Genome-Wide Association Study to Identify Novel Susceptibility Loci for Kawasaki Disease in Han Chinese

Fuu-Jen Tsai[1,2,3] ¶, Yi-Ching Lee[4,5,6]¶, Jeng-Sheng Chang[7]¶, Li-Min Huang[8], Fu-Yuan Huang[9], Nan-Chang Chiu[9], Ming-Ren Chen[9], Hsin Chi[9], Yann-Jinn Lee[9], Li-Ching Chang[4] , Yi-Min Liu[4], Hsiang-Hua Wang[4], Chien-Hsiun Chen[4,5,10], Yuan-Tsong Chen[4,5,11*], Jer-Yuarn Wu[4,5,12*]

The work is generated from Institute of Biomedical Sciences and National Genotyping Center, Academia Sinica, Taipei, Taiwan,

1 School of Chinese Medicine, China Medical University, Taichung, Taiwan,

2 Department of Medical Genetics, Medical Research and Pediatrics, China Medical University Hospital, Taichung, Taiwan,

3 Department of Health and Nutrition Biotechnology, Asia University, Taichung, Taiwan,

4 Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan,

5 National Genotyping Center, Academia Sinica, Taipei, Taiwan,

6 Graduate Institute of Integrated Medicine, China Medical University, Taichung, Taiwan,

7 Divisions of Cardiology, Department of Pediatrics, China Medical University and Hospital, Taichung, Taiwan,

删除: c

8 Department of Pediatrics, National Taiwan University Hospital, College of Medicine, National Taiwan University, Taipei, Taiwan

9 Department of Pediatrics, Mackay Memorial Hospital, Taipei County, Taiwan

10 Graduate Institute of Chinese Medical Science, China Medical University, Taichung, Taiwan,

11 Department of Pediatrics, Duck University Medical Center, Durham, North Carolina, United States of American,

12 Graduate Institute of Chinese Medical Science, China Medical University, Taichung, Taiwan

* Corresponding author. Institute of Biomedical Sciences, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei, 11529, Taiwan.

Tel: 886-2-2789-9081; Fax: 886-2-2782-5573.

E-mail: chen0010@ibms.sinica.edu.tw (Y-TC);

Tel :886-2-2789-9075 ; Fax :886-2782-4066

E-mail: jywu@ibms.sinica.edu.tw (J-YW)

¶ These authors contributed equally to this work.

Brief title: Susceptibility Genes for Kawasaki Disease

Abstract

Kawasaki disease (KD) is an acute systemic vasculitis syndrome primarily affecting infants and young children with unknown etiology. Epidemiological findings suggest that genetic predisposition underlies the disease susceptibility. Taiwan has the third highest incidence of KD in the world, after Japan and Korea. To investigate novel mechanisms that might predispose individuals to KD. We conducted a genome-wide association study (GWAS) in 250 KD patients and 446 controls in a Han Chinese population residing in Taiwan. This was further replicated in an independent Han Chinese cohort of 208 cases and 366 controls. Three novel loci with most strongly associated SNPs were detected in the joint analysis. Among them, three SNPs close to mitochondrial ribosomal protein S22 **(**MRPS22**)**, coatomer protein complex, subunit beta-2 **(**COPB2**),** and retinol-binding protein 2 (PBP2) genes, are associated with KD, with P-values of 9.52×10^{-5} (rs1873668), 9.93×10^{-5} (rs4243399), and 9.93×10^{-5} (rs16849083), respectively. We have also identified a SNP in the intronic region of the endoplasmic reticulum amino peptidase 1 (ERAP1) gene (rs149481, $p_{best} = 4.61 \times 10^{-5}$). Six SNPs, rs17113284; rs8005468; rs10129255, rs2007467, rs10150241; and rs12590667 with p_{best} between 2.08×10^{-5} to 8.93×10^{-6}, clustered in an area containing immunoglobulin heavy chain variable regions were identified. This is the first GWAS performed in subjects of Han

刪除: pathogenesis

刪除: ese

刪除: ve

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

刪除: (a disease prevalence of 0.01%), just behind Japan and Korea

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

刪除: Asian descent

Chinese. The novel KD candidates we identified are implicated in T cell receptor signaling, MHC I peptide presentation, regulation of proinflammatory cytokines, as well as an antibody-mediated immune response. The findings may lead to a better understanding of the underlying molecular pathogenesis of KD.

Introduction

Kawasaki disease (KD) is the most common acquired heart disease in children world-wide. The incidence of KD varies among different countries, with Asian countries having higher incidences than Western countries. Japan has the highest annual incidence in the world, which has 200 KD cases per 100,000 children under age 5 [1,2]. The annual incidence in Korea is 113.1 KD cases per 100,000 children under 5 years old [3]. The annual incidence in Taiwan is 66 to 69 KD cases per 100,000 in children under age 5, after Japan and Korea [4,5]. KD is an acute, self-limited vasculitis of infants and children. Symptoms include prolonged fever unresponsive to antibiotics, polymorphous skin rash, swollen glands, red eyes, inflammation of the mouth, extensive rash, and swollen and red hands and feet. Characteristic peeling of the skin on fingers and toes occurs during convalescence [6]. Coronary aneurysms develop in 15%–25% of untreated children [6,7], making this disease the leading cause of acquired heart disease among children in developed

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

刪除: [4,5]

刪除: [1,2]

刪除: Kawasaki disease (

刪除:)

countries. The cause of this disease is currently unknown. The clinical and epidemiological features strongly suggested that an infectious trigger KD and genetic predisposition may underlie the etiology of the disease. It is proposed that an inflammatory stimulus sets in motion a cascade of events that leads to host immune dysregulation in genetically predisposed individuals. During the acute stage of KD, infiltration of CD8-positive T cells and macrophages, activation of vascular endothelial cells, and increased serum levels of proinflammatory cytokines lead to inflammation and injury of blood vessels [8,9].Thus, identification of genetic factors would greatly facilitate understanding of the disease etiology and pathophsyiology.

