**General information**

**GWAS summary statistics of autoimmune and allergic diseases in BioBank Japan.**

|  |  |
| --- | --- |
| **Genotyping array**: | The Illumina HumanOmniExpressExome BeadChip or a combination of the Illumina HumanOmniExpress and HumanExome BeadChips. |
| **Sample QC**: | We excluded samples with low genotyping call rates (call rate < 98%). We included samples of the estimated East Asian ancestry using PCA. |
| **Variant QC**: | We excluded variants with (1) genotyping call rate < 99%, (2) P value for Hardy–Weinberg equilibrium < 1.0 × 10−6, and (3) number of heterozygotes < 5. |
| **Phasing and imputation**: | Eagle and Minimac3 software |
| **Imputation reference**: | A combined panel of WGS data from the BioBank Japan project (N=1,037) and 1KGP p3v5 ALL (N=2,504). |
| **Post-imputation QC**: | We here report results of whole-genome imputed variants with Rsq > 0.7 and MAF > 0.005. |
| **Association test**: | SAIGE software was used with age,　sex, and top five principal components as covariates. |
| **Meta-analysis:** | RE2C software was used for the multi-trait meta-analysis adjusting for sample overlap between GWAS summary data. We also applied Metasoft to calculate heterogeneity index I2 and P-value based on Cochran's Q test. |

**GWAS summary statistics of autoimmune and allergic diseases in UK biobank.**

|  |  |
| --- | --- |
| **Genotyping array**: | Either the Applied Biosystems UK BiLEVE Axiom Array or the Applied Biosystems UK Biobank Axiom Array. |
| **Sample QC**: | We analyzed individuals of white British genetic ancestry as determined by the PCA-based sample selection criteria. |
| **Variant QC**: | We removed markers that failed quality control in more than one batch, had a greater than 5% overall missing rate, and had a MAF of less than 0.0001. |
| **Phasing and imputation**: | IMPUTE4 software |
| **Imputation reference**: | A combination of the Haplotype Reference Consortium, UK10K, and 1000 Genomes Phase 3 reference panels. |
| **Post-imputation QC**: | We here report results of whole-genome imputed variants with Rsq > 0.7 and MAF > 0.005. |
| **Association test**: | SAIGE software was used with age,　sex, and top five principal components as covariates. RE2C software was used for the meta-analysis to integrate the individual GWAS summary statistic. |
| **Meta-analysis:** | RE2C software was used for the multi-trait meta-analysis adjusting for sample overlap between GWAS summary data. We also applied Metasoft to calculate heterogeneity index I2 and P-value based on Cochran's Q test. |

**GWAS summary statistics of autoimmune and allergic diseases in BioBank Japan and UK biobank.**

Here, we provide summary statistics of the cross-population and multi-trait meta-analysis result of BioBank Japan and UK Biobank.

|  |  |
| --- | --- |
| **Genotyping array**: | BBJ: The Illumina HumanOmniExpressExome BeadChip or a combination of the Illumina HumanOmniExpress and HumanExome BeadChips. UKB: Either the Applied Biosystems UK BiLEVE Axiom Array or the Applied Biosystems UK Biobank Axiom Array. |
| **Sample QC**: | BBJ: We excluded samples with low genotyping call rates (call rate < 98%). We included samples of the estimated East Asian ancestry using PCA. UKB: We analyzed individuals of white British genetic ancestry as determined by the PCA-based sample selection criteria. |
| **Variant QC**: | BBJ: We excluded variants with (1) genotyping call rate < 99%, (2) P value for Hardy–Weinberg equilibrium < 1.0 × 10−6, and (3) number of heterozygotes < 5. UKB: We removed markers that failed quality control in more than one batch, had a greater than 5% overall missing rate, and had a MAF of less than 0.0001. |
| **Phasing and imputation**: | BBJ: Eagle and Minimac3 software UKB: IMPUTE4 software |
| **Post-imputation QC**: | We here report results of whole-genome imputed variants with Rsq > 0.7 and MAF > 0.005. |
| **Association test**: | SAIGE software was used with age,　sex, and top five principal components as covariates. RE2C software was used for the meta-analysis to integrate the individual GWAS summary statistic. |
| **Meta-analysis:** | RE2C software was used for the multi-trait meta-analysis adjusting for sample overlap between GWAS summary data. We also applied Metasoft to calculate heterogeneity index I2 and P-value based on Cochran's Q test. |

**Uploaded files**

**GWAS summary statistics of autoimmune and allergic diseases in BioBank Japan.**

|  |  |
| --- | --- |
| File name | Descriptions |
| GWASsummary\_{TRAITNAME}\_BBJ\_ARD2022.txt.gz | Summary results for autosomal variants |

**GWAS summary statistics of autoimmune and allergic diseases in UK biobank.**

|  |  |
| --- | --- |
| File name | Descriptions |
| GWASsummary\_{TRAITNAME}\_UKB\_ARD2022.txt.gz | Summary results for autosomal variants |

**GWAS summary statistics of autoimmune and allergic diseases in BioBank Japan and UK biobank.**

|  |  |
| --- | --- |
| File name | Descriptions |
| GWASsummary\_{TRAITNAME}\_ALL\_ARD2022.txt.gz | Summary results for autosomal variants |

Allergy: meta-analysis of bronchial asthma [BA], pollinosis [PO], and atopic dermatitis [AD]  
Autoimmune: meta-analysis of rheumatoid arthritis [RA], Graves' disease [GD], and type 1 diabetes [T1D]

**Columns**

Individual trait

| **#** | **column name** |  | **Descriptions** |
| --- | --- | --- | --- |
| 1 | CHR |  | chromosome |
| 2 | POS |  | position (hg19) |
| 3 | Allele1 |  | REF allele |
| 4 | Allele2 |  | ALT allele (This allele is the effect allele.) |
| 5 | AF\_Allele2 |  | allele frequency of Allele2 (ALT) |
| 7 | Rsq |  | Rsq value in imputation |
| 8 | BETA |  | effect size of Allele2 |
| 9 | SE |  | standard error of BETA |
| 10 | Tstat |  | score statistic |
| 11 | p.value |  | P value with SPA (suddle point approximation) applied |
| 12 | p.value.NA |  | P value when SPA is not applied |
| 13 | Is.SPA.converge |  | whether SPA is converged or not |
| 10 | varT |  | estimated variance of score statistic with sample related incorporated |
| 11 | varTstar |  | variance of score statistic without sample related incorporated |
| 12 | AF.Cases |  | allele frequency of Allele2 in cases |
| 13 | AF.Controls |  | allele frequency of Allele2 in controls |

| **#** | **column name** | **Descriptions** |
| --- | --- | --- |
| 1 | CHR | chromosome |
| 2 | POS | position (hg19) |
| 3 | A1 | REF allele |
| 4 | A2 | ALT allele (This allele is the effect allele.) |
| 5 | Nstudy | number of GWAS summary statistic used in the meta-analysis |
| 6 | LSs | Lin-Sullivan’s effect size (fixed effect) |
| 7 | LSse | Lin-Sullivan’s standard error (fixed effect |
| 8 | LSp | Lin-Sullivan’s P value (fixed effect) |
| 9 | RE2Cs1 | RE2C statistic mean effect part |
| 10 | RE2Cs2 | RE2C statistic heterogeneity part |
| 11 | RE2Cp\* | RE2C\* P value |
| 12 | RE2Cp | RE2C P value |
| 13 | I\_SQUARE | heterogeneity index I2 |
| 14 | Q | Cochran's Q statistic |
| 15 | PVALUE\_Q | P value based on Q |

Meta-analysis