Original Article | Annals of Dermatology 2023;35(1):32-37 • https://doi.org/10.5021/ad.22.110

Comparison of Risk Allele Frequencies of Psoriasis-Associated Single-Nucleotide Polymorphisms in Different Population Groups

Donghoon Lee[1], Taehan Koo[1], Jiho Park[2], Hyun-Tae Shin[3,4]

[1]Department of Dermatology, Veterans Health Service Medical Center, Seoul, [2]Oaro Dermatology Clinic, Seoul, [3]Department of Dermatology, Inha University College of Medicine, Incheon, [4]Research Center for Controlling Intercellular Communication (RCIC), Inha University College of Medicine, Incheon, Korea

Received May 4, 2022
Revised August 18, 2022
Accepted October 12, 2022

Corresponding Author
Hyun-Tae Shin
Department of Dermatology, Inha University College of Medicine, 27 Inhang-ro, Jung-gu, Incheon 22332, Korea
Tel: +82-32-890-1040
E-mail: hyuntae.shin@inha.ac.kr
https://orcid.org/0000-0003-1799-5860

Background: The prevalence of psoriasis differs by population, and it appears to be more common among Europeans than in East Asians. Recent genome-wide association studies (GWAS) have identified alleles that increase the risk of psoriasis, and these alleles may present different frequencies in different geographic regions.
Objective: We aimed to gain insights into the causes of differences in disease frequencies according to populations and the factors affecting prevalence and pattern differences.
Methods: We collected a total of 147 psoriasis-associated single-nucleotide polymorphisms (SNPs) from the GWAS catalog and compared the allele frequency differences in 27 populations using public population frequency in the 1000 Genomes Project phase 3 (n=2,504) and the Korean Reference Genome Database (n=1,722). Additionally, we calculated the composited genetic risk scores across the population groups.
Results: There were distinct patterns of allele frequencies in different population groups. In many cases, East Asians exhibited allele frequencies opposite to that of Europeans. The genetic risk score was higher in Europeans (average: 0.487) and Americans (average: 0.492) than in East Asians (average: 0.471). The prevalence of psoriasis correlated with the average genetic risk score of the population.
Conclusion: We observed a difference in the allele frequencies of psoriasis-associated SNPs between the studied populations. This result suggests that the difference in the prevalence of psoriasis between population groups can be interpreted to some extent by the genotype.

Keywords: Allele frequency, Polymorphism, single nucleotide, Prevalence, Psoriasis

INTRODUCTION

Psoriasis is a common chronic inflammatory skin disease associated with a variety of comorbidities, including arthritis, obesity, metabolic syndrome, cardiovascular diseases, and depression. It has no complete cure and can have a significant negative impact on the patient's quality of life. Psoriasis affects people of all ages and ethnic groups, and the prevalence of psoriasis varies widely depending on race, climate, and geographic location. The reported prevalence of psoriasis in various countries ranges from 0.09% to 11.43%, making psoriasis a serious global problem affecting at least 100 million people worldwide[1]. Due to the disease burden of psoriasis, the World Health Organization reported the public health impact of psoriasis to provide policy-makers with practical solutions to improve the health care of patients with psoriasis in their populations[1].

Psoriasis is caused by complex etiologies that are affected by genetic predisposition as well as environmental influences. A higher incidence of psoriasis has been reported in families of psoriasis patients; in twin studies, identical twins showed 2 to 3 times higher susceptibility to psoriasis than dizygotic twins

did[2,3]. Classic genome-wide linkage analysis identified 9 chromosomal loci, including psoriasis susceptibility (PSORS) 1~9[4]. The major genetic determinant is PSORS1, which is located within the major histocompatibility complex in the short arm of chromosome 6, containing *HLA-C*, *CCHCR1*, and *CDSN*[5]. Through recent multiple genome-wide association studies (GWAS) in people of European-Caucasian descent, the initial finding of chromosomal susceptible loci was replicated and various new associations have been identified[6,7]. However, the population frequency of risk alleles associated with psoriasis may show differences in each geographic region, which may contribute to differences in the disease prevalence between populations.

In order to gain insights into the relationship between genetic causes and disease prevalence, we collected the single-nucleotide polymorphisms (SNPs) associated with psoriasis risk and compared the allele frequency differences in 27 populations using public population frequency databases[8,9]. We also calculated the composite genetic risk score for psoriasis at the population level and tested the correlation between the average composite risk score and psoriasis prevalence in the population.