A number of genes participating in immune-regulatory responses and cardiovascular-related loci have been extensively studied the association with KD susceptibility or disease outcome through the candidate gene approach [10]. Several candidate genes have been examined in independent cohorts of the same or different ethnicity, however most of the results were conflicting. Most of the previous studies have been carried out with a single small cohort of KD patients and findings were reported without validation in additional case-control sets. There have been few genome-wide studies of this disease. A genome-wide linkage analysis conducted with Japanese KD sib-pair samples revealed several linked regions, and led to the identification of a functional polymorphism in the inositol 1,4,5-triphosphate 3-

kinase C (ITPKC) gene [11,12]. A GWAS conducted in an international cohort of Caucasian patients. Variants located in genes with potentially functionally related to inflammation, apoptosis, and cardiovascular pathology with significance level of P $=4.7 \times 10^{-2}$ to 1.9×10^{-4} were identified [13].

To identify novel mechanisms that might predispose patients to KD, we conducted a two-stage GWAS scan for KD in a Han Chinese population residing in Taiwan. In this paper, we present results of the first GWAS in KD in a Han-Chinese population, and new KD susceptibility loci involved in the immune response and associated with increased risk of KD were identified.

删除: Taiwanese have the third highest incidence of KD, just behind Japan and Korea.

Results

Data quality

Individual call rates were greater than 0.95 for all of the subjects genotyped in this study. The average individual call rate was 0.993 (standard deviation [SD] = 0.005).

Of the 906,955 SNPs on the Affymetrix 6.0 SNP chip, 96,955 were non-polymorphisc in both KD cases and controls, 27,779 had a call rate of <95%, 30,080 had overall MAF < 5% and had total call rate <99% and 28,093 had a significant difference from Hardy–Weinberg equilibrium (HWE) compared to the control group (p <10^{-7}) deviated significantly from HWE in the control group. A

total of 723,638 SNPs (79.8%) passing a quality control filter with an average SNP call rate of 0.996 (SD = 0.006) were included in the GWAS (Table S1).

删除: , a minor allele frequency ≧0.05, and a significant difference from Hardy–Weinberg equilibrium (HWE) compared to the control group ($p \geqq 10^{-4}$),

Assessment of population stratification

Multidimensional scaling analysis (Figure S1) and the permutation test for between-group identity-by-state (IBS) differences ($p = 0.460$) suggested that no strong population stratification existed. Furthermore, the variance inflation factor for genomic control, $\lambda = 1.118$, also indicated that no very strong stratification existed.

GWAS and cross-platform validation

The analysis was first performed with samples from 250 people with KD and 446 controls (Table S2). A total of 141 SNPs out of the 723,638 SNPs were associated with KD ($p < 10^{-4}$) (Figure 1). Seventy-four out of 141 SNPs that have consecutive significant SNP around and located either in or within 500 kb of known genes and with good Affymetrix calling of clustering in both cases and controls were chosen for cross-platform validation using Sequenom MassARRAY. Fifty-seven SNPs out of these 74 SNPs were retained after cross-platform validation, in which they reached an overall 99% consistency and had association P_{best} values of $<10^{-4}$ with any of the genotype, allele, trend, dominant, and recessive models in both platforms

(Table S3).

Replication of identified SNPs in follow-up cohort

We then genotyped the 57 validated associated SNPs in an additional 208 cases and 366 controls. The results of the 57 SNPs in 208 cases and 366 controls (stage 2) and joint analysis in 458 cases and 812 controls are listed in Table S3. Twelve out of 57 SNPs had a nominal p-values ≤0.05 in the replication cohort. Thirteen SNPs had most significant association with KD in joint analysis ($P_{best} < 10^{-4}$) were listed in Table1 and highlighted in the Table S3. Out of them, ten SNPs in three regions with consecutive significant SNPs around were selected and described below. The estimated recombination rates based on the Chinese HapMap and the association results from GWAS in the three regions are shown in Figure 2.

The rs1873668 ($p_{best} = 9.52 \times 10^{-5}$), rs4243399 ($p_{best} = 9.93 \times 10^{-5}$), and rs16849083 ($p_{best} = 9.93 \times 10^{-5}$) were located on chromosome 3q23 and had association with KD in joint analysis ($P_{best} < 10^{-4}$) (Table 1). Their nearby SNPs rs1873666 ($p_{best} = 0.000123$), rs4894410 ($p_{best} = 0.000103$), and rs16849065 ($p_{best} = 0.000243$) also showed significant association in joint analysis (Table S3). These SNPs were located to a LD region containing genes encoded mitochondrial ribosomal protein S22 (MRPS22), coatomer protein complex, subunit beta-2

删除: of

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

(COPB2) , and retinol-binding protein 2, PBP2 (Figure 2A and Figure S5).

The rs149481 with $p_{best} = 4.61 \times 10^{-5}$ in join analysis was located on chromosome 5 15q (Table 1); lie in the intron region of the endoplasmic reticulum amino peptidase 1 (ERAP1) gene (Figure 2B). The nearby SNP rs27042 showed association (p_{best} =0.000153) in joint analysis (Table S3) is also located in the same intronic region of the ERAP1 gene (Figure 2B). ERAP1 plays a role in peptide trimming for the generation of most HLA class I–binding peptides, which is also involved in regulating the proinflammatory cytokines signaling through cleavage the cell surface receptors for them.

