

Identification of susceptibility loci for Takayasu arteritis through a large multi-ancestral genome-wide association study

Lourdes Ortiz-Fernandez¹, Güher Saruhan-Direskeneli², Fatma Alibaz-Oner³, Sema Kaymaz-Tahra³, Patrick Coit^{1,4}, Xiufang Kong⁵, Allan P Kiprianos⁶, Robert T Maughan⁶, Sibel Z. Aydin^{3,7}, Kenan Aksu⁸, Gokhan Keser⁸, Sevil Kamali⁹, Murat Inanc⁹, Jason Springer¹⁰, Servet Akar^{11,12}, Fatos Onen¹², Nurullah Akkoc^{12,13}, Nader A. Khalidi¹⁴, Curry Koenig¹⁵, Omer Karadag¹⁶, Sedat Kiraz¹⁶, Lindsay Forbess¹⁷, Carol A. Langford¹⁸, Carol A. McAlear¹⁹, Zeynep Ozbalkan^{20,21}, Sule Yavuz²², Gozde Yildirim Çetin²³, Nilufer Alpay-Kanitez²⁴, Sharon Chung²⁵, Askin Ates^{21,26}, Yasar Karaaslan²¹, Kathleen McKinnon-Maksimowicz^{27,28}, Paul A. Monach²⁹, Hüseyin TE Ozer³⁰, Emire Seyahi³¹, Izzet Fresko³¹, Ayse Cefle³², Philip Seo³³, Kenneth J. Warrington³⁴, Mehmet A. Ozturk³⁵, Steven R. Ytterberg³⁴, Veli Cobankara³⁶, Ahmet Mesut Onat^{37,38}, Nurşen Duzgun²⁶, Muge Bıkcıgil³⁹, Sibel P. Yentür², Lindsay Lally⁴⁰, Angelo A. Manfredi⁴¹, Elena Baldissera⁴¹, Eren Erken³⁰, Ayten Yazici³², Bünyamin Kısacık^{37,38}, Timuçin Kaşifoğlu⁴², Ediz Dalkilic⁴³, David Cuthbertson⁴⁴, Christian Pagnoux⁴⁵, Antoine Sreih¹⁹, Guillermo Reales^{46,47}, Chris Wallace^{46,47,48}, Jonathan D. Wren^{49,50}, Deborah S. Cunninghame-Graham⁵¹, Timothy J. Vyse⁵¹, Ying Sun⁵, Huiyong Chen⁵, Peter C Grayson⁵², Enrico Tombetti^{53,54}, Lindi Jiang^{5,55}, Justin C. Mason⁶, Peter A. Merkel^{19,56}, Haner Direskeneli³, Amr H. Sawalha^{1,27,57}

¹ Division of Rheumatology, Department of Pediatrics, University of Pittsburgh, Pittsburgh 15224, PA, USA

² Department of Physiology, Istanbul University, Istanbul Faculty of Medicine, Istanbul 34093, Turkey

³ Department of Internal Medicine, Division of Rheumatology, Marmara University, Faculty of Medicine, Istanbul 34890, Turkey

⁴ Graduate Program in Immunology, University of Michigan, Ann Arbor, MI 48109, USA

⁵ Department of Rheumatology, Zhongshan Hospital, Fudan University, Shanghai, 200032, China

⁶ National Heart and Lung Institute, Imperial College London, London SW3 6LY, UK

⁷ Faculty of Medicine, Division of Rheumatology, the Ottawa Hospital Research Institute, University of Ottawa, ON K1Y 4E9, Canada

⁸ Department of Internal Medicine, Division of Rheumatology, Ege University, Faculty of Medicine, Izmir 35100, Turkey

⁹ Department of Internal Medicine, Division of Rheumatology, Istanbul University, Istanbul Faculty of Medicine, Istanbul 34903, Turkey

- ¹⁰ Division of Allergy, Clinical Immunology and Rheumatology, Department of Medicine, University of Kansas Medical Center, Kansas City, KS 66160, USA
- ¹¹ Department of Internal Medicine, Division of Rheumatology, Katip Çelebi University, Faculty of Medicine, İzmir 35620, Turkey.
- ¹² Department of Internal Medicine, Division of Rheumatology, Dokuz Eylül University, Faculty of Medicine, İzmir 35340, Turkey
- ¹³ Department of Internal Medicine, Division of Rheumatology, Celal Bayar University, Faculty of Medicine, Manisa 45140, Turkey
- ¹⁴ Division of Rheumatology, St. Joseph's Healthcare, McMaster University, Hamilton, ON L8N 1Y2, Canada
- ¹⁵ Division of Rheumatology, University of Utah, Salt Lake City, UT 84132, USA
- ¹⁶ Department of Internal Medicine, Division of Rheumatology, Hacettepe University, Faculty of Medicine, Ankara 06100, Turkey
- ¹⁷ Division of Rheumatology, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
- ¹⁸ Department of Rheumatology, Cleveland Clinic, Cleveland, OH 44195, USA
- ¹⁹ Division of Rheumatology, University of Pennsylvania, Philadelphia, PA 19104, USA
- ²⁰ Department of Internal Medicine, Division of Rheumatology, Hitit University, Faculty of Medicine, Çorum 19200, Turkey
- ²¹ Department of Internal Medicine, Division of Rheumatology, Ankara Numune Training and Research Hospital, Ankara 06100, Turkey
- ²² Division of Rheumatology, Department of Medical Sciences, Uppsala University, Uppsala 751 85, Sweden
- ²³ Department of Internal Medicine, Division of Rheumatology, Sutcu Imam University, Faculty of Medicine, Kahramanmaraş 46050, Turkey
- ²⁴ Department of Internal Medicine, Division of Rheumatology, Koc University, Faculty of Medicine, Istanbul 34450, Turkey
- ²⁵ Russell/Engleman Rheumatology Research Center, Department of Medicine, University of California San Francisco, San Francisco, CA 94143, USA
- ²⁶ Department of Internal Medicine, Division of Rheumatology, Ankara University, Faculty of Medicine, Ankara 06100, Turkey
- ²⁷ Division of Rheumatology and Clinical Immunology, Department of Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA
- ²⁸ Division of Rheumatology, Henry Ford Health System, Detroit, MI 48202, USA
- ²⁹ Section of Rheumatology, Boston University School of Medicine, Boston, MA 02118, USA

- ³⁰ Department of Internal Medicine, Division of Rheumatology, Cukurova University, Faculty of Medicine, Adana 01330, Turkey
- ³¹ Department of Internal Medicine, Division of Rheumatology, Istanbul University-Cerrahpaşa, Faculty of Medicine, Istanbul 34098, Turkey
- ³² Department of Internal Medicine, Division of Rheumatology, Kocaeli University, Faculty of Medicine, Kocaeli 41380, Turkey
- ³³ Division of Rheumatology, Johns Hopkins University, Baltimore, MD 21224, USA
- ³⁴ Division of Rheumatology, Mayo Clinic College of Medicine, Rochester, MN 55905, USA
- ³⁵ Department of Internal Medicine, Division of Rheumatology, Faculty of Medicine, Gazi University, Ankara 06500, , Turkey
- ³⁶ Department of Internal Medicine, Division of Rheumatology, Pamukkale University, Faculty of Medicine, Denizli 20160, Turkey
- ³⁷ Division of Rheumatology, Gaziantep Medical Park Hospital, Gaziantep 27060, Turkey
- ³⁸ Department of Internal Medicine, Division of Rheumatology, Gaziantep University, Faculty of Medicine, Gaziantep 27310, Turkey
- ³⁹ Department of Internal Medicine, Division of Rheumatology, Yeditepe University, Faculty of Medicine, Istanbul 34752, Turkey
- ⁴⁰ Department of Medicine, Division of Rheumatology, Hospital for Special Surgery, New York, NY 10021, USA
- ⁴¹ Division of Immunology, Transplantation and Infectious diseases, San Raffaele Hospital, Milan 20132, Italy
- ⁴² Department of Internal Medicine, Division of Rheumatology, Osman Gazi University, Faculty of Medicine, Eskişehir 26030, Turkey
- ⁴³ Department of Internal Medicine, Division of Rheumatology, Uludag University, Faculty of Medicine, Bursa 16059, Turkey
- ⁴⁴ Department of Biostatistics, University of South Florida, Tampa, FL 33612, USA
- ⁴⁵ Division of Rheumatology, Mount Sinai Hospital, Toronto, ON M5L 3L9, USA
- ⁴⁶ Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, University of Cambridge, Puddicombe Way, Cambridge CB2 0AW, UK
- ⁴⁷ Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge, CB2 0QQ,UK
- ⁴⁸ MRC Biostatistics Unit, University of Cambridge, Cambridge Institute of Public Health, Forvie Site, Robinson Way, Cambridge Biomedical Campus, Cambridge, CB2 0SR, UK
- ⁴⁹ Genes & Human Disease Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA

⁵⁰ Department of Biochemistry and Molecular Biology, The University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

⁵¹ Department of Medical and Molecular Genetics, King's College London, London WC2R 2LS, UK

⁵² Systemic Autoimmunity Branch, NIAMS, National Institutes of Health, Bethesda, MD 20892, USA

⁵³ Department of Biomedical and Clinical Sciences L. Sacco, Milan University, Milan 20157, Italy

⁵⁴ Division of Immunology, Transplantation and Infectious diseases, San Raffaele Hospital, Milan 20132, Italy

⁵⁵ Evidence-Based Medicine Center, Fudan University, Shanghai 200032, China

⁵⁶ Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA 19104, USA

⁵⁷ Lupus Center of Excellence, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, USA

Please address correspondence to Amr H. Sawalha, MD. Address: 7123 Rangos Research Center, 4401 Penn Avenue, Pittsburgh, PA 15224, USA. Phone: (412) 692-8140. Fax: (412) 412-692-5054. Email: asawalha@pitt.edu

Key words: Takayasu arteritis, GWAS, vasculitis, genetic association, eQTL, epigenetic

Running Title: Multi-ancestral GWAS identifies genetic susceptibility in Takayasu arteritis

Abstract

Takayasu arteritis is a rare inflammatory disease of large arteries. We performed a genetic study in Takayasu arteritis comprising 6,670 individuals (1,226 cases) from five different populations. We revealed previously unreported *HLA* risk factors and four novel *non-HLA* susceptibility loci in *VPS8*, *SVEP1*, *CFL2*, and chr13q21, and reinforced *IL12B*, *PTPK2B*, and chr21q22 as robust susceptibility loci shared across ancestries. Functional analysis proposed plausible underlying disease mechanisms and pinpointed *ETS2* as a potential causal gene for chr21q22 association. We also identified >60 candidate loci with suggestive association ($p < 5 \times 10^{-5}$) and devised a genetic risk score for Takayasu arteritis. Takayasu arteritis was compared to hundreds of other traits, revealing the closest genetic relatedness to inflammatory bowel disease. Epigenetic patterns within risk loci suggest roles for monocytes and B cells in Takayasu arteritis. This work enhances understanding of the genetic basis and pathophysiology of Takayasu arteritis and provides clues for potential new therapeutic targets.

Introduction

Takayasu arteritis is a rare inflammatory disease affecting large arteries, most frequently the aorta and its major branches. Clinical manifestations result from vascular complications such as stenoses, occlusions and aneurysms, and may also include non-specific symptoms such as fatigue, fever, or weight loss ^{1;2}. Disease onset typically occurs in the second or third decade of life and the disease is more common in women. Takayasu arteritis has been reported worldwide, but disease prevalence varies across populations being the highest in East Asia ³. Although the etiology of Takayasu arteritis is incompletely understood, there is strong evidence for a role of genetic factors in the disease pathophysiology ⁴.