MATERIALS AND METHODS

Population frequency comparison of psoriasis-related SNPs

We downloaded full association data (v1.0.2-associations_e96_r2019-10-14) from the GWAS catalog (NHGRI-EBI GWAS catalog, https://www.ebi.ac.uk/gwas/home), and 147 SNPs that were associated with at least one of the "psoriasis" or "psoriasis vulgaris" related trait were included in this study (Supplementary Table 1). We used 1000 Genomes Project phase 3 (n=2,504) and Korean Reference Genome Database (KRGDB) (n=1,722) to obtain the population-level allele frequencies of the SNPs[8,9]. The 1000 Genomes Project phase 3 surveyed genetic variations of people who declared themselves to be healthy from 26 populations worldwide, which can be grouped into African (n=661), American (n=347), East Asian (n=504), European (n=503), and South Asian (n=489) based on their geographic locations and ancestries. The 1000 Genomes Project phase 3 genotype data was downloaded from ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/ (last accessed: January 2021). Since the Korean population was excluded from East Asia in the 1000 Genomes Project, genotyping data for the Korean population was obtained through KRGDB. The population frequency of KRGDB was downloaded from the web-based database (http://coda.nih.go.kr/coda/KRGDB/index.jsp, last accessed: January 2021). To compare the distribution of individual risk allele count in the Korean population, individual genotype results of the 2nd phase of KRGDB (n=1,099) were distributed from the National Human Resource Bank of Korea. KRGDB did not provide clinical information about psoriasis.

Calculation of genetic risk scores using SNPs related to psoriasis

To compare the composite genetic risk score (GRS) of psoriasis, we adopted the following equation suggested by Mao et al.[10]:

$$GRS = \frac{\sum_{i=1}^{I} Xi}{2I}$$

where, "I" refers to the number of psoriasis risk SNPs and "Xi" to the copies of risk alleles at the i^{th} SNP. In one extreme case, if a person has two copies of risk alleles at each risk SNP, then the person's GRS will become "1". On the other hand, if a person has no copies of the risk alleles at each risk SNP, then the person's score will become "0". If copies of effect alleles (0/1/2) are randomly assigned to each SNP, the expected value of the GRS will be "0.5". SNPs with a frequency difference of more than 10% between the total (n=1,722) and the 2nd phase (n=1,099) of KRGDB were excluded from the GRS calculation.

Statistical analysis

The Fisher's exact test was used to assess whether the effect allele at a given SNP was significantly enriched or depleted compared to the global population frequency in the 1000 Genome Project database. For generating heat maps to visualize the allele enrichment/depletion patterns in different populations, Fisher testing *p*-values were adjusted, and then the adjusted *p*-value (*q*-value) was $\log_{10}$ transformed. If the effect allele of an SNP was enriched in a population, then the negative $\log_{10}$ of the enrichment *q*-value (a positive number) was used to represent the SNP in association with that population in a heat map. On the other hand, if the allele of an SNP was depleted in a population, the value of $\log_{10}$ of the depletion *q*-value (a negative number) was used to represent the SNP for that population in the heat map. All statistical analyses and visualization presented here were performed using R (v 3.6.0).

Ethics statement

This study was approved by the Institutional Review Boards of Inha University Hospital (IRB No. 2020-03-025).

RESULTS

Population frequency patterns of psoriasis-risk alleles

In total, 147 psoriasis-associated SNPs were obtained from 13 GWAS studies (Supplementary Table 1). Among these, 10 studies involved Europeans, 2 involved East Asians, and 1 involved mixed populations. Clearly, most populations, except Europeans, were understudied. However, there was no significant difference (Kruskal–Wallis rank sum test, p=0.09) in the SNP frequency among these populations (Supplementary Fig. 1). The results showed that the psoriasis-associated SNPs found in the European studies were also distributed in other populations without significant differences. We then compared the allele frequencies of the psoriasis-associated SNPs in each of the 5 continental groups (Africa, America, East Asia, Europe, and South Asia) and Korea from public databases (Fig. 1, Supplementary Table 2). There were distinct patterns of risk allele frequencies depending on the population groups among the psoriasis-associated SNPs. The hierarchical clustering tree showed the differences among the populations: the European, American, and South Asian populations belonged to one cluster, the East Asian and Korean populations to another cluster, and the African population to a different branch. In addition, the heat map clearly showed that the majority pattern of the allele frequency of East Asians in the 1000 Genomes Project was very similar to that of Koreans. The above results confirm that the population frequency of psoriasis-associated SNPs was significantly different. The representatively frequent SNPs related to psoriasis in Koreans are shown in Supplementary Table 3.

