Gene 648 (2018) 76–81

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Fine-mapping analysis of the MHC region for vitiligo based on a new Han-MHC reference panel☆

Chao Yang[a,b], Juan Wu[a,b], Xuelei Zhang[a,b], Leilei Wen[a,b], Jingying Sun[a,b], Yuyan Cheng[a,b], Xianfa Tang[a,b], Bo Liang[a,b], Gang Chen[a,b], Fusheng Zhou[a,b], Yong Cui[c], Anping Zhang[a,b], Xuejun Zhang[a,b], Xiaodong Zheng[a,b,*], Sen Yang[a,b,*], Liangdan Sun[a,b,*]

[a] Institute of Dermatology and Department of Dermatology of the First Affiliated Hospital, Anhui Medical University, Hefei 230032, China
[b] Key Laboratory of Dermatology, Anhui Medical University, Ministry of Education, Hefei 230032, China
[c] Department of Dermatology, China-Japan Friendship Hospital, Beijing 100029, China

ARTICLE INFO

Keywords:
Fine-mapping
Vitiligo
HLA imputation
Han-MHC reference

ABSTRACT

Vitiligo is an immune-related disease with patchy depigmentation of skin and hair caused by selective destruction of melanocytes. In recent decades, many studies have shown the association between vitiligo and HLA genes; however, the results of Han Chinese are scarce. In this study, we performed a fine-mapping analysis of the MHC region in 2818 Han Chinese subjects through a widely used HLA imputation method with a newly built large-scale Han-MHC reference panel. Three new four-digit HLA alleles (*HLA-DQB1 * 02:02*, *HLA-DQA1 * 02:01* and *HLA-DPB1 * 17:01*) were identified to be associated with the risk of vitiligo, and four previously reported alleles were confirmed. Further conditional analysis revealed that two important variants, HLA-DQβ1 amino acid position 135 (OR = 1.79, $P = 1.87 \times 10^{-11}$) and HLA-B amino acid positions 45–46 (OR = 1.44, $P = 5.61 \times 10^{-11}$), conferred most of the MHC associations. Three-dimension ribbon models showed that the former is located within the β2 domain of the HLA-DQβ1 molecule, and the latter lies in the α1 domain of the HLA-B molecule, while both are involved in specific antigen presenting process. Finally, we summarized all significant signals in the MHC region to clarify their complex relationships, and 8.60% of phenotypic variance could be explained based on all reported variants in Han Chinese so far. Our findings highlight the complex genetic architecture of the MHC region for vitiligo in Han Chinese population and expand our understanding of the roles of HLA coding variants in the etiology of vitiligo.

1. Introduction

Vitiligo is an autoimmune disorder characterized by patchy depigmentation of skin and hair resulting from progressive loss of melanocytes, with an estimated prevalence ranging from 0.5% to 1% in most populations worldwide (Ezzedine et al., 2012). Various theories have been proposed for the explanation of the dysfunction and loss of melanocytes in vitiligo, such as the biochemical hypothesis, neural hypothesis, autoimmune hypothesis, oxidative hypothesis and genetic hypothesis (Ongenae et al., 2003; Schallreuter et al., 2008; Spritz, 2008). Among them, much evidence has indicated that autoimmune processes are the major explanation, and genetic data has already provided important insights (Alkhateeb et al., 2003; Ongenae et al., 2003; Sandoval-Cruz et al., 2011; Spritz, 2012).