The first suggestion of a genetic component to Takayasu arteritis arose in the 1960's with the publication of a case report of familial aggregation of the disease ⁵. Later in the 1970's, the human leukocyte antigen (HLA) was the first genetic region identified as a susceptibility factor. The association with *HLA-B*52* allele remains the most robust genetic signal in Takayasu arteritis ⁴. In the last decade, four large-scale genetic analyses have been conducted, but only nine *non-HLA* genetic loci have been revealed with a genome-wide level of significance ⁶⁻⁹. Despite the recent progress achieved, our knowledge of the genetic component of Takayasu arteritis remains significantly limited compared to other immune-mediated diseases. In this regard, two main limitations can be highlighted. The low prevalence of this form of vasculitis makes the collection of large cohorts challenging, thereby limiting the statistical power needed to uncover the genetic basis of the disease. Secondly, even though Takayasu arteritis has been described in all ancestries, large genetic studies have only analyzed individuals from Turkish, European-American, and Japanese populations ⁷⁻⁹. Awareness of the importance of including diverse and under-represented ancestry populations in genetic studies is growing and is positioned high on the priorities of genomic research.

We performed a large genome-wide association study (GWAS) in Takayasu arteritis, using cases and controls from five diverse ancestries. We identified novel (i.e. previously unreported) genetic susceptibility loci for the disease. Functional and

epigenetic enrichment analyses provided insights into the pathogenesis of Takayasu arteritis and revealed potential new therapeutic targets.

Methods

Study population and genotyping

The total study population comprised 1,226 individuals diagnosed with Takayasu arteritis from five populations (Turkish, Northern European descendant- from North America and the United Kingdom, Han Chinese, South Asian, and Italian) and 5,444 ancestry-matched unaffected individuals. All cases fulfilled the American College of Rheumatology 1990 classification criteria for Takayasu arteritis ¹⁰. For 591 individuals diagnosed with Takayasu arteritis, genotyping data were generated using Infinium Global Screening Array-24 v2.0 and Infinium OmniZhongHua-8 v1.3 (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. For the remaining cases and controls included in the analyses, the genotyping data were obtained from a previously published GWAS ⁶ and from the database of Genotypes and Phenotypes (dbGAP) and the 1000 Genome Project ^{11; 12}. A detailed description of the study population and the genotyping platforms used for each cohort can be found in **(Table S1)**. The study was approved by the institutional review boards and the ethics committees at all participating institutions, and written informed consent was obtained from all study participants.

Quality controls and imputation

Unified stringent quality control (QC) of the data was conducted for each cohort separately before imputation using PLINK v.1.9 ¹³. Samples were removed if they had a genotyping call rate <95%. In addition, single nucleotide polymorphisms (SNPs) with a genotyping call rate <98%, minor allele frequencies (MAF) <1% and deviating from Hardy-Weinberg equilibrium in cases and controls (HWE; p value < 1×10^{-3}) were filtered out. Sex chromosomes were not analyzed. In addition, duplicated or multi-allelic SNPs as well as those A/T-C/G SNPs with MAF > 0.4 were removed to avoid errors in the imputation process. Combining QC-filtered genotyping datasets in each population was carried out before imputation unless the overlapping SNPs among different genotyping arrays was too low for imputation. In that case, imputation of each dataset

was performed before merging. Imputation of genetic data was conducted with the Michigan Imputation Server using Minimac3¹⁴ and the Haplotype Reference Consortium (HRC) r1.1 as reference population for all datasets¹⁵. The software SHAPEIT¹⁶ was used for haplotype reconstruction. To control for potential imputation errors as different genotyping platforms were used, only SNPs with stringent correlation values ($R^2 > 0.9$) were used for further analyses. Subsequently, imputed data were subjected to additional QC: MAF < 1% and HWE p value < 1×10^{-3} . After merging different datasets, we used PLINK to detect strand consistency, and any SNP showing inconsistency was discarded from further analysis. To avoid the presence of first-degree relatives and/or duplicate subjects, relatedness was evaluated using Genome function in PLINK v.1.9¹³ and one individual from each pair was removed ($\text{Pi-HAT} > 0.4$). Principal component analysis (PCA) was assessed to control for possible population stratification, Eigensoft 6.1.4 software was used to calculate the principal components for each population independently using around 100,000 independent SNPs¹⁷. Individuals found >6 standard deviations from the cluster centroids were considered outliers and were removed. The corresponding plots were produced using R 3.6 software¹⁸ (**Figure S1**).

Classical *HLA* alleles imputation

Imputation of classical *HLA* alleles was performed with the HLA*IMP:03 v0.1.0 online software¹⁹ after following instructions detailed in its website. Briefly, this software used a multi-ancestral reference panel of 10,561 individuals (8,768 European, 869 Asian, 568 African American and 365 Latino) with genotyping data as well as lab-based typing alleles. A total of 6,114 post-QC phased genotyped markers within the expanded *HLA* region (chr6: 20 to 40 Mb) were used for the imputation of both *HLA* class I and class II loci (*HLA-A*, *HLA-B*, *HLA-C*, *HLA-DQA1*, *HLA-DQB1*, and *HLA-DRB1*) at two field of resolution. Only alleles with an imputation posterior probability of 0.9 were considered successfully imputed and used for the analyses.

Data analysis

First, we conducted logistic regression using the 10 first principal components as covariates for each population independently. Genomic inflation factor (λ) was calculated for each cohort and quantile-quantile (Q-Q) plots were generated with and without the HLA region. Next, the results of the logistic regression were meta-analyzed by means of the inverse variance method. A fixed-effects model was applied for those SNPs without evidence of heterogeneity, more than 90% of all SNPs, (Cochran's Q test p value (Q) > 0.1 and heterogeneity Index (I^2) < 50%), and random-effects model was applied for SNPs displaying heterogeneity of effects between studies (Q \leq 0.1 and $I^2 \geq$ 50%). Genome-wide significance was established at a p value $\leq 5 \times 10^{-8}$ and a threshold of p value $\leq 5 \times 10^{-5}$ was set for suggestive associations. To identify independent associations in the *HLA* region, dependency analyses were conducted using a stepwise strategy at the population level followed by inverse variance meta-analysis. First, the most associated variant identified in the meta-analysis within the *HLA* region was used as covariate to perform the conditional logistic regressions in each population separately. The results from the conditional logistic regressions were then meta-analyzed by the inverse variance method and the SNP showing the most significant association was considered as an independent association and used as covariate in the second round of conditional logistic regressions in each of the populations and meta-analysis. This process was repeated until no variant reached a GWAS level of significance (p value < 5×10^{-8}). In addition, using a similar approach, we also carried out pairwise conditional analysis among the independent SNPs identified and the classical *HLA* alleles found associated with Takayasu arteritis (p value < 5×10^{-8}). PLINK v.1.9¹³ was used for the analyses and qqman R package was used to generate the Manhattan and Q-Q plots.

Functional annotation

To evaluate the potential causality of the identified associated variants we performed a detailed functional annotation by using a combination of *in silico* tools and public datasets. HaploReg v4.1 was used to perform epigenomic annotation²⁰. To explore expression quantitative trait loci (eQTL) across different human tissues, we used

several public databases: Genotype-Tissue Expression (GTEx) project ²¹, Blood NESDA NTR Conditional eQTL Catalog ²², Blood eQTL ²³, and RegulomeDB ²⁴. In addition, GTEx project was also checked to identify genetic variants regulating gene structure, named splicing QTL (sQTL). Finally, chromatin interactions between SNPs and gene promoter regions were evaluated using the webtool Capture Hi-C plotter (CHiCP) ^{25; 26}.

Enrichment analyses

We performed a functional epigenomic enrichment analysis to assess whether disease-associated *non-HLA* SNPs were over- or under- represented among regulatory elements in the genome. The GenomeRunner software ²⁷ was used to examine the SNP-list (5×10^{-8} p value cutoff was used) within a regulatory context. Briefly, 127 reference epigenomes and 7 epigenetic marks obtained from the Encyclopedia of DNA Elements (ENCODE) and the Roadmap Epigenomics Consortium projects were analyzed. H3K4me3 and H3K27me3 mark active and silenced promoters, respectively. H3K4me1 is associated with enhancers and regions downstream of transcription factor binding sites, and H3K4me2 has been related with both promoters and enhancers. H3K9me3 has been associated with repressive heterochromatic state. Finally, acetylation in H3K9 and H3K27 has been associated with transcriptional initiation and open chromatin structure. The processed gapped peak annotation panels were selected for these analyses ^{28; 29}. To control for functional enrichments occurring by chance, all common SNPs from dbSNP v142 were used as background. The enrichment p values shown were corrected for multiple testing using a False Discovery Rate (FDR) calculation (FDR p values < 0.05 were considered significant).

The identification of key pathways was performed by Gene Ontology (GO) enrichment analysis using Metascape ³⁰. Disease-associated genetic variants (p value < 5×10^{-5}) located within 50 kb of a gene were annotated and those genes were used for this analysis. The Biological Process ontology branch was selected and only terms with p value < 1×10^{-3} and at least three genes were considered. Bonferroni corrected p values (q values) are also shown. Since GO terms often overlap, the results were categorized

using a cluster-based approach based on GO similarities to provide better interpretation.

Cumulative genetic risk score

We investigated if an accumulative genetic risk score (GRS) for Takayasu arteritis varies across populations. We examined the GRS of a total of 2,504 individuals from the five major populations included in the 1000 Genome project phase 3 release: AFR, African, N=661; AMR, Admixed American, N=347; EAS, East Asian, N=504; EUR, European, N=503; SAS, South Asian, N=489 ¹². The GRS calculation was carried out following the method described by Hughes et al³¹. Briefly, the GRS of each individual is calculated by the sum of the number of risk alleles (0,1 or 2) from each SNP multiplied by the natural logarithm of its OR. The SNPs and the OR values used are listed in **Table S2**. We included the most significant genetic variant in each locus identified with a suggestive level of significance (p value $< 5 \times 10^{-5}$) in our meta-analysis, plus the variant rs12524487, which tags *HLA-B*52*, the most robust genetic association in Takayasu arteritis confirmed in every population studied ⁷. We analyzed a total of 70 genetic variants (two loci were excluded due to the lack of the associated variant information in the 1000 Genome Project data). ORs used were those obtained in the meta-analysis. To determine if the cumulate GRS was different across populations, one-way ANOVA following by Tukey's multiple comparison test was performed using GraphPad Prism version 8.1.1 (GraphPad Software, La Jolla California USA). ANOVA p values < 0.05 and Tukey's adjusted p values < 0.05 were considered significant.

Takayasu arteritis genetic relationship with other immune-mediated diseases

We projected Takayasu arteritis into a previously constructed lower-dimensional representation of immune-mediated disease (IMD) genetics, an "IMD basis", using the `project_sparse` function from the `cupcake` R package. To represent all the ancestral groups included in this study, the summary statistic of the meta-analysis was used for these analyses. Projection requires knowledge of the linkage disequilibrium (LD)

between variants, and we used an LD matrix estimated from our GWAS samples to allow for the different ancestries included.

Takayasu arteritis was compared to: 1) IMD GWAS studies, 2) UK Biobank self-reported disease traits "round 2" version), which contains the most complete coverage of self-reported IMD, but is restricted to the white European subset of participants, 3) UK Biobank ICD code traits from GeneAtlas ³² which does not restrict by ancestry, though most genotyped participants included in the UK Biobank were also White European ³³, and 4) GWAS summary data from FinnGen, and restricted to traits starting with the codes D-S, as we considered this subset to represent non-infectious, non-cancer diseases.

Classes 1-3 were available in supplementary tables of ³⁴, while GWAS summary data from FinnGen was accessed from its website. Classes 1-2 were used to generate the hierarchical clustering, while classes 3-4 were used for comparisons based on Mahalanobis distance. Mahalanobis distance was calculated as $(P_j - P_{TAK})' S^{-1} (P_j - P_{TAK})$ where P_j and P_{TAK} represent the projected vectors of trait j and TAK respectively, and S the correlation between components calculated from the European LD reference matrix. We confirmed that results were visually indistinguishable if we used the LD matrix calculated from the samples in this study.

