J., Duan, X., Zeng, X. (2017). Chinese Registry of rheumatoid arthritis (CREDIT): II. prevalence and risk factors of major comorbidities in Chinese patients with rheumatoid arthritis. *Arthritis Res Ther*, 19(1), 251. doi:10.1186/s13075-017-1457-z
- Julià, A., Ballina, J., Cañete, J. D., Balsa, A., Tornero-Molina, J., Naranjo, A., Marsal, S. (2008). Genome-wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility. *Arthritis Rheum*, 58(8), 2275-2286. doi:10.1002/art.23623
- Jutley, G., Raza, K., & Buckley, C. D. (2015). New pathogenic insights into rheumatoid arthritis. *Curr Opin Rheumatol*, 27(3), 249-255. doi:10.1097/bor.0000000000000174

- Kaler, A. S., & Purcell, L. C. (2019). Estimation of a significance threshold for genome-wide association studies. *BMC Genomics*, *20*(1), 618. doi:10.1186/s12864-019-5992-7
- Karami, J., Aslani, S., Jamshidi, A., Garshasbi, M., & Mahmoudi, M. (2019). Genetic implications in the pathogenesis of rheumatoid arthritis; an updated review. *Gene*, *702*, 8-16. doi:10.1016/j.gene.2019.03.033
- Kichaev, G., Bhatia, G., Loh, P. R., Gazal, S., Burch, K., Freund, M. K., Price, A. L. (2019). Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet*, *104*(1), 65-75. doi:10.1016/j.ajhg.2018.11.008
- Kochi, Y., Suzuki, A., & Yamamoto, K. (2014). Genetic basis of rheumatoid arthritis: a current review. *Biochem Biophys Res Commun*, *452*(2), 254-262. doi:10.1016/j.bbrc.2014.07.085
- Kuo, C. F., Grainge, M. J., Valdes, A. M., See, L. C., Yu, K. H., Shaw, S. W. S., Doherty, M. (2017). Familial aggregation of rheumatoid arthritis and co-aggregation of autoimmune diseases in affected families: a nationwide population-based study. *Rheumatology (Oxford)*, *56*(6), 928-933. doi:10.1093/rheumatology/kew500
- Kuo, C. F., Luo, S. F., See, L. C., Chou, I. J., Chang, H. C., & Yu, K. H. (2013). Rheumatoid arthritis prevalence, incidence, and mortality rates: a nationwide population study in Taiwan. *Rheumatol Int*, *33*(2), 355-360. doi:10.1007/s00296-012-2411-7
- Laufer, V. A., Tiwari, H. K., Reynolds, R. J., Danila, M. I., Wang, J., Edberg, J. C., Bridges, S. L., Jr. (2019). Genetic influences on susceptibility to rheumatoid arthritis in African-Americans. *Hum Mol Genet*, *28*(5), 858-874. doi:10.1093/hmg/ddy395
- Liu, D. J., Peloso, G. M., Yu, H., Butterworth, A. S., Wang, X., Mahajan, A., Kathiresan, S. (2017). Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet*, *49*(12), 1758-1766. doi:10.1038/ng.3977
- Mack, S., Coassin, S., Rueedi, R., Yousri, N. A., Seppälä, I., Gieger, C., Kronenberg, F. (2017). A genome-wide association meta-analysis on lipoprotein (a) concentrations adjusted for apolipoprotein (a) isoforms. *J Lipid Res*, *58*(9), 1834-1844. doi:10.1194/jlr.M076232
- McBurney, C. A., & Vina, E. R. (2012). Racial and ethnic disparities in rheumatoid arthritis. *Curr Rheumatol Rep*, *14*(5), 463-471. doi:10.1007/s11926-012-0276-0
- McInnes, I. B., & Schett, G. (2011). The pathogenesis of rheumatoid arthritis. *N Engl J Med*, *365*(23), 2205-2219. doi:10.1056/NEJMra1004965
- Minichiello, E., Semerano, L., & Boissier, M. C. (2016). Time trends in the incidence,