A cluster of variants were located in a LD region of the chromosome 14 q32.33(Figure S5), including: rs17113284 ($p_{best} = 1.08 \times 10^{-5}$), rs8005468 ($p_{best} = 8.93 \times 10^{-6}$), rs10129255 ($p_{best} = 1.14 \times 10^{-5}$), rs2007467 ($p_{best} = 1.28 \times 10^{-5}$), rs10150241 ($p_{best} = 1.08 \times 10^{-5}$), and rs12590667 ($p_{best} = 1.27 \times 10^{-5}$) (Table1 and Table S3). This region contains genes coding for immunoglobulin heavy chain variable regions (Figure 2 C).

Haplotypic Association

We further performed a haplotypic association analysis. Fifty seven haptotype blocks showed $p<10^{-4}$ in the original cohort analysis. 171 SNPs within the haplotype blocks were genotyped using a Sequenom iPLEX platform in the

刪除: The rs1873668 and rs4243399 are not located on any known gene but the two nearest genes mitochondrial ribosomal protein S22 (MRPS22) and coatomer protein complex, subunit beta-2 (COPB2) are about 20 kb away (Figure 2A). The rs16849083 is located in the intronic region of retinol-binding protein 2, PBP2 (Figure 2A)

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

刪除: region

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 縮排: 第一行: 1.27 cm

replication cohort (Table S4). Significant associations were verified for three haplotypes (with $p<10^{-4}$ in the replication cohort and with $p<10^{-5}$ in joint analysis) (Table S4).

格式化: 字型: 非粗體, 字型色彩: 自訂色彩(RGB(51,51,255))

Comparison of our results with previously identified candidate regions

Some genes with known functions or roles related to KD pathophysiology have been intensively studied the association with KD by the candidate gene approach.

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

刪除: showed

Therefore, we examined the SNPs in these gene regions in our GWAS to identify their association with KD in Han population. HLA-Bw22 (Bw54) was implicated to be associated with KD in Japanese [14]. A SNP located in the HLA-E gene was suggested to be associated with KD in the Han Chinese population [15]. Additionally, the SNP, TNF-α -308A, was shown to be associated with susceptibility to KD in a Chinese population [16], and a Taiwanese group reported that an insertion in the angiotensin I converting enzyme (ACE) was associated with KD susceptibility [17]. More importantly, the inositol 1,4,5-trisphosphate 3-kinase C (ITPKC) region was identified by linkage study to be associated with KD susceptibility in Japanese cohort and US cohort, and its function involved in T-cell activation was demonstrated in cell system [12]. Our results showed none of the SNPs located in or near these gene regions had a −log10 (p_{best}) higher than 4 in our GWAS. However, some SNPs located in HLA-A, HLA-B, TNF-alpha, and ITPKC

刪除: predominant

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

刪除: KD patients

刪除: reported

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

region did show a nominal association (P<0.01) in our GWAS (Figure S3).

The SNPs located in 8 regions with P-value less than 0.05 were identified in the combined analysis of an international cohort of Caucasian KD patients and in family-based KD cohort [13]. Therefore, we further examined the SNPs in these gene regions included in our GWAS. Our results showed none of the SNPs had a −log10 (p_{best}) higher than 4. However, 10 and 42 SNPs in NAALADL2 (N-acetylated alpha-linked acidic dipeptidase-like 2) and CSMD1 (CUB and Sushi multiple domains) region respectively showed a nominal association (P<0.01) in our GWAS (Figure S4).

Discussion

This study reported that to the best of our knowledge, the first GWAS in KD conducted in Han Chinese. We identified novel candidates associated with KD susceptibility. The identification of these genes may provide new insights into the development of KD.

刪除: and second GWAS in KD

Three KD association SNPs, rs1873668, rs4243399, and rs16849083, are located on chromosome 3q23 region, in which three genes are included. The head to tail overlapping genes MRPS22 and COPB2 are located nearest to the rs1873668 and rs4243399. The rs16849083 is located in the intronic region of PBP2. MRPS22,

a mitochondrial ribosomal protein, the mutations of NRPS22 lead to antenatal mitochondrial disease, which presented with antenatal skin oedema, hypotonia, cardiomyopathy, and tubulopathy [18]. COPB2 constitutes the coat of nonclathrin-coated vesicles and is essential for Golgi budding and vesicular trafficking. PBP2 is thought to participate in the uptake and/or intracellular metabolism of vitamin A. The roles of NRPS22 and PBP2 have no obvious connection to the pathogenesis of KD. In a phosphoproteomic analysis of T cell receptor signaling, which reveals that the proteins involved in receptor and membrane trafficking in endosomes and vesicles, play a role in T cell receptor signaling and are critical for T cell activation [19]. The possibility of COPB2 involvements in T cell activation and KD pathogenesis need to be further investigated.