Results

Multi-ancestral meta-analysis

We report a cross-ancestry meta-analysis comprised of individuals from five different populations: Turkish, Northern European descendant, Han Chinese, South Asian, and Italian. From the original 6,670 individuals (1,226 cases) a total of 6,221 individuals (1,091 cases) and around 4 million variants were maintained after stringent QC filtering (a summary of sample/variant QC is shown in **Table S3**). First, we performed logistic regression analyses in each population independently (**Figure S2**) followed by a meta-analysis including all cohorts. Results of genomic control analysis showed no evidence of population stratification for any cohort (**Figure S3 and Table S3**). The results of the meta-analysis are illustrated in **Figure 1**. As expected, the most robust association signal was observed within the *HLA* region.

To further evaluate the *HLA* association with Takayasu arteritis we first conducted a stepwise conditional analysis of more than 60,000 SNPs located in the extended *HLA* region (chr6: 20 to 40 Mb). As shown in **Figure 2**, five SNPs were identified as independent associations with Takayasu arteritis (p value $< 5 \times 10^{-8}$). The strongest effect corresponds with several polymorphisms in complete LD located in the 5' of *MUC21* (index SNP: rs2844678, p value = 5.70×10^{-47} ; OR = 2.71). In addition, two intergenic SNPs located between *HLA-B* and *MICA* (rs12524487, p value = 5.72×10^{-38} ; OR = 3.12, and rs17193507, p value = 7.08×10^{-30} ; OR = 3.82), the rs12526858 within the *HLA-B* (p value = 1.55×10^{-10} ; OR = 2.70) and the rs28749167 located in the non-classical *HLA-G* gene (p value = 1.90×10^{-09} ; OR = 1.43) were identified as independent risk factors. Detailed results are displayed in **Table 1**.

Next, we imputed the *HLA* classical alleles to determine their association with Takayasu arteritis. Consistent with the previous knowledge, the *HLA-B*52:01* (p value = 4.16×10^{-36} ; OR = 4.37) showed the strongest association among all markers. The results of the conditional logistic regressions detected independent associations surpassing the genome-wide level of significance of other class I and class II *HLA* alleles: *HLA-B*13:02*, p value = 2.96×10^{-25} ; OR = 3.21; *HLA-B*15:01*, p value = 1.09×10^{-15} ; OR = 3.07; *HLA-DQB1*05:02*, p value = 3.88×10^{-08} ; OR = 1.90. To determine the relationship

between the *HLA* variants and the classical alleles, we performed conditional regression analysis (**Table S4**). We can observe that some signals, even though remained significant, showed an attenuated effect after adjusting for other markers as is the case of *HLA-B*5201*-rs12524487, *HLA-B*15:01*-rs12527787, and *HLA-B*15:01*-rs12526858. These results suggest that we cannot rule out an independent effect for these markers in Takayasu arteritis. In contrast, the association of rs17193507 was completely abrogated by conditioning on *HLA-B*15:01* which indicates that the association of this variant relies on the association of the classical allele.

Finally, we also evaluated each population independently (**Table S5**). Interestingly, some SNPs located in other non-classical HLA genes were identified associated at a genome-wide level of significance in specific populations. These include a variant located in the *PSORS1C1* in Turkish population (rs61444472, p value= 2.37×10^{-17}), a polymorphism within the *HLA-V* in the Northern European descendant population (rs45479291, p value= 1.92×10^{-16}), and rs1264697 (p value= 1.33×10^{-08}) located near the *TRIM31* gene in Chinese population.

Outside of the *HLA*, 41 SNPs shared between at least two independent populations surpassed the genome-wide level of significance (p value < 5×10^{-8}), showing no odds ratio heterogeneity among populations (**Table S6**). These variants were located within 7 genomic regions, of which 4 corresponded with previously unreported Takayasu arteritis-associated loci. The most significant variants within each genomic region are shown in **Table 2**.

We detected evidence of association with a genetic variant located upstream of *SVEP1*, sushi Von Willebrand Factor Type A, EGF And Pentraxin Domain Containing 1, (rs7038415, p value= 4.90×10^{-8} ; OR=1.45) and a variant upstream of *CFL2*, cofilin 2, (rs76457959, p value= 2.84×10^{-9} ; OR=1.62). Notably, these genes have been associated with coronary disease and with atrial fibrillation, respectively³⁵⁻³⁸. We also identified a genetic association signal within *VPS8*, VPS8 subunit of CORVET complex, (rs58693904, p value= 3.73×10^{-9} ; OR=1.51) which encodes a protein involved in vesicular transport³⁹ and with an intergenic variant located in chr13q21 genomic region

(rs9540128, p value= 1.89×10^{-8} ; OR=2.02). None of these variants showed high LD with any other genotyped or imputed SNPs (**Figure S4**).

Along with the previously unreported identified genetic susceptibility loci for Takayasu arteritis, our findings also reinforced the association of three previously reported loci. The strongest effect after the *HLA* region corresponded with chr21q22 (rs4817983, p value= 1.34×10^{-12} ; OR= 0.64). This intergenic region was first described in Turkish and European-American populations and was replicated later in Japanese^{6,9}. We confirmed and extended this genetic association to Takayasu arteritis in Han Chinese and Italian cohorts (p value= 1.43×10^{-2} ; OR= 0.51 and p value= 2.98×10^{-4} ; OR= 0.33, respectively). Although the results for the South Asian cohort were not statistically significant, a consistent odds ratio direction was detected. Furthermore, our results validated the association of *IL12B*, interleukin 12B, (rs4379175, p value= 3.13×10^{-9} ; OR= 1.36) and *PTK2B*, protein tyrosine kinase 2 beta, (rs28834970, p value= 1.40×10^{-9} ; OR= 0.71) with homogeneity across all 5 cohorts. A nominal association was observed in Han Chinese population for both genes. These results further establish the association of these loci with Takayasu arteritis and provide evidence of a common genetic background among different ancestral groups. Finally, the high LD structure among the identified GWAS-level SNPs restrained the localization to a single genetic variant in any of these three genomic regions (Regional plots are shown in **Figure S4**).

In addition, 308 SNPs corresponding to 66 genomic regions (including *PTK2B* and chr21q22 loci) showed suggestive evidence of association with Takayasu arteritis (p value $\leq 5 \times 10^{-5}$) (**Table S7**). It should be noted that some of these loci have been related to other immune-mediated diseases such as *PLCG2* (phospholipase C gamma 2) with Inflammatory bowel disease, *RBPJ* (recombination signal binding protein for immunoglobulin kappa J region) with rheumatoid arthritis or *ZMIZ1* (zinc finger MIZ-type containing 1) with Crohn's disease among others⁴⁰⁻⁴². Interestingly, a cross-phenotype analysis of systemic vasculitis suggested *ZMIZ1* as a shared risk factor⁴³. Other promising candidate genes are *CCR7* (C-C motif chemokine receptor 7) and *RGS6* (regulator of G protein signaling 6).

Finally, we checked the known GWAS-level variants associated with Takayasu arteritis in our meta-analysis. Detailed data are shown in **Table S8**. Briefly, the variants identified in *KLHL33* and *LILRB3* were not included in our meta-analysis, due to inability to impute or low-quality imputation, respectively. The results of our meta-analysis showed significant associations in *FCGR2A/FCGR3A* and *DUSP22* variants (rs10919543 (p value = 2.68×10^{-3} , OR= 1.45) and rs17133698 (p value = 4.89×10^{-3} , OR= 0.64), respectively), consistent with previous studies. It should be noted that the variant in *FCGR2A/FCGR3A* was not evaluated in the Turkish and Northern European descendant population in this study and, similarly, the SNP in *DUSP22* was filtered out (MAF <1%) in our Northern European descendant and Italian populations. We detected an association in the SNP reported in *IL6*, rs2069837 (p value = 2.05×10^{-2} , OR= 0.84), but interestingly, we find that odds ratios directions in this locus are different across populations. Finally, the variant in *LILRA3*, rs103294, reported previously in a Japanese cohort showed no tendency of association in the meta-analysis (this variant only passed the QC in Turkish and Italian populations), but was found to be associated in our Turkish cohort (p value = 9.36×10^{-9} , OR= 1.66).

Functional annotations and epigenetic enrichment analysis of genome-wide associated SNPs

In line with most of genetic susceptibility loci detected in immune-mediated diseases, Takayasu arteritis-associated genetic variants reside in non-coding regions suggesting that they might influence the disease by disrupting regulatory elements. To assign biological significance to those variants, we carried out a comprehensive functional annotation using *in silico* approaches.

First, we evaluated the regulatory potential of disease-associated variants reported with a GWAS level of significance in our study (p value < 5×10^{-8}). Epigenetic annotation indicates that most of these variants co-localize with epigenetic marks, and therefore, likely contribute to chromatin state across multiple primary tissues and cells (**Table S9**). We found that 85% of variants may alter regulatory motifs and therefore, might disrupt transcription factor (TF) binding. DNase hypersensitivity sites were also observed in

44% of the variants analyzed. Finally, protein binding was identified in some of these variants. For example, in lymphoblastoid cell lines, a Takayasu arteritis-associated genetic variant in *PTPK2B*, rs755951, is located in a binding site for CTCF, and binding of PU.1 was detected in at least two disease-associated variants in chr21q22.

As most of the associated variants co-localize with epigenetic features, we performed a regulatory enrichment analysis to evaluate whether the presence of epigenetic marks at these loci was more frequent than expected by chance, and to determine the specific cell subsets that are more likely to be affected by these genetic variants. We analyzed seven different histone marks (H3K4me1, H3K4me2, H3K4me3, H3K9ac, H3K27ac, H3K27me3 and H3K9me3) across 127 reference epigenomes. Of 889 combinations of regulatory features and tissues examined, we found 170 associated enrichments in 86 cell types/tissues (**Table S10**). Most epigenetic mark enrichment patterns were observed in blood cells which includes a wide repertoire of immune cells. In agreement with our current understanding of the immunopathology of Takayasu arteritis, the most significant enrichments were observed in monocytes (**Figure 3**). Our results showed enrichment of H3K27ac (p value= 2.58×10^{-51}), H3K4me1 (p value= 5.43×10^{-37}), H3K9ac (p value= 3.80×10^{-21}), H3Kme2 (p value= 2.22×10^{-26}), H3K4me3 (p value = 1.03×10^{-32}) in primary human monocytes. Three specific B cells lines presented strong enrichments as well, including primary B cells from cord blood: H3K4me1 (p value= 5.43×10^{-37}) and H3K4me3 (p value= 1.20×10^{-32}), primary B cells from peripheral blood: H3K27ac (p value= 5.22×10^{-15}) and H3K4me1 (p value= 3.02×10^{-17}) and GM12878 lymphoblastoid cells: H3K27ac (p value= 2.59×10^{-6}), H3K9ac (p value= 7.61×10^{-6}), H3K9me3 (p value= 2.02×10^{-7}), H3K4me2 (p value= 6.68×10^{-6}), H3K4me1 (p value= 3.22×10^{-12}) and H3K4me3 (p value= 4.89×10^{-2}). Other robust epigenetic mark co-localizations were also detected in other immune cells, such as neutrophils, natural killer cells, and T cell subsets, and in other tissues such as the heart and digestive tissues.