- prevalence, and severity of rheumatoid arthritis: A systematic literature review. *Joint Bone Spine*, 83(6), 625-630. doi:10.1016/j.jbspin.2016.07.007
- Moore, A., Enquobahrie, D. A., Sanchez, S. E., Ananth, C. V., Pacora, P. N., & Williams, M. A. (2012). A genome-wide association study of variations in maternal cardiometabolic genes and risk of placental abruption. *Int J Mol Epidemiol Genet*, 3(4), 305-313.
- Morales, J., Welter, D., Bowler, E. H., Cerezo, M., Harris, L. W., McMahon, A. C., MacArthur, J. A. L. (2018). A standardized framework for representation of ancestry data in genomics studies, with application to the NHGRI-EBI GWAS Catalog. *Genome Biol*, 19(1), 21. doi:10.1186/s13059-018-1396-2
- Ndosi, M., Vinall, K., Hale, C., Bird, H., & Hill, J. (2011). The effectiveness of nurse-led care in people with rheumatoid arthritis: a systematic review. *Int J Nurs Stud*, 48(5), 642-654. doi:10.1016/j.ijnurstu.2011.02.007
- Panagiotou, O. A., & Ioannidis, J. P. (2012). What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *Int J Epidemiol*, 41(1), 273-286. doi:10.1093/ije/dyr178
- Qi, G., & Chatterjee, N. (2018). Heritability informed power optimization (HIPO) leads to enhanced detection of genetic associations across multiple traits. *PLoS Genet*, 14(10), e1007549. doi:10.1371/journal.pgen.1007549
- Qi, Q., Workalemahu, T., Zhang, C., Hu, F. B., & Qi, L. (2012). Genetic variants, plasma lipoprotein(a) levels, and risk of cardiovascular morbidity and mortality among two prospective cohorts of type 2 diabetes. *Eur Heart J*, 33(3), 325-334. doi:10.1093/eurheartj/ehr350
- Rheumatoid arthritis. (2018). *Nat Rev Dis Primers*, 4, 18002. doi:10.1038/nrdp.2018.2
- Riancho, J. A. (2012). Genome-wide association studies (GWAS) in complex diseases: advantages and limitations. *Reumatol Clin*, 8(2), 56-57. doi:10.1016/j.reuma.2011.07.005
- Ronald, J., Rajagopalan, R., Cerrato, F., Nord, A. S., Hatsukami, T., Kohler, T., Jarvik, G. P. J. S. (2011). Genetic variation in LPAL2, LPA, and PLG predicts plasma lipoprotein (a) level and carotid artery disease risk. *42(1)*, 2-9.
- Sasayama, D., Hiraishi, A., Tatsumi, M., Kamijima, K., Ikeda, M., Umene-Nakano, W., . . . Kunugi, H. (2013). Possible association of CUX1 gene polymorphisms with antidepressant response in major depressive disorder. *Pharmacogenomics J*, 13(4), 354-358. doi:10.1038/tpj.2012.18
- Scott, D. L., Wolfe, F., & Huizinga, T. W. (2010). Rheumatoid arthritis. *Lancet*, 376(9746), 1094-1108. doi:10.1016/s0140-6736(10)60826-4
- Stahl, E. A., Raychaudhuri, S., Remmers, E. F., Xie, G., Eyre, S., Thomson, B. P.,

- Plenge, R. M. (2010). Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet*, 42(6), 508-514. doi:10.1038/ng.582
- Suzuki, A., Yamada, R., Kochi, Y., Sawada, T., Okada, Y., Matsuda, K., Yamamoto, K. (2008). Functional SNPs in CD244 increase the risk of rheumatoid arthritis in a Japanese population. *Nat Genet*, 40(10), 1224-1229. doi:10.1038/ng.205
- Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat Rev Genet*, 20(8), 467-484. doi:10.1038/s41576-019-0127-1
- Tang, C. H. (2019). Research of Pathogenesis and Novel Therapeutics in Arthritis. *Int J Mol Sci*, 20(7). doi:10.3390/ijms20071646
- Uh, H. W., Deelen, J., Beekman, M., Helmer, Q., Rivadeneira, F., Hottenga, J. J., Houwing-Duistermaat, J. J. (2012). How to deal with the early GWAS data when imputing and combining different arrays is necessary. *Eur J Hum Genet*, 20(5), 572-576. doi:10.1038/ejhg.2011.231
- Viatte, S., & Barton, A. (2017). Genetics of rheumatoid arthritis susceptibility, severity, and treatment response. *Semin Immunopathol*, 39(4), 395-408. doi:10.1007/s00281-017-0630-4
- Wagan, A. A., Mahmud, T. E., Rasheed, A., Zafar, Z. A., Rehman, A. U., & Ali, A. (2016). Cardiovascular risk score in Rheumatoid Arthritis. *Pak J Med Sci*, 32(3), 534-538. doi:10.12669/pjms.323.9779
- Wasserman, A. (2018). Rheumatoid Arthritis: Common Questions About Diagnosis and Management. *Am Fam Physician*, 97(7), 455-462.
- Wasserman, A. M. (2011). Diagnosis and management of rheumatoid arthritis. *Am Fam Physician*, 84(11), 1245-1252.
- Willems, J. M., Trompet, S., Blauw, G. J., Westendorp, R. G., & de Craen, A. J. (2010). White blood cell count and C-reactive protein are independent predictors of mortality in the oldest old. *J Gerontol A Biol Sci Med Sci*, 65(7), 764-768. doi:10.1093/gerona/glq004
- Wu, X., Yang, H. J., Jung Kim, M., Zhang, T., Qiu, J. Y., & Park, S. (2019). Association between PTPN22-1123G/C and susceptibility to rheumatoid arthritis: A systematic review and meta-analysis. *Int J Rheum Dis*, 22(5), 769-780. doi:10.1111/1756-185x.13535
- Yamamoto, K., Okada, Y., Suzuki, A., & Kochi, Y. (2015). Genetics of rheumatoid arthritis in Asia--present and future. *Nat Rev Rheumatol*, 11(6), 375-379. doi:10.1038/nrrheum.2015.7
- Yamanaka, H., Sugiyama, N., Inoue, E., Taniguchi, A., & Momohara, S. (2014). Estimates of the prevalence of and current treatment practices for rheumatoid

arthritis in Japan using reimbursement data from health insurance societies and the IORRA cohort (I). *Mod Rheumatol*, 24(1), 33-40.

doi:10.3109/14397595.2013.854059

Yarwood, A., Huizinga, T. W., & Worthington, J. (2016). The genetics of rheumatoid arthritis: risk and protection in different stages of the evolution of RA.