The major histocompatibility complex (MHC) region on chromosome 6p21.3, which contains the human leukocyte antigen (HLA) genes, was thought to play an essential role in immune response. The association between vitiligo and HLA genes has been documented for a few decades. In the early stage, case-control studies with a small sample size based on serological typing or DNA typing of HLA have shown an association. However, only the associations of *HLA-A * 02*, *HLA-DRB1 * 04* and *HLA-DRB1 * 07* with vitiligo were somehow consistent among different studies (Zhang et al., 2005; Liu et al., 2007; Spritz, 2008). Since the application of genome-wide association study (GWAS), more and more vitiligo susceptibility loci have been identified, including MHC region (Jin et al., 2010; Quan et al., 2010; Jin et al., 2011; Birlea et al., 2013). In 2010, Jin et al. identified two major MHC signals associated with generalized vitiligo in the European population,

Abbreviations: GWAS, genome-wide association study; MHC, major histocompatibility complex; HLA, human leukocyte antigen; SNP, single-nucleotide polymorphism; LD, linkage disequilibrium; QC, quality control; MAF, minor allele frequency

☆ Any conflict of interest disclosures: None.

* Corresponding authors at: Science and teaching building, 81 Meishan Road, Hefei, 230032, Anhui Province, China.
E-mail addresses: zhengxiaodong@ahmu.edu.cn (X. Zheng), yangsen@medmail.com.cn (S. Yang), ahmusld@163.com (L. Sun).

https://doi.org/10.1016/j.gene.2018.01.053
Received 20 October 2017; Received in revised form 30 December 2017; Accepted 16 January 2018
Available online 02 February 2018
0378-1119/

represented by rs12206499 (*HLA-A ~ HCG9*) and rs532098 (*HLA-DRB1 ~ HLA-DQA1*). Moreover, another two association signals within the MHC region were reported in the Han Chinese population, including rs11966200 (might reflect *HLA-A * 30:01*, *HLA-B * 13:02*, *HLA-C * 06:02* and *HLA-DRB1 * 07:01*) and rs9468925.

Although we have established the association between HLA genes and generalized vitiligo in the Han Chinese subjects by both linkage study and GWAS (Liang et al., 2007; Quan et al., 2010), due to the complexity of the MHC region, it is difficult to fully reveal the genetic architecture of the MHC region for vitiligo. Recently, the newly developed computational strategy SNP2HLA (Jia et al., 2013) enables researchers to perform fine-mapping studies within the MHC region. This has resulted in many achievements in the research of autoimmune diseases of the European population, including idiopathic achalasia (Gockel et al., 2014), psoriasis (Okada et al., 2014), ankylosing spondylitis (Cortes et al., 2015), inflammatory bowel diseases (Goyette et al., 2015), and type 1 diabetes (Hu et al., 2015). However, the susceptibility on HLA genes for diseases such as vitiligo may differ among different populations and continents. To accelerate the fine-mapping strategy in the Han Chinese population, our group constructed the Han-MHC reference panel, which has shown better performance for the imputation of HLA genes (Zhou et al., 2016). Recent studies also implied that the autoimmunity risk conferred by HLA genes was more likely determined by the variants of their amino acid residues (Raychaudhuri et al., 2012; Foo et al., 2013; Kim et al., 2014). However, a fine-mapping study of the MHC region in vitiligo to assess the contribution of amino acid variants of HLA genes has not been implemented. To this end, we conducted a fine-mapping study on 2818 subjects to systematically investigate the coding variants of HLA genes within the MHC region for generalized vitiligo in the Han Chinese population based on a HLA imputation method with our newly built Han-MHC reference panel.

2. Materials and methods

2.1. Study subjects and SNP genotyping data

We used single-nucleotide polymorphism (SNP) genotyping data (Illumina 610-Quad BeadChip, Illumina, San Diego, CA) from our previous extended vitiligo GWAS, including 1117 cases with generalized vitiligo and 1701 healthy controls from the Han Chinese population (Tang et al., 2013). All the vitiligo patients fulfilled the diagnostic criteria of Vitiligo European Task Force (Taieb et al., 2007). Clinical and demographic information was obtained from both patients and controls through a structured questionnaire. All the participating subjects provided written informed consent, and the study obtained approval from the appropriate institutional review board of Anhui Medical University and conformed to the ethical principles of the Declaration of Helsinki. Stringent quality control (QC) procedures were carried out as described elsewhere (Quan et al., 2010). The result of principal component analysis showed that the vitiligo cases and controls were genetically matched and the genomic inflation factor was $\lambda gc = 1.07$.