Given that epigenetic marks may correlate with gene expression changes we examined several public repositories of eQTLs to explore the relationship between gene expression and genetic variants associated with Takayasu arteritis. We found that 78% of variants detected with a GWAS level of significance in our study have been linked to

altered expression of at least one gene in multiple tissues and/or cell types **Table S11**). Takayasu arteritis-associated variants in *PTK2B* are associated with expression changes of several genes in whole blood, lymphoblastoid cell lines, monocytes, and left ventricular heart tissue, among others. Notably, the disease risk alleles in *PTK2B* are associated with reduced expression of *PTK2B*. Furthermore, these polymorphisms also alter the expression levels of *PTK2B* isoforms, acting as sQTL. In addition, Takayasu arteritis susceptibility variants in *PTK2B* are associated with altered expression levels of other genes such as *DPYSL2* (dihydropyrimidinase like 2) in lymphoblastoid cell lines, *EPHX2* (epoxide hydrolase 2) in tibial artery tissue, and *CHRNA2* (Cholinergic Receptor Nicotinic Alpha 2 Subunit) and *TRIM35* (tripartite motif containing 35) in whole blood. Takayasu arteritis-associated variants upstream of *IL12B* change the expression of *UBLCP1* (ubiquitin like domain containing CTD phosphatase 1) in whole blood, and those in the intergenic region chr21q22 alter the expression of several genes including *ETS2* (ETS proto-oncogene 2, transcription factor) and *LCA5L* (lebercilin LCA5 like).

Under the assumption that chromatin architecture may influence transcriptional regulation, we used the Capture HiC Plotter webtool to detect and visualize physical chromatin interactions between Takayasu arteritis-associated variants and gene promoter regions across multiple cell types ²⁵. The vast majority of examined variants showed multiple interactions in different cell types and several of them interact with the promoters of genes whose expression levels were affected (**Table S11 and Figure S5**). It is worth highlighting the physical interaction between variants in the chr21q22 locus and *ETS2* promoter in multiple immune cells such as monocytes, macrophages, and neutrophils. The functional relevance of this interaction is supported by the above mentioned eQTL results. Altogether, these findings pinpointed *ETS2* as a potential target gene for Takayasu arteritis influenced by the disease susceptibility locus in chr21q22 (**Figure 4**).

Pathway analysis

To highlight biological processes that might be relevant in the pathogenesis of Takayasu arteritis, we performed a Gene Ontology term enrichment analysis using all genes annotated to genetic variants with at least a suggestive level of association with

Takayasu arteritis in our study (p value $< 5 \times 10^{-5}$). Given that ontologies are often redundant, the results from these analyses commonly reveal related and even overlapping terms. To simplify their interpretation, GO terms were clustered by their similarities revealing seven different clusters: lymphocyte differentiation (p value = 3.23×10^{-6}), blood vessel morphogenesis (p value = 2.80×10^{-5}), heart morphogenesis (p value = 1.07×10^{-4}), regulation of ion transmembrane transport (p value = 2.30×10^{-4}), cell recognition (p value = 6.38×10^{-4}), peptidyl-tyrosine phosphorylation (p value = 4.29×10^{-3}), and regulation of cysteine-type endopeptidase activity involved in apoptotic process (p value = 7.05×10^{-3}) (**Figure 5**). A detailed list of all GO terms showing included genes is displayed in **Table S12**.

Cumulative Genetic Risk Score (GRS)

To investigate if the diversity of Takayasu arteritis prevalence among ancestries may have genetic bases, we assessed the cumulative GRS of more than 2,500 individuals from the five major populations and tested if mean GRS differences across populations exist. The GRS values are illustrated in **Figure 6**. There was a significant difference between the genetic risk scores among populations (ANOVA p value < 0.0001). The highest GRS were observed in African and East Asian populations, while European and South Asian showed the lowest values. The multiple comparison test on mean GRS between each pair of populations revealed statistically significant differences for all pairs except African versus East Asian and European versus South Asian (Tukey's adjusted p value = 0.924 and 0.457 respectively).

Takayasu arteritis genetic relationship with multiple immune-mediated diseases

Relating different diseases through their GWAS summary statistics has been used to reveal etiological relationships between diseases. We chose to compare Takayasu arteritis to other immune-mediated diseases through projecting it into a previously constructed lower-dimensional representation of IMD genetics, an “IMD basis”⁴⁴. This representation uses 13 components to summarize axes of risk shared between different

combinations of diseases. We found Takayasu arteritis differed significantly ($\text{FDR} < 1\%$) from control on eight of the 13 components, five of which were also associated with Crohn's disease and/or ulcerative colitis. In the original report of the IMD basis, relationships between different IMD GWAS and self-reported disease traits in European UK Biobank subjects were inferred from hierarchical clustering of their projections, and we reclustered the same set of traits here together with Takayasu arteritis to examine with which IMD it shared genetic risk components. We found that Takayasu arteritis clustered with Crohn's disease, ulcerative colitis, and ankylosing spondylitis (**Figure S6**). However, clustering across multiple dimensions can be sensitive to choices of algorithm, and although the method appears robust to different ancestries (because it is based on summary odds ratios which have been calculated with appropriate ancestry adjustment), we also wanted to confirm this was not an effect of comparing to European subjects. We therefore also compared Takayasu arteritis to projections of all diseases in two large cohorts (FinnGen and all UK Biobank subjects including non-Europeans) using a one-dimensional summary, the Mahalanobis distance (see Methods). This found that in both cohorts, Crohn's disease was the closest genetically to Takayasu arteritis, followed by ulcerative colitis, out of the hundreds of traits considered (**Figure S7**).

Discussion

We report the results of a large genetic association study in Takayasu arteritis using cases and controls from five diverse populations. In addition to the well-known *HLA-B*52* association, we detected several independent signals within the *HLA* region and characterized four novel genetic susceptibility non-*HLA* loci in Takayasu arteritis with a GWAS level of significance (*SVEP1*, *CFL2*, *VPS8* and chr13q21). Further, we identified >60 genetic loci with a suggestive level of association with the disease.

Consistent with our current knowledge of the disease, the strongest association signals were observed within the HLA region. Multiple studies have previously analyzed the association of this region with Takayasu arteritis confirming *HLA-B*52:01* as a robust genetic factor in several ancestries. Other HLA classical alleles (class I and class II) and variants along the extended HLA region have also been identified in specific populations. However, the complex linkage disequilibrium structure of this region and the low statistical power of most of these studies have limited the localization of the specific effects and the identification of new susceptibility loci at a GWAS level of significance⁴. Here, we report five variants showing independent associations with the disease, three are located near the well-known susceptibility locus *HLA-B* (rs12524487-*HLA-B/MICA*, rs17193507-*HLA-B/MICA* and rs12526858-*HLA-B*) and the other two variants might indicate previously unreported risk loci for Takayasu arteritis (rs2844678-*MUC21* and rs28749167-*HLA-G*). In addition, four independent HLA classical alleles were identified as susceptibility factors for Takayasu arteritis (*HLA-B*52:01*, *HLA-B*13:02*, *HLA-B*15:01*, and *HLA-DQB1*05:02*). Conditional regression analysis revealed that the association in rs17193507 is dependent on *HLA-B*15:01*. Our results also suggested genetic relationship among the *HLA-B* variants and the *HLA-B* classical alleles but independent effect on the disease etiology cannot be discarded. Although our data strongly suggest additional effects exist beyond *HLA-B*52:01* in Takayasu arteritis, further studies are needed to functionally characterize these complex effects.

Of interest, giant cell arteritis which is another large vessel vasculitis that can present with clinical and angiographic findings that are sometimes similar to Takayasu arteritis, appears to be associated with a distinctly different genetic susceptibility within the HLA

region. Unlike Takayasu arteritis, the most robust genetic association in giant cell arteritis is within the HLA class II region ⁴⁵.

The strongest association outside the HLA region correspond with *SVEP1* which encodes a cell-adhesion molecule also known as Polydom ⁴⁶. This protein acts as a ligand for integrin $\alpha 9\beta 1$, which has been recently described to play an important role in the development of autoimmune diseases including rheumatoid arthritis and multiple sclerosis ^{47; 48}. In this context, Slep1 is necessary for lymphatic vessel remodeling which is involved in immune cell trafficking and surveillance ⁴⁹. In addition, silencing *SVEP1* in endothelial cell lines was associated with increased expression of key molecules involved in endothelial cell-leukocyte adhesion, including interleukin-8 (IL-8), growth-regulated oncogene- α (GRO- α), monocyte chemoattractant protein 1 (MCP-1), and monocyte chemoattractant protein 3 (MCP-3) ⁵⁰. Furthermore, *SVEP1* polymorphisms have been also associated with cardiovascular conditions such as increased risk of coronary disease and blood pressure ^{36-38; 51}.

Cofilin 2 gene (*CFL2*), encodes an actin-binding protein that is a key regulator of actin dynamics ⁵². It has been shown that the actin cytoskeleton and its regulators play an important role in T cell-B cell interaction ⁵³. A related gene family member, cofilin, is involved in the clustering of T cell receptors during T-cell activation through actin reorganization ⁵⁴. In addition, previous studies reported associations between *CFL2* and atrial fibrillation, and mutations in this gene have been described in individuals with congenital myopathy ^{35; 55; 56}.

In contrast, *VPS8* and chr13q21 loci have not been previously associated with any related disease. *VPS8* encodes a vacuolar protein involved in vesicular-mediated protein trafficking. A protein-protein interaction analysis, using STRING webtool ⁵⁷, revealed other related vacuolar proteins that have been associated with cardiovascular disease. For example, polymorphisms in *VPS33B* are associated with hypertension and myocardial infarction in Japanese individuals ⁵⁸.

With regard to the four novel genetic association signals we identified in Takayasu arteritis, it should be noted that no additional variants in strong LD ($r^2 > 0.6$) with any of the associated variants were identified in our study, and that we did not detect clear

functional implications of the identified polymorphisms on the pathophysiology of the disease. These variants might disrupt regulatory elements in a cell type-specific context or might tag unidentified genetic variants that could be responsible for the causality of these associations. Replication and fine-mapping studies in independent populations followed by functional studies should be conducted to confirm and characterize the functional consequences of these novel genetic susceptibility loci in Takayasu arteritis.

The results of our multi-ancestral approach also extended the genetic associations of *PTK2B*, *IL12B*, and a locus on chr21q22 to other populations. Of interest, these genetic associations are examples of shared risk loci among different immune-mediated diseases. The rs17057051 polymorphism in *PTK2B* has been previously reported as a genetic susceptibility factor in inflammatory bowel disease⁵⁹. The genetic effect in this SNP, which is in strong LD ($r^2 > 0.8$) with the lead *PTK2B* variant associated with Takayasu arteritis in our study, is in the same direction in both disease conditions. The polymorphisms rs6556412 and rs4379175, upstream of *IL12B*, have been associated with inflammatory bowel disease and psoriasis, respectively⁵⁹⁻⁶¹. Interestingly, while the direction of the genetic effect in this locus is similar in Takayasu arteritis and inflammatory bowel disease, it appears that the risk allele in rs4379175-*IL12B* in Takayasu arteritis is protective against psoriasis. Finally, the genetic region chr21q22 has been also reported to be associated with ankylosing spondylitis and inflammatory bowel disease^{59; 62; 63}, and with a genetic effect in the same direction as we report in Takayasu arteritis.

PTK2B encodes a tyrosine kinase, PYK2, which is expressed in multiple cell types, including immune cells. Indeed, experiments in mice showed that deficiency in Pyk-2 leads to abnormal function of marginal zone B cells⁶⁴. In agreement with results published in a Japanese cohort⁹, we observed that the minor alleles of disease-associated SNPs in *PTK2B* are protective and are associated with increased expression of *PTK2B* in different tissues. Furthermore, these SNPs are linked with the expression of different isoforms of *PTK2B*. One causal variant, rs2322599, was proposed in the Japanese study, however, a total of ten SNPs (including rs2322599) surpassed the GWAS level of significance in our meta-analysis. No single variant could be identified as

causal given the high LD in this region and the broad epigenetic enrichment patterns involving these variants. Importantly, disease-associated variants in *PTK2B* are also associated with gene expression changes involving other genes in this locus, such as *DPYSL2*, *EPHX2*, *CHRNA2*, and *TRIM35*. In this context, we cannot rule out that these other genes might also play a role in Takayasu arteritis. Indeed, chromatin interaction was revealed between disease-associated variants in *PTK2B* and promoter regions of *EPHX2* and *CHRNA2*. Interestingly, *PTK2B* variant rs755951 is located in a CTCF binding site, suggesting the involvement of this Takayasu arteritis risk variant in long-range chromatin interactions.