Rheumatology (Oxford), 55(2), 199-209. doi:10.1093/rheumatology/keu323

Zhang, K., Liu, S., Li, W., Liu, S., Li, X., Fang, Y., Ning, H. (2018). Identification of QTNs Controlling Seed Protein Content in Soybean Using Multi-Locus

Genome-Wide Association Studies. 9(1690). doi:10.3389/fpls.2018.01690

Zhang, Y. M., Jia, Z., & Dunwell, J. M. (2019). Editorial: The Applications of New Multi-Locus GWAS Methodologies in the Genetic Dissection of Complex

Traits. *Front Plant Sci*, 10, 100. doi:10.3389/fpls.2019.00100

Zhu, H., Xia, W., Mo, X. B., Lin, X., Qiu, Y. H., Yi, N. J., Lei, S. F. (2016).

Gene-Based Genome-Wide Association Analysis in European and Asian Populations Identified Novel Genes for Rheumatoid Arthritis. *PLoS ONE*,

11(11), e0167212. doi:10.1371/journal.pone.0167212

臺灣人體生物資料庫. (2020). 資料庫進展. Retrieved from

https://www.twbiobank.org.tw/new_web/about-development.php



	score
Target population (Who should be tested?): Patients who	
1. have at least 1 joint with definite clinical synovitis (swelling)*	
2. with the synovitis not better explained by another disease†	
Classification criteria for RA (score-based algorithm: add score of categories A-D; A score of $\geq 6/10$ is needed for classification of a patient as having definite RA)‡	
A. Joint involvement§	
1 large joint¶	0
2-10 large joints	1
1-3 small joints (with or without involvement of large joints)#	2
4-10 small joints (with or without involvement of large joints)	3
>10 joints (at least 1 small joint)**	5
B. Serology (at least 1 test result is needed for classification) ††	
Negative RF <i>and</i> negative ACPA	0
Low-positive RF <i>or</i> low-positive ACPA	2
High-positive RF <i>or</i> high-positive ACPA	3
C. Acute-phase reactants (at least 1 test result is needed for classification) †††	
Normal CRP <i>and</i> normal ESR	0
Abnormal CRP <i>or</i> abnormal ESR	1
D. Duration of symptoms§§	
<6 weeks	0
≥ 6 weeks	1

* The criteria are aimed at classification of newly presenting patients. In addition, patients with erosive disease typical of rheumatoid arthritis (RA) with a history compatible with prior fulfillment of the 2010 criteria should be classified as having RA. Patients with longstanding disease, including those whose disease is inactive (with or without treatment) who, based on retrospectively available data, have previously fulfilled the 2010 criteria should be classified as having RA.

† Differential diagnoses vary among patients with different presentations, but may include conditions such as systemic lupus erythematosus, psoriatic arthritis, and gout. If it is unclear about the relevant differential diagnoses to consider, an expert rheumatologist should be consulted.

‡ Although patients with a score of $<6/10$ are not classifiable as having RA, their status can be reassessed and the criteria might be fulfilled cumulatively over time.

§ Joint involvement refers to any *swollen* or *tender* joint on examination, which may be confirmed by imaging evidence of synovitis. Distal interphalangeal joints, first carpometacarpal joints, and first metatarsophalangeal joints are *excluded from assessment*. Categories of joint distribution are classified according to the location and number of involved joints, with placement into the highest category possible based on the pattern of joint involvement.

¶ "Large joints" refers to shoulders, elbows, hips, knees, and ankles.

"Small joints" refers to the metacarpophalangeal joints, proximal interphalangeal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints, and wrists.

** In this category, at least 1 of the involved joints must be a small joint; the other joints can include any combination of large and additional small joints, as well as other joints not specifically listed elsewhere (e.g., temporomandibular, acromioclavicular, sternoclavicular, etc.).

†† Negative refers to IU values that are less than or equal to the upper limit of normal (ULN) for the laboratory and assay; low-positive refers to IU values that are higher than the ULN but ≤ 3 times the ULN for the laboratory and assay; high-positive refers to IU values that are > 3 times the ULN for the laboratory and assay. Where rheumatoid factor (RF) information is only available as positive or negative, a positive result should be scored as low-positive for RF. ACPA = anti-citrullinated protein antibody.

††† Normal/abnormal is determined by local laboratory standards. CRP = C-reactive protein; ESR = erythrocyte sedimentation rate.

§§ Duration of symptoms refers to patient self-report of the duration of signs or symptoms of synovitis (e.g., pain, swelling, tenderness) of joints that are clinically involved at the time of assessment, regardless of treatment status.