We identified two KD association SNPs, rs149481 and rs27042, located in an intron of the ERAP1 gene. ERAP1 plays a critical role in trimming peptides to the optimal length for HLA class I presentation [20] and cleaving cell surface receptors for the proinflammatory cytokines. It is ubiquitously expressed in every tissue, and is expressed at higher levels in the trachea, thymus, and lymph nodes (based on a National Center for Biotechnology Information expressed sequence tags profile). From a functional perspective, ERAP1 represent an excellent biological candidate for KD pathogenesis. ERAP1 trimming the peptides for HLA class I presentation

删除: , because

could play a critical role in initiating the immune response. Furthermore, ERAP1 cleaves cell surface receptors for the proinflammatory cytokines IL-1 [21], IL-6 [22], and TNF [23], thereby downregulating their signaling. Genetic variants that alter the functioning of ERAP1 could, therefore, have proinflammatory effects through this mechanism. In addition to the association with KD, polymorphisms in ERAP1 have been found to be associated with ankylosing spondylitis, an HLA class I–mediated autoimmune disease, and enthesitis, a common inflammatory arthritis characterized by axial skeletal inflammation [24]. These findings suggest that KD could be triggered by peptides presented by HLA class I molecules on antigen presenting cells, or the augment of proinflammation cytokines signaling could be involved in KD, leading to initiation of an acquired immune response. However, the association between HLA and KD remains inconclusive. None of the SNPs located in or near these gene regions had a −log10 (p_{best}) higher than 4 in our GWAS, but some SNPs located in HLA-A, HLA-B region did show a nominal association (P<0.01). Whether the HLA play a role in the KD susceptibility will be further investigated in a larger sample size and by traditional HLA genotyping method. The clarification of association between HLA and KD will be help to elucidate the possible involvement pathway of ERAP1 associated with KD.

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

格式化: 字型色彩: 藍色

刪除: Furthermore, ERAP1 cleaves cell surface receptors for the proinflammatory cytokines IL-1 [21], IL-6 [22], and TNF [23], thereby downregulating their signaling. Genetic variants that alter the functioning of ERAP1 could, therefore, have proinflammatory effects through this mechanism. In addition to the association with KD, polymorphisms in ERAP1 have been found to be associated with ankylosing spondylitis, an HLA class I–mediated autoimmune disease, and enthesitis, a common inflammatory arthritis characterized by axial skeletal inflammation [24]. These findings suggest that KD could be triggered by peptides presented by HLA class I molecules on antigen presenting cells, or the augment of proinflammation cytokines signaling could be involved in KD, leading to initiation of an acquired immune response.

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

We found that a cluster of SNPs located in chromosome 14 immunoglobulin

heavy chain variable regions with the strongest associated in the joint analysis. The heavy chain molecule is a major contributor to the generation of immunoglobulin diversity and specificity. Therefore, in addition to cell-mediated immunity, antibody-mediated immunity may be critical in KD pathogenesis. IgA plasma cells infiltrate inflamed tissues, including coronary arteries in acute KD [25,26], and oligoclonal KD antibodies bind to an antigen in acute KD-inflamed ciliated bronchial epithelium [27]. Specific immunoglobulins produced by particular heavy chain recombinants following a KD antigen trigger could play a role in KD pathogenesis. Identification of these specific antibodies could lead to the elucidation of disease-relevant mechanisms and potential treatments. We also performed a haplotypic association analysis in the GWAS cohort. Seven haptotype blocks showed $p<10^{-4}$ and had the haplotype frequency higher than 0.05 (Table S4). The association of these haplotype blocks with KD will be further validated by Sequnome platform and in the replication cohort.

The major limitation of this study is that the sample size, which reflects the difficulties that we had in recruiting patients with such a rare disease, could only reach a statistical power of 0.78 to detect a disease allele with a frequency of 0.10 and a relative risk of 3.0, assuming a disease incidence of 0.01% under an additive model. Also, similar to other GWAS, our study was underpowered to detect very

刪除:

格式化: 縮排: 第一行: 0 cm

刪除: prevalence

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

small effect size. With the additional replication sample, the joint analysis marginally strengthened some of the SNPs in the GWAS. Many of the significantly associated SNPs in the discovery phase were either insignificant or of marginal significance in the replication phase because of small sample size. This raises the possibility, as with every GWAS, that there is some false discovery here. Replication of the findings in large well-powered independent samples is crucial. Therefore, a multicenter collaboration is required to increase the size of the study cohorts and to detect any smaller effects. Several genes implicated in previous candidate gene studies were not shown to have a −log10 (p_{best}) higher than 4 in our studies, probably due to the different relative importance of particular genetic variants in susceptibility of complex diseases between ethnicity or haplotype structures among ethnicity. However, some of the SNPs located in the candidate gene regions did show a KD association *p* value of <0.01. Further replication with a larger sample size and different populations will helpful to evaluate the involvement of these genes in KD pathogenesis. There was higher proportion of male in KD patients then control in our study because of the high incidence of KD in male. Although it is unclear whether and how this might bias the results, it may have lead to some degree of bias.

删除: With the additional replication sample, the joint analysis marginally strengthened the GWAS signals.

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

删除:

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

In summary, we conducted a GWAS in the Han Chinese population for KD.

Ten SNPs located in three novel loci are associated with KD in a Han Chinese population. The novel KD risk loci containing genes that are implicated in T cell receptor signaling, HLA class I presentation, signal of proinflammatory cytokines regulation, as well as an antibody-mediated immune response. Our finding suggests that antigen initiation, uncontrolled inflammation, and antibodies are involved in the pathogenesis of KD, which may provide new directions for future studies. In particular, to further characterize of causative variants and to elucidate the mechanisms of them may lead to the identification of novel diagnostic and therapeutic targets for KD.

Materials and Methods

Ethical statement

The study was approved by the institutional review board and the ethics committee of the institution review board of China Medical University Hospital, Mackay Memorial Hospital, and Academia Sinica, Taiwan. Written informed consents were obtained from the subjects or their parents in accordance with institutional requirements and Declaration of Helsinki principles.