Epigenomic annotation, chromatin interaction maps, and eQTL analysis also revealed the possible regulatory mechanisms underlying the genetic association between Takayasu arteritis and *IL12B*. Our data identified *UBLCP1* as a possible target gene affected by this genetic association. In whole blood, almost all Takayasu arteritis-associated variants in *IL12B* alter *UBLCP1* mRNA expression. Interestingly, *UBLCP1* has been found to be associated with increased genetic risk in psoriasis ⁶⁵.

Among findings from functional annotation analyses in the genetic risk locus in chr21q22 region, it should be highlighted that PU.1 protein binding was detected in this locus in ENCODE Chip-Seq experiments. PU.1 is a transcription factor member of the Ets family which is required for hematopoietic development ⁶⁶. Mice deficient in Pu.1 showed a lack of mature B cells and macrophages ⁶⁷. Our eQTL analysis of Takayasu arteritis risk variants in this locus revealed higher expression of *ETS2* in whole blood, and HiC data showed chromatin interaction between these variants and *ETS2* promoter in monocytes and macrophages, among other immune cell types. *ETS2* also encodes a member of the Ets family with a key role in macrophage development. Recently, experimental data confirmed the functional synergy between Ets2 and PU.1 to activate the transcription of the *bcl-x* gene which encodes Bcl-xL, a necessary protein for macrophage survival ⁶⁸. Interestingly, another member of the Ets family, *ETS1* has been identified as a susceptibility risk locus for other immune-mediated diseases such as lupus and inflammatory bowel disease ^{69; 70}.

The finding of enrichment in active chromatin epigenetic marks among the identified Takayasu arteritis risk loci was consistent with our understanding of the disease, reflecting a robust immunological signature. Specifically, monocytes and B cells presented the highest enrichment patterns. Monocytes and macrophages are thought to play a crucial role in the pathophysiology of large vessel vasculitis, including Takayasu arteritis ⁷¹. Although the role of B cells in Takayasu arteritis is less clear, a seemingly successful response to rituximab treatment in some cases, the presence of tissue infiltrating B cells in affected arteries, higher frequency of circulating plasmablasts, and the presence of anti-endothelial cell antibodies in individuals diagnosed with Takayasu arteritis points to a possible role for B cells in this disease ^{72; 73}. Our analysis also identified other important immune cell types implicated in the pathophysiology of this form of vasculitis, including T cell subsets and NK cells, the latter consistent with results from enrichment analysis in a recent Japanese report ⁹. We also revealed enrichment patterns across non-immune related tissues. In agreement with the pathophysiology of Takayasu arteritis, several heart-related tissue types were detected. The enrichments found in digestive tissues could be explained by a plausible genetic overlap between Takayasu arteritis and inflammatory bowel disease and an observed higher prevalence of inflammatory bowel disease in individuals diagnosed with Takayasu arteritis ^{74; 75}. This is also supported by our data of projecting the genetic component of multiples immune-mediated disease in which Takayasu arteritis was found to cluster with Crohn's disease, ulcerative colitis, and ankylosing spondylitis.

The most significant biological process identified among genetic risk loci detected in our study was lymphocyte differentiation, which includes key specific pathways such as leukocyte differentiation, T cell activation, inflammatory response to antigenic stimulus, cytokine production, among others. Along with the disease-associated loci in *IL12B* and *PTK2B*, other genetic loci with suggestive association are included in these pathways such as *CCR7*, *RBPJ*, *PLCG2*, *RORA*, *ZMIZ1*, *CD44*, *LMO1*, and *CARD9*. Among these genes, *PLCG2* has been associated with inflammatory bowel disease, *RBPJ* with rheumatoid arthritis, and *ZMIZ1* with Crohn's disease ⁴⁰⁻⁴². Genes in blood vessel and heart morphogenesis pathways were also observed. These include *CFL2* and *SVEP1*, and several genes with a suggestive level of association such as *EPHB4* and *PRDM16*.

which are risk loci for inflammatory bowel disease and rheumatoid arthritis, respectively^{59; 76}.

The cumulative genetic risk scores, which is a simple and intuitive approach to convert genetic data into a predictive measure of disease susceptibility, suggest that the differences in Takayasu arteritis prevalence have genetic bases. Although the GRS value of the African population was higher than expected, the data were in line with Takayasu prevalence described to date, being the highest in East Asia, Mexico, and South East Asia³. Epidemiological data of Takayasu arteritis in African populations are scarce. However, a recent review analyzing the current state of systemic vasculitis in Africa concluded that these pathologies might be under-diagnosed⁷⁷. Other plausible explanation for these results might be the existence of unidentified protective genetic effects against Takayasu arteritis in African populations. It should be noted that these results should be interpreted as estimates based on our current knowledge of the genetic component of Takayasu arteritis. The accuracy of our estimations could be limited given that the ORs of the associated signals might vary among populations. Further, additional genetic loci involved in Takayasu arteritis remain to be identified in currently studied and additional populations.

Finally, genetic studies have been successful in proposing new molecular targets for drug discovery⁷⁸. In this regard, the potential of *IL12B* as a drug target for Takayasu arteritis has been revealed by genetic studies. Furthermore, it is worth mentioning that PYK2, EST2 and the SVEP1-integrin interaction might represent promising drug targets for Takayasu arteritis. Indeed natalizumab, an anti- $\alpha 4$ integrin antibody, has already shown beneficial therapeutic effects in multiple sclerosis and Crohn's disease^{79; 80}. Leflunomide which is a PYK2 antagonist used to treat rheumatoid arthritis⁸¹, has also been investigated in other large vessel vasculitides including Takayasu arteritis and giant cell arteritis⁸²⁻⁸⁴.

In summary, we performed a large multi-ancestral genome-wide association study in Takayasu arteritis. We revealed previously unreported risk factors within the *HLA-region* and four novel *non-HLA* genetic associations with a GWAS level of significance,

extended the genetic association of previously reported risk loci to other populations, and uncovered over 60 additional candidate genetic susceptibility loci for the disease. Using functional and epigenetic enrichment characterization we identified pathways, biological processes, and target genes affected by these genetic risk loci. Our results also suggest a close genetic relationship between Takayasu arteritis and inflammatory bowel diseases. Finally, the results of this work highlight immune cell types that are crucial in mediating genetic risk in Takayasu arteritis.

Acknowledgements

This work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health grant number R01 AR070148 to AHS. The Vasculitis Clinical Research Consortium has received support from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (U54 AR057319 and U01 AR51874 04), the National Center for Research Resources (U54 RR019497), and the Office of Rare Diseases Research of the National Center for Advancing Translational Sciences. JCM, APK and RMM acknowledge support from the Imperial College, National Institute for Health Research, Biomedical Research Centre. CW and GR acknowledge support from the Wellcome Trust (WT107881) and the Medical Research Council (MC_UU_00002/4). This work was supported by the use of study data downloaded from the dbGaP web site, under dbGaP accessions phs000272.v1.p1, phs000431.v2.p1, phs000583.v1.p1, and phs000444.v1.p1.

Data and Code availability

Summary statistics of the meta-analysis are available and can be provided upon reasonable request. The code to reproduce the comparison figures of the genetic relationship among immune-mediated diseases analyses is available from <https://github.com/chr1swallace/tak-basis>. The rest of the data/code used in this study are included in the main manuscript or its supplementary material.

Declaration of Interests

The authors declare no competing interests

Web Resources

HLA*IMP:03 v0.1.0, <http://imp.science.unimelb.edu.au/hla/instructions>

Capture Hi-C plotter (CHiCP), <https://www.chicp.org/chicp/>

Cupcake R package, <https://github.com/ollyburren/cupcake>

UK Biobank, <http://www.nealelab.is/uk-biobank/>

FinnGen, <https://www.finnngen.fi>

ENCODE, <https://www.encodeproject.org/>

Roadmap Epigenomics Consortium, <http://www.roadmapepigenomics.org/>

References

1. Jennette, J.C., Falk, R.J., Bacon, P.A., Basu, N., Cid, M.C., Ferrario, F., Flores-Suarez, L.F., Gross, W.L., Guillevin, L., Hagen, E.C., et al. (2013). 2012 revised International Chapel Hill Consensus Conference Nomenclature of Vasculitides. *Arthritis Rheum* 65, 1-11.
2. de Souza, A.W., and de Carvalho, J.F. (2014). Diagnostic and classification criteria of Takayasu arteritis. *J Autoimmun* 48-49, 79-83.
3. Johnston, S.L., Lock, R.J., and Gompels, M.M. (2002). Takayasu arteritis: a review. *J Clin Pathol* 55, 481-486.
4. Renauer, P., and Sawalha, A.H. (2017). The genetics of Takayasu arteritis. *Presse Med* 46, e179-e187.
5. Hirsch, M.S., Aikat, B.K., and Basu, A.K. (1964). Takayasu's Arteritis. Report of Five Cases with Immunologic Studies. *Bull Johns Hopkins Hosp* 115, 29-64.
6. Renauer, P.A., Saruhan-Direskeneli, G., Coit, P., Adler, A., Aksu, K., Keser, G., Alibaz-Oner, F., Aydin, S.Z., Kamali, S., Inanc, M., et al. (2015). Identification of Susceptibility Loci in IL6, RPS9/LILRB3, and an Intergenic Locus on Chromosome 21q22 in Takayasu Arteritis in a Genome-Wide Association Study. *Arthritis Rheumatol* 67, 1361-1368.
7. Saruhan-Direskeneli, G., Hughes, T., Aksu, K., Keser, G., Coit, P., Aydin, S.Z., Alibaz-Oner, F., Kamali, S., Inanc, M., Carette, S., et al. (2013). Identification of multiple genetic susceptibility loci in Takayasu arteritis. *Am J Hum Genet* 93, 298-305.
8. Terao, C., Yoshifuji, H., Kimura, A., Matsumura, T., Ohmura, K., Takahashi, M., Shimizu, M., Kawaguchi, T., Chen, Z., Naruse, T.K., et al. (2013). Two susceptibility loci to Takayasu arteritis reveal a synergistic role of the IL12B and HLA-B regions in a Japanese population. *Am J Hum Genet* 93, 289-297.
9. Terao, C., Yoshifuji, H., Matsumura, T., Naruse, T.K., Ishii, T., Nakaoka, Y., Kirino, Y., Matsuo, K., Origuchi, T., Shimizu, M., et al. (2018). Genetic determinants and an epistasis of LILRA3 and HLA-B*52 in Takayasu arteritis. *Proc Natl Acad Sci U S A* 115, 13045-13050.
10. Arend, W.P., Michel, B.A., Bloch, D.A., Hunder, G.G., Calabrese, L.H., Edworthy, S.M., Fauci, A.S., Leavitt, R.Y., Lie, J.T., Lightfoot, R.W., Jr., et al. (1990). The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. *Arthritis and rheumatism* 33, 1129-1134.
11. Tryka, K.A., Hao, L., Sturcke, A., Jin, Y., Wang, Z.Y., Ziyabari, L., Lee, M., Popova, N., Sharopova, N., Kimura, M., et al. (2014). NCBI's Database of Genotypes and Phenotypes: dbGaP. *Nucleic Acids Res* 42, D975-979.
12. Genomes Project, C., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., et al. (2015). A global reference for human genetic variation. *Nature* 526, 68-74.
13. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7.
14. Das, S., Forer, L., Schonherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation service and methods. *Nat Genet* 48, 1284-1287.
15. McCarthy, S., Das, S., Kretschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 48, 1279-1283.
16. Delaneau, O., Zagury, J.F., and Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 10, 5-6.
17. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38, 904-909.