2.2. HLA imputation

We extracted the genotyped SNPs located on the region of chromosome 6 at 25–34 Mb (UCSC hg19 assembly), and then imputed classical HLA alleles, their corresponding amino acid polymorphisms, and untyped SNPs by using software SNP2HLA package v1.03 with our newly built Han-MHC reference panel. This reference panel consists of 10,689 healthy individuals recruited from different regions of China and contains eight HLA genes (*HLA-A*, *HLA-B*, *HLA-C*, *HLA-DRB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1* and *HLA-DPB1*), amino acid polymorphisms and SNPs. For the post-imputed data, we removed variants with low imputation quality ($r^2 < 0.50$), or those with minor allele frequency (MAF) below 0.01 or significant deviation from Hardy-Weinberg equilibrium ($P < 1.00 \times 10^{-4}$) using the software PLINK version 1.07.

2.3. Single-variant association analysis and conditional analysis

All of the HLA variants were encoded as binary markers as described elsewhere (Jia et al., 2013). To assess the variant effect for vitiligo risk, we applied the logistic regression model with the software PLINK, assuming each allele has an additive effect on the log-odds scale. The statistical results were presented without genomic control correction, as there is negligible evidence of population stratification in our study in terms of the λgc shown above, which is consistent with our previous study (Tang et al., 2013). To detect distinctive and independent variants of HLA alleles and their corresponding amino acid polymorphisms, we applied stepwise conditional analysis by including the most significantly associated variant as a fixed effect in the logistic regression model until no significant HLA variants were observed.

2.4. LD analysis and multivariate regression analysis

We performed two-loci linkage disequilibrium (LD) analysis by calculating the correlation coefficient r^2 using software PLINK. To further determine the effect of each independent variant identified within the MHC region, we performed a multivariate analysis using the logistic model based on the "Forward: LR" method, assuming an additive effect of the alleles tested. The analysis was accomplished in software SPSS version 19.0.

2.5. Calculation of the phenotypic variance explained

We calculated an estimate of the phenotypic variance explained by the set of variants associated with vitiligo as the Pseudo R2 of a logistic regression model (Nagelkerke, 1991) in R. MHC variants identified in the present study and non-MHC variants reported in our previous vitiligo GWA studies were both included in the analysis.

3. Results

3.1. HLA imputation with Han-MHC reference panel

By applying the imputation method with software SNP2HLA, we successfully inferred two-digit and four-digit allele genotypes for the eight HLA genes and their corresponding amino acid polymorphisms, as well as SNPs within the MHC region on 2818 subjects. After stringent QC, we finally obtained 21,580 variants to further evaluate their associations with the risk of vitiligo. To reduce type I error rate in subsequent statistical analysis, we set $P < 2.32 \times 10^{-6}$ as the significance threshold for our analysis after the Bonferroni correction.

3.2. Association of HLA variants with vitiligo susceptibility

We first evaluated the association of both genotyped and imputed SNPs in 1117 vitiligo cases and 1701 controls. We confirmed eight previously reported MHC SNPs in our analysis. Among them, two SNPs achieved genome-wide significance and others reached at least nominal association (Table 1).

In the current analysis, we mainly examined associations of 152 classical HLA alleles and 628 amino acid polymorphisms with vitiligo susceptibility. When we tested the imputed variants, seven four-digit classical HLA alleles (Table 2) and forty amino acid polymorphisms (Supplementary Table 1) achieved the study-wide statistical significance threshold ($P < 2.32 \times 10^{-6}$), and the strongest association signal mapped to *HLA-DQB1* region. Among them, three HLA alleles were newly identified, including *HLA-DQB1 * 02:02* ($P = 1.87 \times 10^{-11}$), *HLA-DQA1 * 02:01* ($P = 2.63 \times 10^{-11}$) and *HLA-*