Study subjects and phenotype definition

Unrelated patients with KD (250) were consecutively recruited from China Medical University Hospital in Taichung, Taiwan, in collaboration with the National Clinical Core (NCC) of Taiwan. There were 156 males (62.4%) and 94 females (37.6%) (Table S2), with an average age at onset of 1.76 ± 1.61 y. The follow-up 208 KD patients were recruited from China Medical University Hospital in Taichung, National Taiwan University Hospital in Taipei, Changhua Christian Hospital in Changhua, Taipei Velerans General Hospital in Taipei, and Chang Gung Memorial Hospital in Kaohsiung and Linkou, Taiwan. There were 142 male (68.3%) and 66 female (31.7%) (Table S2). All of the patients were diagnosed according to criteria for KD [28,29], including fever lasting at least 5 d and at least four of the following symptoms: (1) changes in extremities (e.g., erythema, edema, and desquamation), (2) bilateral conjunctival infection, (3) polymorphous rash, (4) cervical lymphadenopathy, (5) changes in lips and oral cavity (e.g., pharyngeal erythema, dry/fissured or swollen lips, and strawberry tongue). Only Han Chinese, who comprise 98% of the residents in Taiwan, were considered for recruitment. The ethnic background was defined by self-report questionnaires. The control group, 446 in GWAS study and 366 in follow-up study, was randomly selected from the Taiwan Han Chinese Cell and Genome Bank in Taiwan, as reported previously [30]. The prevalence of KD in the Taiwan population is less than 0.01%; hence, the

刪除: (Table S1)

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

刪除:

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

刪除: conjunctivitis

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

刪除: (*N* =

刪除:)

格式化: 字型色彩: 自訂色彩(RGB(0,0,204))

controls were presumably disease free.

Genotyping and quality control

Genomic DNA was extracted from patient blood using the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN, USA). Each individual was genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 with total 906,600 SNPs, according to the manufacturer's protocols, by National Genotyping Center at Academia Sinica, Taipei, Taiwan. Genotype calling was done using the BRNNP algorithm implemented in Genotype Console, with default parameters suggested by the platform manufacturer. All of the sample call rates were greater than 0.95, and the average call rate of the sample was 0.993 (SD = 0.005). Quality control of genotype data was performed by examining several summary statistics. First, the total call rate (successful call rate) and minor allele frequency (MAF) of cases and controls was calculated for each SNP. SNPs were excluded for further analysis if one of following conditions occurred: (1) only one allele appeared in cases and controls; (2) the total call rate was less than 0.95; and (3) the total MAF was less than 0.05 and the total call rate was less than 0.99. In addition, SNPs that significantly departed from Hardy–Weinberg equilibrium proportions (HWEP) were also excluded ($p < 10^{-4}$).

Population stratification

Detection of possible population stratification that could influence association analysis was carried out with multidimensional scaling analysis implemented in PLINK (http://pngu.mgh.harvard.edu/purcell/plink [30]). We also estimated the variance inflation factor for genomic control.

Genome-wide association analysis

Genome-wide association analysis was carried out to compare allele/genotype frequencies between cases and controls using five single-point methods: genotype, allele type, and Cochran–Armitage trend test, along with tests considering dominant and recessive inheritance modes. Quantile-quantile (Q–Q) plots were then used to examine *p*-value distributions (Figure S1). Two-point analysis was performed using a logistic regression model, regressing the affected status of two SNPs and their interaction. SNPs were coded as 0,1, and 2 for the number of minor alleles and treated as continuous variables.

删除:

Validation and replication

Top SNPs ($p < 10^{-4}$) from the genome-wide association analysis in the 250 cases

and 446 controls were further validated using MALDI-TOF mass spectrometry (Sequenom MassARRAY; Sequenom, San Diego, CA, USA). SNPs retained after cross-platform validation were then genotyped in an additional 208 cases and 366 controls for replication.

Haplotypic Analysis

Analysis of haplotypic associations was performed using a …method, a haplotype analysis method for population-based case-control association studies. It uses a fixed three SNPs sliding window.

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 字型: 非粗體, 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 字型色彩: 自訂色彩(RGB(51,51,255))

格式化: 字型: 非粗體, 字型色彩: 自訂色彩(RGB(51,51,255))

Supporting Information

Figure S1. Multidimensional scaling analysis plot.

The multidimensional scaling (MDS) plot shows the first two principal components, estimated by PLINK, based on the genotype data of 100,000 SNPs with equal spacing across the human genome randomly selected from 723,638 high-quality SNPs. No population stratification between the 250 KD cases (red) and 446 controls (green) was detected (IBS group difference empirical $p = 0.460$ for T1:

Figure S2. Q–Q plot for the trend test.

Q–Q plots are shown for the trend test based on the 723,638 high-quality SNPs of the initial analysis with 250 cases and 446 controls. Blue lines represent the upper and lower boundaries of the 95% confidence interval bands. Black dots showing deviations from the line of equality indicate either that the theoretical distribution was incorrect, or that the sample was contained with values generated by a true association.

Figure S3. Comparisons to previously identified KD susceptibility gene regions.
The figure shows comparisons to the susceptible gene regions reported by previous candidate genes studies and one genome wild linkage study. Triangles represent the −log10 of the minimal *p* values of the SNPs in the gene region and within a region of ±200 kb, based on the initial analysis of 250 KD cases and 446 controls.

Figure S4. Comparisons to previous GWAS.
The figure shows comparisons to the susceptible gene regions reported by previous GWAS. Squares represent the −log10 of the minimal *p* values of the SNPs in the gene region, based on the initial analysis of 250 KD cases and 446 controls.