18. R Development Core Team. (2019). R: A language and environment for statistical computing. In. (Vienna, Austria, R Foundation for Statistical Computing.
19. Motyer, A., Vukcevic, D., Dilthey, A., Donnelly, P., McVean, G., and Leslie, S. (2016). Practical Use of Methods for Imputation of HLA Alleles from SNP Genotype Data. *bioRxiv*, 091009.
20. Ward, L.D., and Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 40, D930-934.
21. Consortium, G.T. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 348, 648-660.
22. Jansen, R., Hottenga, J.J., Nivard, M.G., Abdellaoui, A., Laport, B., de Geus, E.J., Wright, F.A., Penninx, B., and Boomsma, D.I. (2017). Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Hum Mol Genet* 26, 1444-1451.
23. Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., et al. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 45, 1238-1243.
24. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 22, 1790-1797.
25. Schofield, E.C., Carver, T., Achuthan, P., Freire-Pritchett, P., Spivakov, M., Todd, J.A., and Burren, O.S. (2016). CHiCP: a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets. *Bioinformatics* 32, 2511-2513.
26. Capture HiC Plotter (CHiCP). In. (
27. Dozmorov, M.G., Cara, L.R., Giles, C.B., and Wren, J.D. (2016). GenomeRunner web server: regulatory similarity and differences define the functional impact of SNP sets. *Bioinformatics* 32, 2256-2263.
28. Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317-330.
29. Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57-74.
30. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 10, 1523.
31. Hughes, T., Adler, A., Merrill, J.T., Kelly, J.A., Kaufman, K.M., Williams, A., Langefeld, C.D., Gilkeson, G.S., Sanchez, E., Martin, J., et al. (2012). Analysis of autosomal genes reveals gene-sex interactions and higher total genetic risk in men with systemic lupus erythematosus. *Ann Rheum Dis* 71, 694-699.
32. Canela-Xandri, O., Rawlik, K., and Tenesa, A. (2018). An atlas of genetic associations in UK Biobank. *Nat Genet* 50, 1593-1599.
33. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203-209.
34. Burren, O.S., Reales, G., Wong, L., Bowes, J., Lee, J.C., Barton, A., Lyons, P.A., Smith, K.G.C., Thomson, W., Kirk, P.D.W., et al. (2020). Informed dimension reduction of clinically-related genome-wide association summary data characterises cross-trait axes of genetic risk. *BioRxiv*.
35. Roselli, C., Chaffin, M.D., Weng, L.C., Aeschbacher, S., Ahlberg, G., Albert, C.M., Almgren, P., Alonso, A., Anderson, C.D., Aragam, K.G., et al. (2018). Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet* 50, 1225-1233.

36. Myocardial Infarction, G., Investigators, C.A.E.C., Stitzel, N.O., Stirrups, K.E., Masca, N.G., Erdmann, J., Ferrario, P.G., Konig, I.R., Weeke, P.E., Webb, T.R., et al. (2016). Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *N Engl J Med* 374, 1134-1144.
37. van der Harst, P., and Verweij, N. (2018). Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res* 122, 433-443.
38. Kichaev, G., Bhatia, G., Loh, P.R., Gazal, S., Burch, K., Freund, M.K., Schoech, A., Pasaniuc, B., and Price, A.L. (2019). Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet* 104, 65-75.
39. Jonker, C.T.H., Galmes, R., Veenendaal, T., Ten Brink, C., van der Welle, R.E.N., Liv, N., de Rooij, J., Peden, A.A., van der Sluijs, P., Margadant, C., et al. (2018). Vps3 and Vps8 control integrin trafficking from early to recycling endosomes and regulate integrin-dependent functions. *Nat Commun* 9, 792.
40. Banerjee, S., Biehl, A., Gadina, M., Hasni, S., and Schwartz, D.M. (2017). JAK-STAT Signaling as a Target for Inflammatory and Autoimmune Diseases: Current and Future Prospects. *Drugs* 77, 521-546.
41. Stahl, E.A., Raychaudhuri, S., Remmers, E.F., Xie, G., Eyre, S., Thomson, B.P., Li, Y., Kurreeman, F.A., Zhernakova, A., Hinks, A., et al. (2010). Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet* 42, 508-514.
42. Ellinghaus, D., Ellinghaus, E., Nair, R.P., Stuart, P.E., Esko, T., Metspalu, A., Debrus, S., Raelson, J.V., Tejasvi, T., Belouchi, M., et al. (2012). Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *Am J Hum Genet* 90, 636-647.
43. Ortiz-Fernandez, L., Carmona, F.D., Lopez-Mejias, R., Gonzalez-Escribano, M.F., Lyons, P.A., Morgan, A.W., Sawalha, A.H., Merkel, P.A., Smith, K.G.C., Gonzalez-Gay, M.A., et al. (2018). Cross-phenotype analysis of Immunochip data identifies KDM4C as a relevant locus for the development of systemic vasculitis. *Ann Rheum Dis* 77, 589-595.
44. Burren, O.S., Reales, G., Wong, L., Bowes, J., Lee, J.C., Barton, A., Lyons, P.A., Smith, K.G., Thomson, W., Kirk, P.D., et al. (2020). Informed dimension reduction of clinically-related genome-wide association summary data characterises cross-trait axes of genetic risk. *bioRxiv*, 2020.2001.2014.905869.
45. Carmona, F.D., Coit, P., Saruhan-Direskeneli, G., Hernandez-Rodriguez, J., Cid, M.C., Solans, R., Castaneda, S., Vaglio, A., Direskeneli, H., Merkel, P.A., et al. (2017). Analysis of the common genetic component of large-vessel vasculitides through a meta-Immunochip strategy. *Sci Rep* 7, 43953.
46. Gilges, D., Vinit, M.A., Callebaut, I., Coulombel, L., Cacheux, V., Romeo, P.H., and Vigon, I. (2000). Polydom: a secreted protein with pentraxin, complement control protein, epidermal growth factor and von Willebrand factor A domains. *Biochem J* 352 Pt 1, 49-59.
47. Sato-Nishiuchi, R., Nakano, I., Ozawa, A., Sato, Y., Takeichi, M., Kiyozumi, D., Yamazaki, K., Yasunaga, T., Futaki, S., and Sekiguchi, K. (2012). Polydom/SVEP1 is a ligand for integrin alpha9beta1. *J Biol Chem* 287, 25615-25630.
48. Kon, S., and Uede, T. (2018). The role of alpha9beta1 integrin and its ligands in the development of autoimmune diseases. *J Cell Commun Signal* 12, 333-342.
49. Morooka, N., Futaki, S., Sato-Nishiuchi, R., Nishino, M., Totani, Y., Shimono, C., Nakano, I., Nakajima, H., Mochizuki, N., and Sekiguchi, K. (2017). Polydom Is an Extracellular Matrix Protein Involved in Lymphatic Vessel Remodeling. *Circ Res* 120, 1276-1288.
50. Nakada, T.A., Russell, J.A., Boyd, J.H., Thair, S.A., and Walley, K.R. (2015). Identification of a nonsynonymous polymorphism in the SVEP1 gene associated with altered clinical outcomes in septic shock. *Crit Care Med* 43, 101-108.

51. Giri, A., Hellwege, J.N., Keaton, J.M., Park, J., Qiu, C., Warren, H.R., Torstenson, E.S., Kovesdy, C.P., Sun, Y.V., Wilson, O.D., et al. (2019). Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat Genet* 51, 51-62.
52. Maciver, S.K., and Hussey, P.J. (2002). The ADF/cofilin family: actin-remodeling proteins. *Genome Biol* 3, reviews3007.
53. Burbage, M., and Keppler, S.J. (2018). Shaping the humoral immune response: Actin regulators modulate antigen presentation and influence B-T interactions. *Mol Immunol* 101, 370-376.
54. Lee, K.H., Meuer, S.C., and Samstag, Y. (2000). Cofilin: a missing link between T cell co-stimulation and rearrangement of the actin cytoskeleton. *Eur J Immunol* 30, 892-899.
55. Nielsen, J.B., Thorolfsdottir, R.B., Fritsche, L.G., Zhou, W., Skov, M.W., Graham, S.E., Herron, T.J., McCarthy, S., Schmidt, E.M., Sveinbjornsson, G., et al. (2018). Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat Genet* 50, 1234-1239.
56. Agrawal, P.B., Greenleaf, R.S., Tomczak, K.K., Lehtokari, V.L., Wallgren-Pettersson, C., Wallefeld, W., Laing, N.G., Darras, B.T., Maciver, S.K., Dormitzer, P.R., et al. (2007). Nemaline myopathy with minicores caused by mutation of the CFL2 gene encoding the skeletal muscle actin-binding protein, cofilin-2. *Am J Hum Genet* 80, 162-167.
57. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47, D607-D613.
58. Yamada, Y., Kato, K., Oguri, M., Horibe, H., Fujimaki, T., Yasukochi, Y., Takeuchi, I., and Sakuma, J. (2018). Identification of 13 novel susceptibility loci for early-onset myocardial infarction, hypertension, or chronic kidney disease. *Int J Mol Med* 42, 2415-2436.
59. Liu, J.Z., van Sommeren, S., Huang, H., Ng, S.C., Alberts, R., Takahashi, A., Ripke, S., Lee, J.C., Jostins, L., Shah, T., et al. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 47, 979-986.
60. Franke, A., McGovern, D.P., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R., et al. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 42, 1118-1125.
61. Tsoi, L.C., Spain, S.L., Knight, J., Ellinghaus, E., Stuart, P.E., Capon, F., Ding, J., Li, Y., Tejasvi, T., Gudjonsson, J.E., et al. (2012). Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nat Genet* 44, 1341-1348.
62. Australo-Anglo-American Spondyloarthritis, C., Reveille, J.D., Sims, A.M., Danoy, P., Evans, D.M., Leo, P., Pointon, J.J., Jin, R., Zhou, X., Bradbury, L.A., et al. (2010). Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* 42, 123-127.
63. Anderson, C.A., Boucher, G., Lees, C.W., Franke, A., D'Amato, M., Taylor, K.D., Lee, J.C., Goyette, P., Imielinski, M., Latiano, A., et al. (2011). Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 43, 246-252.
64. Guinamard, R., Okigaki, M., Schlessinger, J., and Ravetch, J.V. (2000). Absence of marginal zone B cells in Pyk-2-deficient mice defines their role in the humoral response. *Nat Immunol* 1, 31-36.
65. Nair, R.P., Duffin, K.C., Helms, C., Ding, J., Stuart, P.E., Goldgar, D., Gudjonsson, J.E., Li, Y., Tejasvi, T., Feng, B.J., et al. (2009). Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat Genet* 41, 199-204.
66. Iwasaki, H., and Akashi, K. (2007). Myeloid lineage commitment from the hematopoietic stem cell. *Immunity* 26, 726-740.
67. McKercher, S.R., Torbett, B.E., Anderson, K.L., Henkel, G.W., Vestal, D.J., Baribault, H., Klemsz, M., Feeney, A.J., Wu, G.E., Paige, C.J., et al. (1996). Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J* 15, 5647-5658.