Fig. 2 presents the effect alleles of each East Asian country, specifically Korea. In the 1000 genome project, East Asian populations were divided into Chinese Dai in Xishuangbanna, China (CDX); Han Chinese in Beijing, China (CHB); Southern Han Chinese in China (CHS); Kinh in Ho Chi Minh City, Vietnam (KHV); and Japanese in Tokyo, Japan (JPT). Even though East Asians and Koreans have similar allele frequency patterns for psoriasis, there were some differences in the allele frequency pattern for each detailed population (Supplementary Table 4). The hierarchical clustering tree showed the differences among populations, and the populations most different

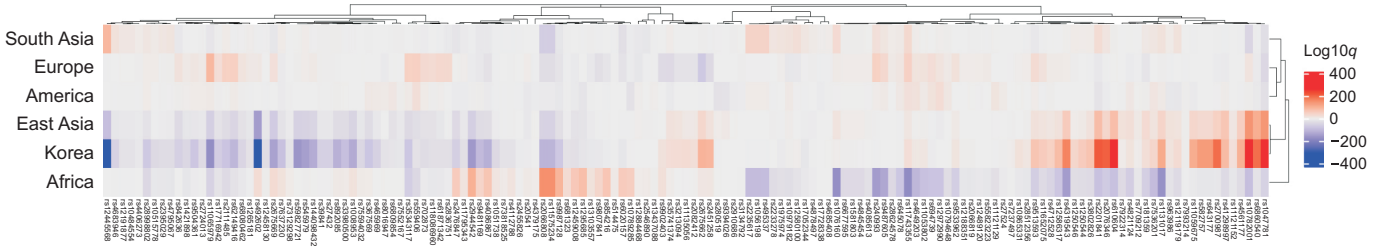

Fig. 1. The heat map of risk alleles of 147 psoriasis-associated single-nucleotide polymorphisms (SNPs) in the population groups. Heat map shows how significantly are the effect alleles enriched or depleted in each population. Each row represents SNPs, and each column shows population groups. Red color indicates that the effect allele is enriched, whereas purple color indicates that the effect allele is depleted.

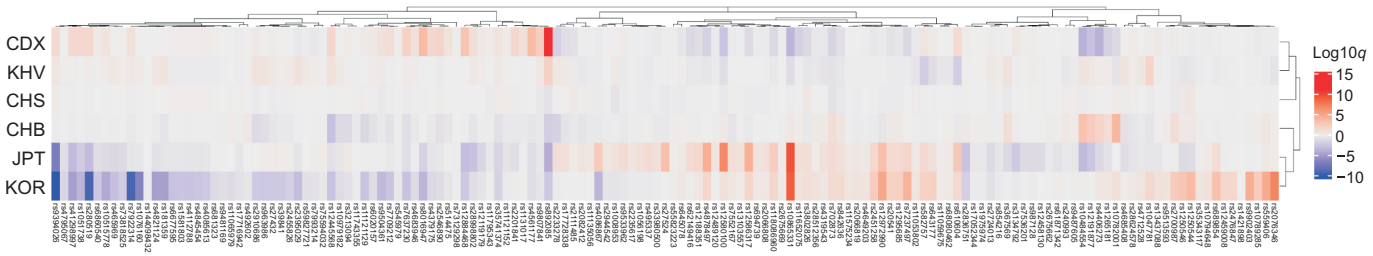

Fig. 2. Heat map of risk alleles of 147 psoriasis-associated single-nucleotide polymorphisms in each East Asian population. Heat map shows how significantly are the effect alleles enriched or depleted in each East Asian population, including Koreans. CDX: Chinese Dai in Xishuangbanna, China, KHV: Kinh in Ho Chi Minh City, Vietnam, CHS: Southern Han Chinese in China, CHB: Han Chinese in Beijing, China, JPT: Japanese in Tokyo, Japan, KOR: Korean.

from Koreans were the CDX, CHS, and KHV that resided in regions that were geographically more distant from Korea than the other East Asian regions (Fig. 2).