Table S1. Quality control of participant data.

Table S2. Demographic and clinical characteristics of participants in the GWAS

and replication study.

Table S3. Association results and concordance rates for validated significant association SNPs in initiated cohort, and association results in replication study.

Acknowledgements:

We thank other doctors in China Medical University Hospital in Taichung, National Taiwan University Hospital in Taipei, Changhua Christian Hospital in Changhua, Taipei Velerans General Hospital in Taipei, and Chang Gung Memorial Hospital in Kaohsiung and Linkou, Taiwan for their contributions in the recruitment of KD patients and the patients and their families who devoted their time and effort to participate in this study.

格式化: 字型色彩: 自動

刪除: We would like to thank all of

References

1. Yanagawa H, Nakamura Y, Yashiro M, Ojima T, Tanihara S, et al. (1998) Results of the nationwide epidemiologic survey of Kawasaki disease in 1995 and 1996 in Japan. Pediatrics 102: E65.
2. Yanagawa H, Nakamura Y, Yashiro M, Oki I, Hirata S, et al. (2001) Incidence survey of Kawasaki disease in 1997 and 1998 in Japan. Pediatrics 107: E33.
3. Park YW, Han JW, Hong YM, Ma JS, Cha SH, et al. (2010) Epidemiological Features of Kawasaki disease in Korea, 2006-2008. Pediatr Int.
4. Chang LY, Chang IS, Lu CY, Chiang BL, Lee CY, et al. (2004) Epidemiologic features of Kawasaki disease in Taiwan, 1996-2002. Pediatrics 114: e678-682.
5. Huang WC, Huang LM, Chang IS, Chang LY, Chiang BL, et al. (2009) Epidemiologic features of Kawasaki disease in Taiwan, 2003-2006. Pediatrics 123: e401-405.
6. Kato H, Koike S, Yamamoto M, Ito Y, Yano E (1975) Coronary aneurysms in infants and young children with acute febrile mucocutaneous lymph node syndrome. J Pediatr 86: 892-898.
7. Kato H, Sugimura T, Akagi T, Sato N, Hashino K, et al. (1996) Long-term consequences of Kawasaki disease. A 10- to 21-year follow-up study of 594 patients. Circulation 94: 1379-1385.
8. Matsubara T, Furukawa S, Yabuta K (1990) Serum levels of tumor necrosis factor, interleukin 2 receptor, and interferon-gamma in Kawasaki disease involved coronary-artery lesions. Clin Immunol Immunopathol 56: 29-36.
9. Lin CY, Lin CC, Hwang B, Chiang B (1992) Serial changes of serum interleukin-6, interleukin-8, and tumor necrosis factor alpha among patients with Kawasaki disease. J Pediatr 121: 924-926.
10. Onouchi Y (2009) Molecular genetics of Kawasaki disease. Pediatr Res 65: 46R-54R.
11. Onouchi Y, Tamari M, Takahashi A, Tsunoda T, Yashiro M, et al. (2007) A genomewide linkage analysis of Kawasaki disease: evidence for linkage to chromosome 12. J Hum Genet 52: 179-190.
12. Onouchi Y, Gunji T, Burns JC, Shimizu C, Newburger JW, et al. (2008) ITPKC functional polymorphism associated with Kawasaki disease susceptibility and formation of coronary artery aneurysms. Nat Genet 40: 35-42.
13. Burgner D, Davila S, Breunis WB, Ng SB, Li Y, et al. (2009) A genome-wide association study identifies novel and functionally related susceptibility Loci for Kawasaki disease. PLoS Genet 5: e1000319.

14. Kato S, Kimura M, Tsuji K, Kusakawa S, Asai T, et al. (1978) HLA antigens in Kawasaki disease. Pediatrics 61: 252-255.
15. Lin YJ, Wan L, Wu JY, Sheu JJ, Lin CW, et al. (2009) HLA-E gene polymorphism associated with susceptibility to Kawasaki disease and formation of coronary artery aneurysms. Arthritis Rheum 60: 604-610.
16. Cheung YF, Huang GY, Chen SB, Liu XQ, Xi L, et al. (2008) Inflammatory gene polymorphisms and susceptibility to kawasaki disease and its arterial sequelae. Pediatrics 122: e608-614.
17. Wu SF, Chang JS, Peng CT, Shi YR, Tsai FJ (2004) Polymorphism of angiotensin-1 converting enzyme gene and Kawasaki disease. Pediatr Cardiol 25: 529-533.
18. Saada A, Shaag A, Arnon S, Dolfin T, Miller C, et al. (2007) Antenatal mitochondrial disease caused by mitochondrial ribosomal protein (MRPS22) mutation. J Med Genet 44: 784-786.
19. Mayya V, Lundgren DH, Hwang SI, Rezaul K, Wu L, et al. (2009) Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions. Sci Signal 2: ra46.
20. Chang SC, Momburg F, Bhutani N, Goldberg AL (2005) The ER aminopeptidase, ERAP1, trims precursors to lengths of MHC class I peptides by a "molecular ruler" mechanism. Proc Natl Acad Sci U S A 102: 17107-17112.
21. Cui X, Rouhani FN, Hawari F, Levine SJ (2003) Shedding of the type II IL-1 decoy receptor requires a multifunctional aminopeptidase, aminopeptidase regulator of TNF receptor type 1 shedding. J Immunol 171: 6814-6819.
22. Cui X, Rouhani FN, Hawari F, Levine SJ (2003) An aminopeptidase, ARTS-1, is required for interleukin-6 receptor shedding. J Biol Chem 278: 28677-28685.
23. Cui X, Hawari F, Alsaaty S, Lawrence M, Combs CA, et al. (2002) Identification of ARTS-1 as a novel TNFR1-binding protein that promotes TNFR1 ectodomain shedding. J Clin Invest 110: 515-526.
24. Harvey D, Pointon JJ, Evans DM, Karaderi T, Farrar C, et al. (2009) Investigating the genetic association between ERAP1 and ankylosing spondylitis. Hum Mol Genet 18: 4204-4212.
25. Rowley AH, Eckerley CA, Jack HM, Shulman ST, Baker SC (1997) IgA plasma cells in vascular tissue of patients with Kawasaki syndrome. J Immunol 159: 5946-5955.
26. Rowley AH, Shulman ST, Mask CA, Finn LS, Terai M, et al. (2000) IgA plasma cell infiltration of proximal respiratory tract, pancreas, kidney, and coronary artery in acute Kawasaki disease. J Infect Dis 182: 1183-1191.
27. Rowley AH, Baker SC, Shulman ST, Garcia FL, Guzman-Cottrill JA, et al.