68. Sevilla, L., Zaldumbide, A., Carlotti, F., Dayem, M.A., Pognonec, P., and Boulukos, K.E. (2001). Bcl-XL expression correlates with primary macrophage differentiation, activation of functional competence, and survival and results from synergistic transcriptional activation by Ets2 and PU.1. *J Biol Chem* 276, 17800-17807.
69. Han, J.W., Zheng, H.F., Cui, Y., Sun, L.D., Ye, D.Q., Hu, Z., Xu, J.H., Cai, Z.M., Huang, W., Zhao, G.P., et al. (2009). Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet* 41, 1234-1237.
70. de Lange, K.M., Moutsianas, L., Lee, J.C., Lamb, C.A., Luo, Y., Kennedy, N.A., Jostins, L., Rice, D.L., Gutierrez-Achury, J., Ji, S.G., et al. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 49, 256-261.
71. Espinoza, J.L., Ai, S., and Matsumura, I. (2018). New Insights on the Pathogenesis of Takayasu Arteritis: Revisiting the Microbial Theory. *Pathogens* 7.
72. Hoyer, B.F., Mumtaz, I.M., Loddenkemper, K., Bruns, A., Sengler, C., Hermann, K.G., Maza, S., Keitzer, R., Burmester, G.R., Buttgerit, F., et al. (2012). Takayasu arteritis is characterised by disturbances of B cell homeostasis and responds to B cell depletion therapy with rituximab. *Ann Rheum Dis* 71, 75-79.
73. Zaldivar Villon, M.L.F., de la Rocha, J.A.L., and Espinoza, L.R. (2019). Takayasu Arteritis: Recent Developments. *Current rheumatology reports* 21, 45.
74. Terao, C., Matsumura, T., Yoshifuji, H., Kirino, Y., Maejima, Y., Nakaoka, Y., Takahashi, M., Amiya, E., Tamura, N., Nakajima, T., et al. (2015). Takayasu arteritis and ulcerative colitis: high rate of co-occurrence and genetic overlap. *Arthritis Rheumatol* 67, 2226-2232.
75. Kwon, O.C., Lee, S.W., Park, Y.B., Oh, J.S., Lee, S.H., Hong, S., Lee, C.K., Yoo, B., and Kim, Y.G. (2018). Extravascular manifestations of Takayasu arteritis: focusing on the features shared with spondyloarthritis. *Arthritis Res Ther* 20, 142.
76. Saad, M.N., Mabrouk, M.S., Eldeib, A.M., and Shaker, O.G. (2019). Studying the effects of haplotype partitioning methods on the RA-associated genomic results from the North American Rheumatoid Arthritis Consortium (NARAC) dataset. *J Adv Res* 18, 113-126.
77. Genga, E., Oyoo, O., and Adebajo, A. (2018). Vasculitis in Africa. *Curr Rheumatol Rep* 20, 4.
78. Plenge, R.M., Scolnick, E.M., and Altshuler, D. (2013). Validating therapeutic targets through human genetics. *Nat Rev Drug Discov* 12, 581-594.
79. Brandstadter, R., and Katz Sand, I. (2017). The use of natalizumab for multiple sclerosis. *Neuropsychiatr Dis Treat* 13, 1691-1702.
80. Guagnozzi, D., and Caprilli, R. (2008). Natalizumab in the treatment of Crohn's disease. *Biologics* 2, 275-284.
81. Breedveld, F.C., and Dayer, J.M. (2000). Leflunomide: mode of action in the treatment of rheumatoid arthritis. *Ann Rheum Dis* 59, 841-849.
82. Cui, X., Dai, X., Ma, L., Yang, C., Tan, W., Zhang, L., Zhang, Z., Feng, X., Wu, R., Zou, Y., et al. (2020). Efficacy and safety of leflunomide treatment in Takayasu arteritis: Case series from the East China cohort. *Semin Arthritis Rheum* 50, 59-65.
83. de Souza, A.W., de Almeida Agustinelli, R., de Cinque Almeida, H., Oliveira, P.B., Pinheiro, F.A., Oliveira, A.C., and Sato, E.I. (2016). Leflunomide in Takayasu arteritis - A long term observational study. *Rev Bras Reumatol Engl Ed* 56, 371-375.
84. Adizie, T., Christidis, D., Dharmapaliah, C., Borg, F., and Dasgupta, B. (2012). Efficacy and tolerability of leflunomide in difficult-to-treat polymyalgia rheumatica and giant cell arteritis: a case series. *Int J Clin Pract* 66, 906-909.

Figure Legends

Figure 1. Manhattan plot showing the results of the meta-analysis among the 5 populations included in this study. The $-\log_{10}$ p value for the genetic variants analyzed are plotted against their physical chromosomal position. The red and blue lines represent the genome-wide level of significance (p value $< 5 \times 10^{-8}$) and the suggestive level of significance (p value $< 5 \times 10^{-5}$), respectively.

Figure 2. Manhattan plots showing the results of stepwise conditional analysis of the *HLA* region. A) Unconditioned. B) Conditioning on rs2844678. C) Conditioning on rs2844678 + rs12524487. D) Conditioning on rs2844678 + rs12524487 + rs17193507. E) Conditioning on rs2844678 + rs12524487 + rs17193507 + rs12526858. The $-\log_{10}$ p value for the genetic variants analyzed are plotted against their physical chromosomal position. The red line represents the genome-wide level of significance (p value $< 5 \times 10^{-8}$).

Figure 3. Column bar plot representing the results from histone marks enrichment analysis among genetic susceptibility loci identified with a GWAS level of significance (p value $< 5 \times 10^{-8}$) in Takayasu arteritis in this study. Y-axis corresponds with the $-\log_{10}$ corrected p value of H3K27ac, H3K4me1 and H3K4me3 enrichments in different blood cells types. These data were derived using ENCODE and Roadmap Epigenomics Consortium projects. PBMCs, peripheral blood mononuclear cells; HSCs, hematopoietic stem cells.

Figure 4. Functional annotation results point to *ETS2* as a causal gene for the chr21q22 association. A) Summary of the epigenetic marks found at rs2836882 in chr21q22. B) Circular view of the chromatic interaction between rs2836882 in chr21q22 and *ETS2* promoter obtained from Capture Hi-C data. This interaction is observed in several cell types: endothelial precursors, megakaryocytes, monocytes, M0 macrophages, M1 macrophages, M2 macrophages, neutrophils, and CD34+ cells. The red color of the line indicates a strong confidence for this interaction. C) Violin plot

representing the difference in expression levels of *ETS2* depending on rs2836882 genotypes obtained from GTEx project.

Figure 5. Takayasu arteritis genetic susceptibility loci identified and the main Gene Ontology terms represented. A) Schematic illustration depicting the susceptibility loci identified in this study. Loci surpassing the genome-wide level of significance (p value $< 5 \times 10^{-8}$) are highlighted in red. B) Gene Ontology enrichment analysis. Biological Process Gene Ontology (GO) clusters revealed using genes annotated to genetic susceptibility loci identified in this study (p value $< 5 \times 10^{-5}$). Only clusters detected with a p value $< 1 \times 10^{-3}$ and that include at least three genes are shown. Each node corresponds with a GO term, and edges represent connections between GO terms. Each GO cluster is defined by a different color and the most associated Biological Process within each cluster is highlighted.

Figure 6. Dot plot representing the cumulative Genetic Risk Score (GRS) for Takayasu arteritis across the five major populations included in the 1000 Genome Project. Each dot illustrates the GRS value of an individual and the black lines represent the means and ranges. Data were used from the 1000 Genome Project, phase 3 release. African, N=661; Admixed American, N=347; East Asian, N=504; European, N=503; South Asian, N=489. There was a significant difference between the genetic risk scores among populations (ANOVA p value < 0.0001). The multiple comparison test on mean GRS between each pair of populations revealed statistically significant differences for all pairs (p value < 0.0001) except African versus East Asian and European versus South Asian (Tukey's adjusted p value = 0.924 and 0.457 respectively).