Distribution of composite GRS

We calculated the GRS based on copies of effect alleles for psoriasis-associated SNPs (see MATERIALS AND METHODS section). Although the majority of psoriasis-associated SNPs were detected in the European population, we assumed that these variants would also be associated with the pathogenesis in non-European populations. This assumption was partially validated in the analysis of multi-population comparison studies for complex traits, and most associations of GWAS-identifying variants were replicated in various populations[11-13]. The GRS was calculated for each person present in the 1000 Genomes Project (n=2,504) and 2nd phase of KRGDB (n=1,099) using 82 SNPs that reported a risk allele base in the original study (Fig. 3A). Although the CHS population scored higher than some European and other populations, most European and American GRS scores were higher than those of the East Asian population. The average GRS of each of the five continental groups was 0.484 for Africa, 0.492 for America, 0.471 for East Asia, 0.487 for Europe, and 0.505 for South Asia. This pattern correlates with the known prevalence of each population (R=0.519) (Fig. 3B, Supplementary Table 5).

DISCUSSION

This study compared the allele frequency patterns of psoriasis-associated SNPs between different population groups to provide more insights into the relationship between genetic factors and differences in prevalence. The frequency of 147 psoriasis-risk SNPs showed distinct patterns of difference in each population group, while geographically close regions showed relatively similar patterns of allele frequencies.

Examples of SNPs that differ between populations include rs1265181 of the *HCG27* mapped in candidate interval, which is the shortest haplotype segment identified in PSORS1[14]. The risk allele frequency was 18% in Europe and 4% to 5% in the East Asian and Korean populations. The allele frequency in