(2004) Detection of antigen in bronchial epithelium and macrophages in acute Kawasaki disease by use of synthetic antibody. J Infect Dis 190: 856-865.
28. Newburger JW, Takahashi M, Gerber MA, Gewitz MH, Tani LY, et al. (2004) Diagnosis, treatment, and long-term management of Kawasaki disease: a statement for health professionals from the Committee on Rheumatic Fever, Endocarditis, and Kawasaki Disease, Council on Cardiovascular Disease in the Young, American Heart Association. Pediatrics 114: 1708-1733.
29. Kim S, Dedeoglu F (2005) Update on pediatric vasculitis. Curr Opin Pediatr 17: 695-702.
30. Pan WH, Fann CS, Wu JY, Hung YT, Ho MS, et al. (2006) Han Chinese cell and genome bank in Taiwan: purpose, design and ethical considerations. Hum Hered 61: 27-30.

Figure Legends

Figure 1. Graphical summary of KD GWAS in a Han Chinese population.

KD association was determined for SNPs on the Affymetrix SNP6.0 chip. The y-axis represents the $-\log_{10} p$ value, and the x-axis represents each of the 723,638 SNPs used in the primary scan of 250 KD cases and 446 controls.

Figure 2. LD structure and association results for disease-association regions.

The $-\log_{10} p$ values (left y-axis) for the best test from the primary scan were plotted as a function of genomic positions based on NCBI Build 36. The SNPs with the strongest signal in the joined analysis are denoted by blue diamonds. Estimated recombination rates (right y-axis) based on the Chinese HapMap population were plotted in blue across the region to reflect the local LD structure around the

significant SNPs. Gene annotations were taken from NCBI.

Table 1. SNPs showing the strongest association results for Kawasaki disease in Han Chinese

					Stage 1 (250 cases and 446 control))				Stage 2 (208 cases and 366 control))				Joint analysis stage1+2 (458 cases and 812 control)			
SNP	Chr	Nearest Gene(s)	Non-risk Allele	Risk Allele	RAF (KD)	RAF (Control)	OR (95% CI)	*p* value (best)	RAF (KD)	RAF (Control)	OR (95% CI)	*p* value (best)	RAF (KD)	RAF (Control)	OR (95% CI)	*p* value (best)
rs1873668	3	MRPS22	C	A	0.05	0.02	3.549 (1.835-6.865)	3.68E-05	0.03	0.02	1.579 (0.754-3.303)	2.22E-01	0.04	0.02	2.501 (1.545-4.047)	9.52E-05
rs4243399	3	COPB2	A	G	0.05	0.02	3.541 (1.831-6.850)	3.82E-05	0.03	0.02	1.579 (0.754-3.303)	2.22E-01	0.04	0.02	2.494 (1.541-4.037)	9.93E-05
rs16849083	3	PBP2	G	A	0.05	0.02	3.564 (1.842-6.895)	3.46E-05	0.03	0.02	1.497 (0.705-3.178)	2.91E-01	0.04	0.02	2.837 (1.714-4.697)	2.19E-05
rs13128867	4	SLC7A11	C	T	0.33	0.23	1.618 (1.268-2.065)	7.42E-05	0.33	0.27	1.316 (1.010-1.716)	4.16E-02	0.33	0.25	1.469 (1.228-1.758)	2.23E-05
rs149481	5	ERAP1	G	T	0.83	0.73	1.825 (1.377-2.418)	2.25E-05	0.78	0.74	1.261 (0.949-1.676)	6.84E-02	0.80	0.73	1.506 (1.235-1.837)	4.61E-05
rs362794	7	RELN	A	G	0.11	0.04	2.537 (1.652-3.897)	4.41E-06	0.07	0.06	1.308 (0.808-2.119)	2.34E-01	0.09	0.05	1.900 (1.386-2.604)	3.03E-05
rs17113284	14	IGHV	T	C	0.42	0.33	1.441 (1.152-1.803)	7.22E-05	0.38	0.34	1.182 (0.926-1.509)	2.52E-02	0.40	0.34	1.314 (1.114-1.549)	1.08E-05
rs8005468	14	IGHV	C	T	0.42	0.33	1.449 (1.158-1.813)	7.98E-05	0.38	0.34	1.184 (0.928-1.511)	2.34E-02	0.40	0.34	1.321 (1.120-1.558)	8.93E-06
rs10129255	14	IGHV	G	A	0.42	0.33	1.462 (1.168-1.830)	5.17E-05	0.38	0.34	1.160 (0.910-1.480)	2.19E-02	0.40	0.34	1.315 (1.116-1.551)	6.77E-06
rs2007467	14	IGHV	G	A	0.42	0.33	1.479 (1.181-1.853)	4.66E-05	0.38	0.34	1.143 (0.896-1.458)	3.54E-02	0.40	0.34	1.307 (1.108-1.542)	1.28E-05
rs10150241	14	IGHV	T	A	0.42	0.33	1.447 (1.157-1.810)	5.73E-05	0.38	0.35	1.152 (0.902-1.470)	3.06E-02	0.40	0.34	1.304 (1.106-1.538)	1.08E-05
rs12590667	14	IGHV	G	A	0.42	0.33	1.456 (1.164-1.822)	4.81E-05	0.38	0.35	1.141 (0.893-1.457)	3.90E-02	0.40	0.34	1.303 (1.105-1.537)	1.27E-05
rs1568657	15	BTBD1	A	G	0.63	0.52	1.542 (1.228-1.936)	8.69E-05	0.58	0.53	1.261 (0.981-1.620)	1.80E-02	0.61	0.52	1.409 (1.191-1.667)	6.61E-06