Figure 1

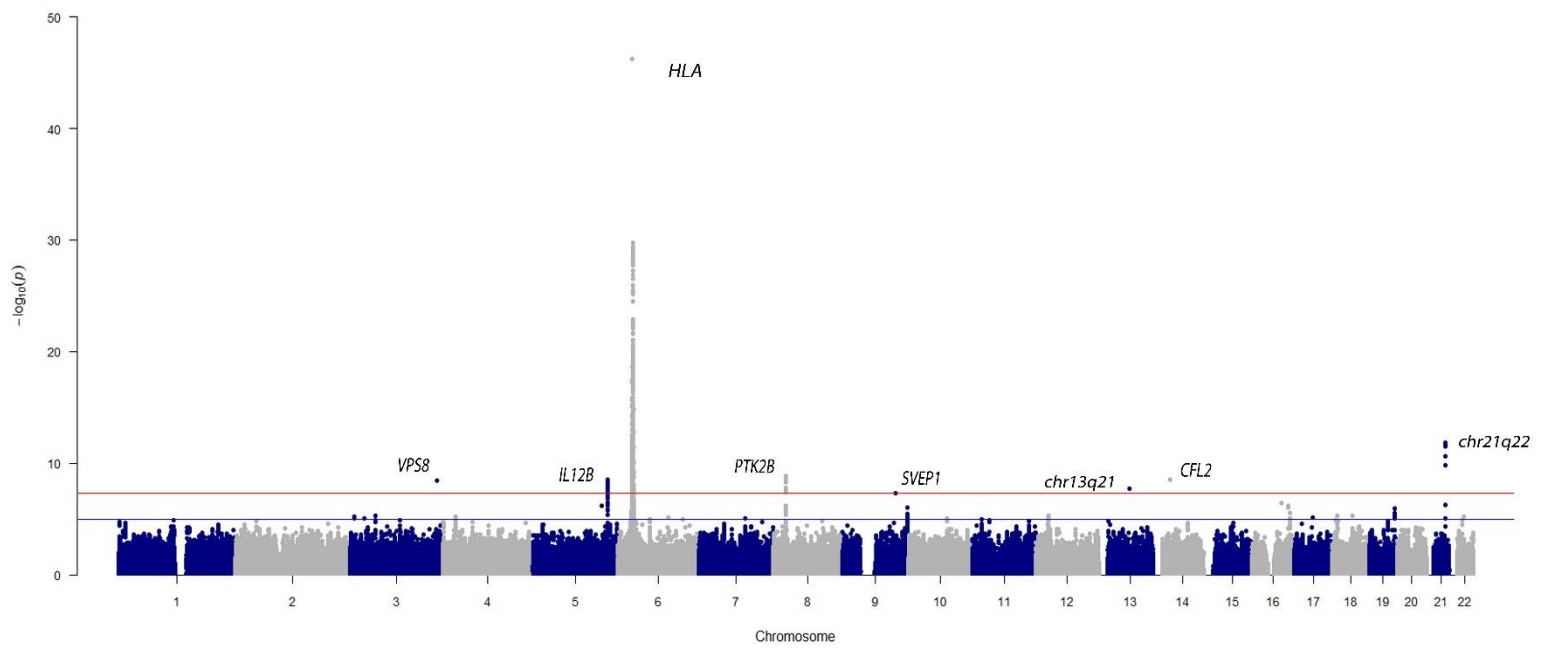


Figure 2

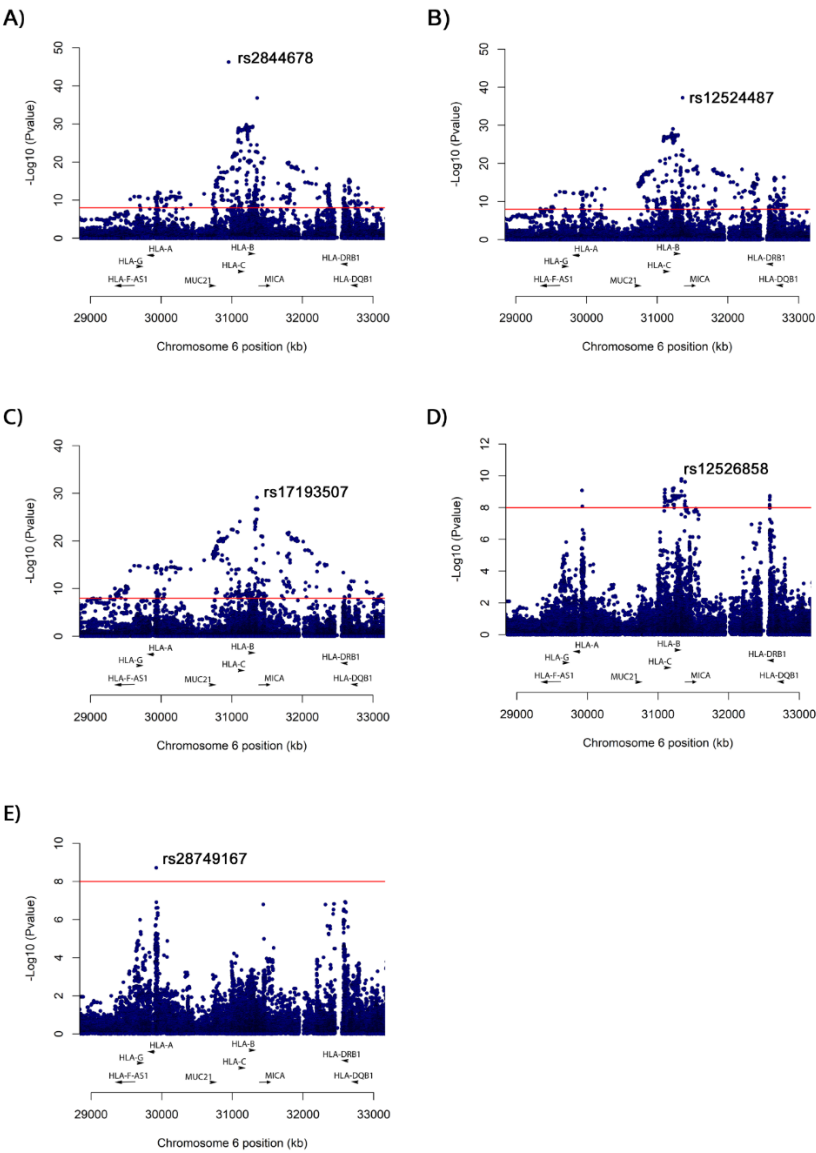


Figure 3

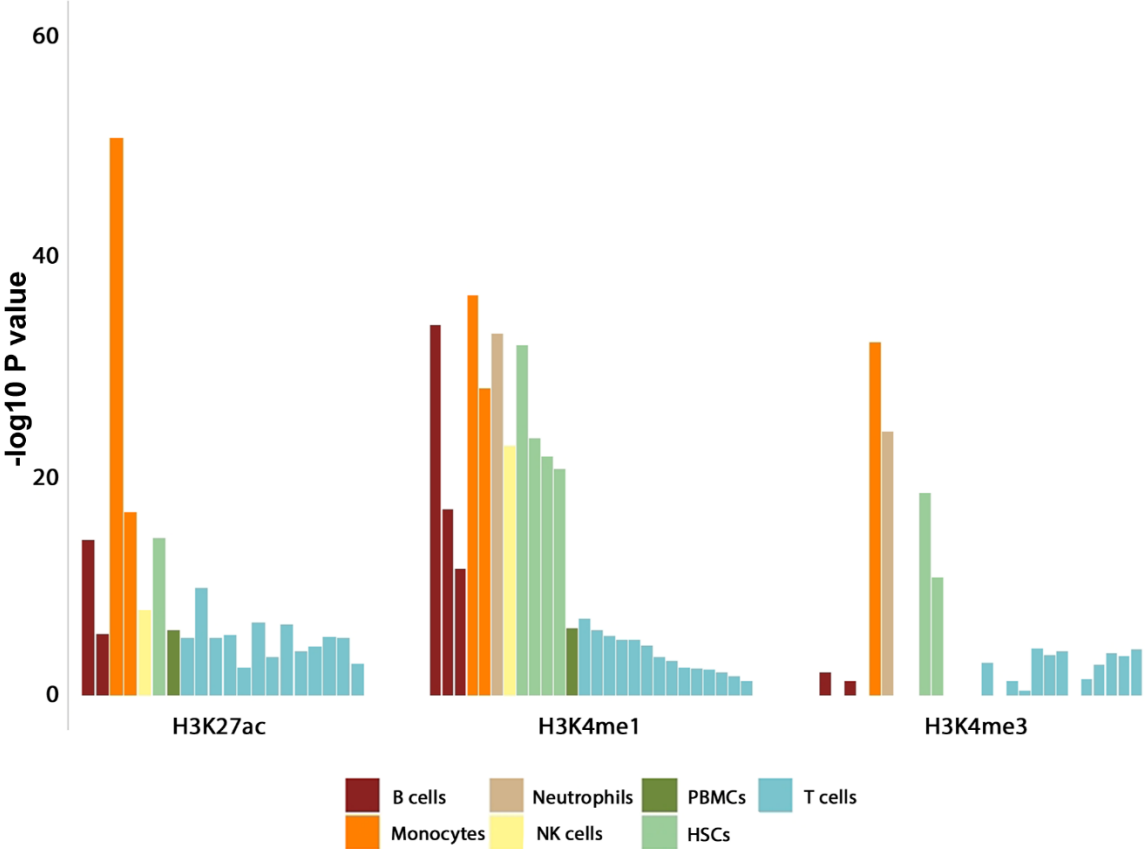
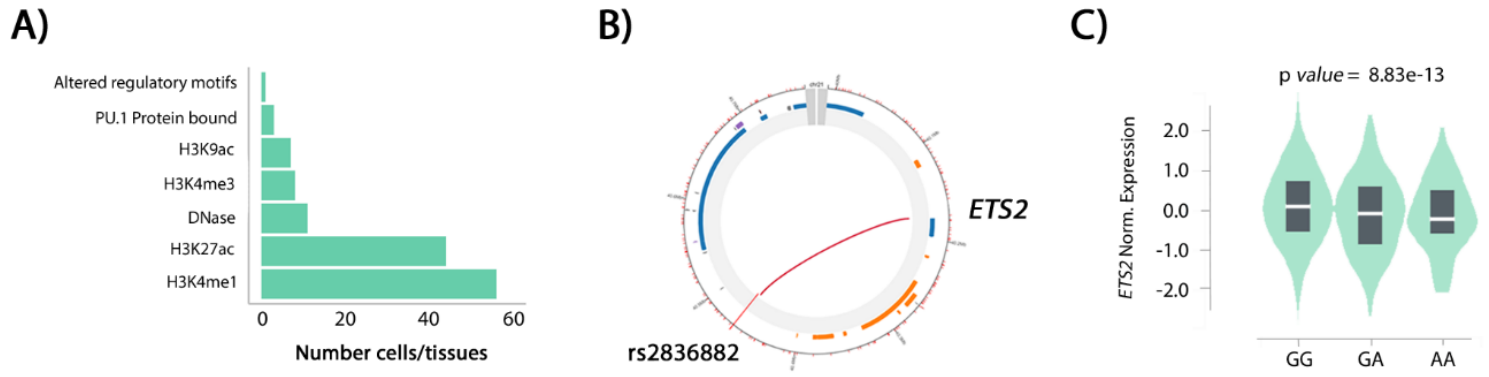


Figure 4



A)



Figure 6

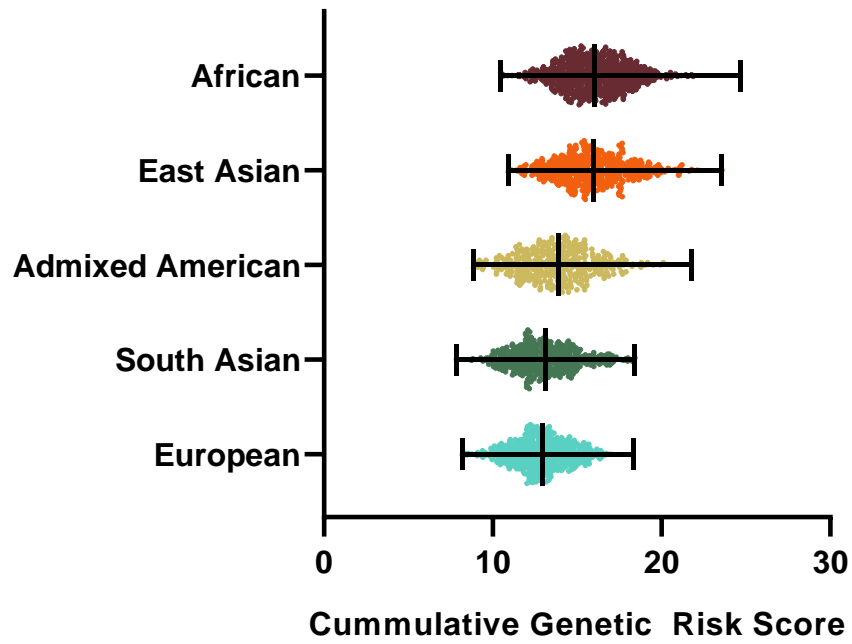


Table 1. Step-wise conditional analysis of the *HLA* variants and the imputed classical *HLA* alleles to detect independent associations with Takayasu arteritis.

SNPs	Covariates	Risk factor	P value	OR	Position (HG19)	Locus
	None	rs2844678	5.70E-47	2.71	30950050	5' <i>MUC21</i>
	rs2844678	rs12524487	5.72E-38	3.12	31354238	<i>HLA-B/MICA</i>
	rs2844678 + rs12524487	rs17193507	7.08E-30	3.28	31351054	<i>HLA-B/MICA</i>
	rs2844678 + rs12524487 + rs17193507	rs12526858	1.55E-10	2.70	31323721	<i>HLA-B</i>
	rs2844678 + rs12524487 + rs17193507 + rs12526858	rs28749167	1.90E-09	1.43	29921034	<i>HLA-G</i>
Classical Alleles	Covariates	Risk factor	P value	OR		
	None	<i>HLA-B*5201</i>	4.16E-36	4.37		
	<i>HLA-B*5201</i>	<i>HLA-B*1302</i>	2.96E-25	3.21		
	<i>HLA-B*5201</i> + <i>HLA-B*1302</i>	<i>HLA-B*1501</i>	1.09E-15	3.07		
	<i>HLA-B*5201</i> + <i>HLA-B*1302</i> + <i>HLA-B*1501</i>	<i>HLA-DQB1*0502</i>	3.88E-08	1.90		

SNP, Single Nucleotide Polymorphism; OR, odds ratio

Table 2. Meta-analysis results showing the most significant genetic variant in each locus associated with Takayasu arteritis at a genome-wide level of significance (p value <5 X 10⁻⁸)

Locus	Chr	Position (HG19)	SNP	Location	Minor Allele	P value	OR	Q Statistic p value
VPS8	3	184612321	rs58693904	Intronic	T	3.73E-09	1.51	0.550
<i>IL12B</i>	5	158804928	rs4379175	Upstream	T	3.13E-09	1.36	0.173
<i>PTK2B</i>	8	27195121	rs28834970	Intronic	C	1.40E-09	0.71	0.957
SVEP1	9	113117183	rs7038415	Upstream	A	4.90E-08	1.45	0.174
chr13q21	13	65078467	rs9540128	Intergenic	T	1.89E-08	2.02	0.793
CFL2	14	35189914	rs76457959	Upstream	T	2.84E-09	1.62	0.682
chr21q22	21	40464924	rs4817983	Intergenic	C	1.34E-12	0.64	0.125

Novel previously unreported genetic loci in Takayasu arteritis are depicted in bold font.

Chr, Chromosome; SNP, Single Nucleotide Polymorphism; OR, odds ratio