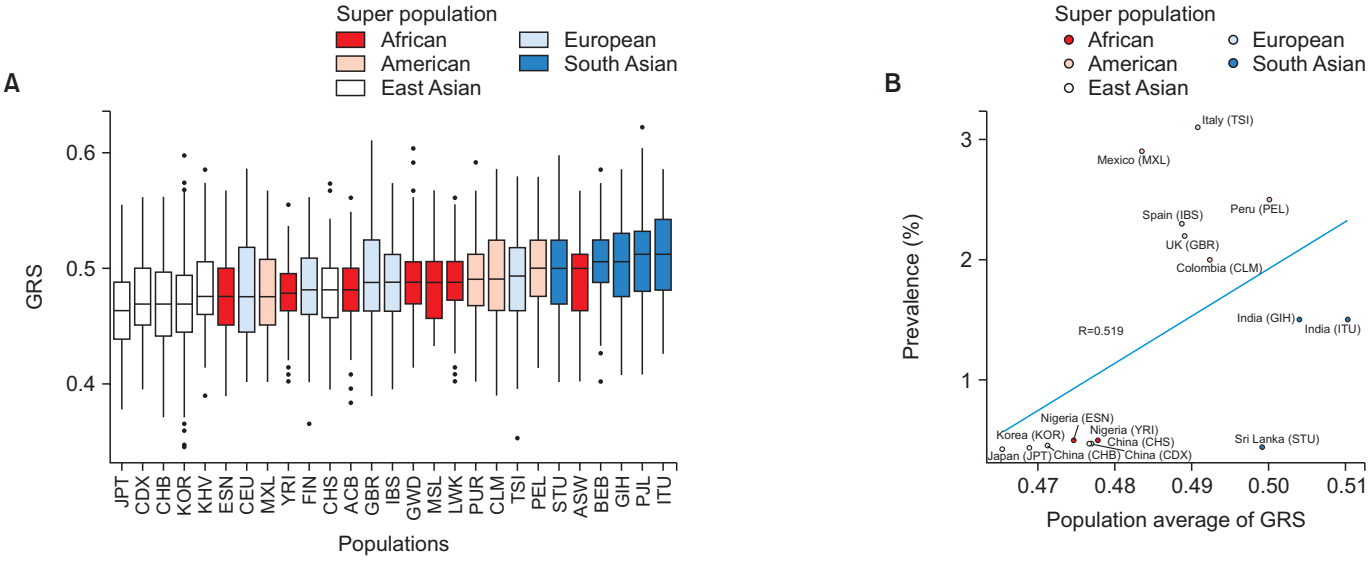

Fig. 3. Genetic risk score of psoriasis-associated single-nucleotide polymorphisms. (A) Distribution of composite GRS for psoriasis; (B) correlation between psoriasis prevalence and average of composite GRS at population level. GRS: genetic risk score, ACB: African Caribbean in Barbados, ASW: African Ancestry in Southwest US, BEB: Bengali in Bangladesh, CDX: Chinese Dai in Xishuangbanna, CEU: Utah residents with Northern and Western European ancestry, CHB: Han Chinese in Beijing, China, CHS: Southern Han Chinese in China, CLM: Colombian in Medellin, Colombia, ESN: Esan in Nigeria, FIN: Finnish in Finland, GBR: British in England and Scotland, GIH: Gujarati Indian in Houston, Texas, GWD: Gambian in Western Division, The Gambia, IBS: Iberian populations in Spain, ITU: Indian Telugu in the UK, JPT: Japanese in Tokyo, Japan, KOR: Korean in the Republic of Korea, KHV: Kinh in Ho Chi Minh City, Vietnam, LWK: Luhya in Webuye, Kenya, MSL: Mende in Sierra Leone, MXL: Mexican Ancestry in Los Angeles, California, PEL: Peruvian in Lima, Peru, PJL: Punjabi in Lahore, Pakistan, PUR: Puerto Rican in Puerto Rico, STU: Sri Lankan Tamil in the UK, TSI: Toscani in Italy, YRI: Yoruba in Ibadan, Nigeria.

Europeans was more than that in East Asians, including Koreans. This gene is a well-known psoriasis-associated gene, which can be interpreted as a causative factor associated with a higher prevalence of psoriasis in Europeans. Another example of rs2395029 mapped in *HCP5* is the SNP with the highest odds ratio among the psoriasis-related SNPs[15]. The odds ratio of rs2395029 was 4.1, which was the highest among the SNPs in this study (Supplementary Table 1). The allele frequencies of rs2395029 of South Asian (9.3%) and European (4.4%) populations were higher than that of East Asian (1.2%) and Korean (0.2%) populations.

Interestingly, The GRS of South Asians highest among all populations, and Africans scored higher than the East Asians did. There is a lack of studies about the prevalence of psoriasis in South Asian and African populations as compared to the European, American, and East Asian populations, which limits its interpretation. Although there are some groups of populations that had patterns that did not fit the well-known prevalence distribution, it can be inferred that this difference in the prevalence between populations can be interpreted to some extent by the genotype.

Most studies on psoriasis in GWAS involve the European population. This report included 13 studies, of which 10 were conducted on the European population. There were evidences showing that variants identified as GWAS are multi-ethnically reproducible. Carlson et al.[13] reported that significant majority of GWAS-identified variants about several traits in European ancestry population have allelic associations in the same direction in non- European ancestry. GWAS of type 2 diabetes in a range of ancestry groups also revealed that most common-variant susceptibility loci are shared across ethnic groups[11]. Also, Wojcik et al.[12] validated the replication of 8,979 SNPs associated various traits from GWAS catalog and they showed that 1,444 associations were replicated at the $p<0.05$ significance threshold. Although some genetic factors could have different effects in each population, we hypothesized that most SNPs identified in European populations also have similar effects on non-European populations.

Two studies were conducted on the East Asian population, including 5 SNPs (rs1265181, rs2233278, rs3213094, rs4085613, and rs9394026). The majority of these SNPs were reported to have a higher allele frequency in East Asians than in Americans and Europeans. Furthermore, GRS calculated by these 5 SNPs did not replicate the positive correlations between disease prevalence and GRS (data not shown). This could be attributed to the limited number of psoriasis-risk SNPs. Therefore, further studies, including East Asian, African, and South Asian populations outside of Europe, are expected to provide results that are more accurate. The fact that the calculated GRS was higher in South Asians than in Europeans, even though the SNPs used were identified in the European population, leads to the assumption that South Asians may have a higher prevalence than what is known. South Asian population-based studies are expected to identify more psoriasis-related SNPs and genes.

This study had some limitations. First, 1722 individuals belonged to the Korean reference genome, which was a large number, assuming that the number of 26 populations belonging to the 1000 genome project was about 100, ranging from 61 to 113. Therefore, the statistical significance for Koreans may have been highly evaluated. Second, when calculating GRS, we did not consider the effect size of the SNPs. This was because some studies did not report effect sizes, and the same SNPs reported different effect sizes across studies. If a sophisticated model considering the effect size is applied, results that are more accurate will be obtained.