Chr, chromosome. Risk allele, allele with higher frequency in cases compared to controls. RAF (KD) and RAF (control), risk allele frequencies in cases and controls, respectively. OR, odds ratio for risk allele. *p* value (best), Minimal p-value of the five association tests:genotype, allele, trend, dominant, and recessive. Stage 1 (Genome scan) included 250 cases and 446 controls. Stage 2 (replication stage) included 208 cases and 366 controls. Alleles were indexed to the forward strand of NCBI Build 36.

格式化表格
格式化 ... [1]
格式化: 置中
格式化: 置中
格式化 ... [2]
格式化 ... [3]
格式化 ... [4]
格式化 ... [5]
格式化 ... [6]
格式化 ... [7]
格式化 ... [8]
格式化 ... [9]
格式化 ... [10]
格式化 ... [11]
格式化 ... [12]
格式化 ... [13]
格式化 ... [14]
删除: ... [15]

Abbreviations

KD, Kawasaki disease

GWAS, genome-wide association study

MRPS22, mitochondrial ribosomal protein S22

COPB2, coatomer protein complex, subunit beta-2

PBP2, retinol-binding protein 2

ITPKC, inositol 1,4,5-triphosphate 3-kinase C

HWE, Hardy–Weinberg equilibrium

第 0 頁: [1] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [1] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [2] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [2] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [3] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [3] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [4] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [4] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [5] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [5] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [6] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [6] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [7] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [7] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [8] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [8] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [9] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [9] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [10] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [10] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [11] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [11] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [12] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [12] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [13] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [13] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [14] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [14] 格式化	yiching	2010/11/19 3:36:00 PM

置中

第 0 頁: [15] 刪除	yiching	2010/11/19 3:43:00 PM

					Stage 1 (250 cases and 446 control))				**Stage 2 (208 cases and 366 control))**		
SNP	**Chr**	**Nearest Gene(s)**	**Non-risk Allele**	**Risk Allele**	**RAF (KD)**	**RAF (Control)**	**OR (95% CI)**	***p* value (best)**	**RAF (KD)**	**RAF (Control)**	**OR (95% CI)**
rs1873668	3	MRPS22	C	A	0.054	0.015695	3.549 (1.835-6.865)	3.6824E-05	0.033654	0.021858	1.579 (0.754-3.303)
rs4243399	3	COPB2	A	G	0.054	0.01573	3.541 (1.831-6.85)	3.81619E-05	0.033654	0.021858	1.579 (0.754-3.303)
rs16849083	3	PBP2	G	A	0.054435	0.015766	3.564 (1.842-6.895)	3.46461E-05	0.03202	0.021858	1.497 (0.705-3.178)
rs13128867	4	SLC&A11	C	T	0.333333	0.234694	1.618 (1.268-2.065)	7.41669E-05	0.326923	0.270492	1.316 (1.01-1.716)
rs149481	5	ERAP1	G	T	0.828	0.72809	1.825 (1.377-2.418)	2.24525E-05	0.779412	0.736339	1.261 (0.949-1.676)
rs362794	7	RELN	A	G	0.106	0.042601	2.537 (1.652-3.897)	4.41E-06	0.072464	0.056011	1.308 (0.808-2.119)
rs17113284	14	IGHV	T	C	0.422	0.333708	1.441 (1.152-1.803)	7.21533E-05	0.382212	0.34153	1.182 (0.926-1.509)
rs8005468	14	IGHV	C	T	0.422	0.33296	1.449 (1.158-1.813)	7.97887E-05	0.382212	0.341096	1.184 (0.928-1.511)
rs10129255	14	IGHV	G	A	0.424	0.33296	1.462 (1.168-1.83)	5.17305E-05	0.377404	0.341096	1.16 (0.91-1.48)
rs2007467	14	IGHV	G	A	0.422	0.328442	1.479 (1.181-1.853)	4.65829E-05	0.375	0.342466	1.143 (0.896-1.458)
rs10150241	14	IGHV	T	A	0.422	0.332584	1.447 (1.157-1.81)	5.73004E-05	0.379808	0.345628	1.152 (0.902-1.47)
rs12590667	14	IGHV	G	A	0.423695	0.33296	1.456 (1.164-1.822)	4.80665E-05	0.379808	0.347945	1.141 (0.893-1.457)
rs1568657	15	BTBD1	A	G	0.628514	0.523702	1.542 (1.228-1.936)	8.69426E-05	0.580097	0.525956	1.261 (0.981-1.62)