Our study showed substantial population differentiation in allele frequencies in psoriasis-related SNPs in different populations. Furthermore, a correlation between GRS and the prevalence of psoriasis was observed. We observed differences in allele frequencies associated with SNPs related to psoriasis between East Asian and other populations, suggesting that this difference in the prevalence between populations can be interpreted to some extent by the genotype.

ACKNOWLEDGMENT

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. 2021M3E5D1A02015185). This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2021R1A5A2031612). This work was supported by INHA UNIVERSITY Research Grant. This study was conducted with bioresources from National Biobank of Korea, the Korea Disease Control and Prevention Agency, Republic of Korea (KBN-2021-005).

SUPPLEMENTARY MATERIALS

Supplementary data can be found via http://anndermatol.org/src/sm/ad-22-110-s001.pdf.

CONFLICTS OF INTEREST

The authors have nothing to disclose.

FUNDING SOURCE

None.

DATA SHARING STATEMENT

All data of the 1000 Genomes Project phase 3 are available from ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/ (last accessed: January 2021). Allele frequency data of KRGDB are available from http://coda.nih.go.kr/coda/KRGDB/index.jsp (last accessed: January 2021). Individual VCF of KRGDB cannot be shared publicly because the VCF were distributed from the National Human Resource Bank of Korea after reviewing distribution application (Distributed Number: KBN-2021-005, distribution guide site: http://nih.go.kr/contents.es?mid=a50402040100, last accessed: March 2021).

ORCID

Donghoon Lee, https://orcid.org/0000-0002-6341-7351
Taehan Koo, https://orcid.org/0000-0003-1053-9387
Jiho Park, https://orcid.org/0000-0003-2125-2488
Hyun-Tae Shin, https://orcid.org/0000-0003-1799-5860

REFERENCES

1. World Health Organization. Global report on psoriasis. Geneva: World Health Organization, 2016.
2. Huang YH, Kuo CF, Huang LH, Hsieh MY. Familial aggregation of psoriasis and co-aggregation of autoimmune diseases in affected families. J Clin Med 2019;8:115.
3. Lønnberg AS, Skov L, Skytthe A, Kyvik KO, Pedersen OB, Thomsen SF. Heritability of psoriasis in a large twin sample. Br J Dermatol 2013;169:412-416.
4. Bowcock AM, Krueger JG. Getting under the skin: the immunogenetics of psoriasis. Nat Rev Immunol 2005;5:699-711. Erratum in: Nat Rev Immunol 2005;5:826.
5. Nestle FO, Kaplan DH, Barker J. Psoriasis. N Engl J Med 2009;361:496-509.
6. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet 2012;44:291-296.
7. Tsoi LC, Stuart PE, Tian C, Gudjonsson JE, Das S, Zawistowski M, et al. Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. Nat Commun 2017;8:15382.
8. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature 2015;526:68-74.
9. Jung KS, Hong KW, Jo HY, Choi J, Ban HJ, Cho SB, et al. KRGDB: the large-scale variant database of 1722 Koreans based on whole genome sequencing. Database (Oxford) 2020;2020:baz146. Erratum in: Database (Oxford) 2020;2020:baaa030.
10. Mao L, Fang Y, Campbell M, Southerland WM. Population differentiation in allele frequencies of obesity-associated SNPs. BMC Genomics 2017;18:861.
11. DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium; Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; Mexican American Type 2 Diabetes (MAT2D) Consortium; Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in muylti-Ethnic Samples (T2D-GENES) Consortium, Mahajan A, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat Genet 2014;46:234-244.
12. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, et al. Genetic analyses of diverse populations improves discovery for complex traits. Nature 2019;570:514-518.
13. Carlson CS, Matise TC, North KE, Haiman CA, Fesinmeyer MD, Buyske S, et al. Generalization and dilution of association results from European GWAS in populations of non-European ancestry: the PAGE study. PLoS Biol 2013;11:e1001661.
14. Zhang XJ, Huang W, Yang S, Sun LD, Zhang FY, Zhu QX, et al. Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21. Nat Genet 2009;41:205-210.
15. Liu Y, Helms C, Liao W, Zaba LC, Duan S, Gardner J, et al. A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. PLoS Genet 2008;4:e1000041.