**Finding PCSK9 SNPs affecting LDL-C levels in Chinese populations**

Finding PCSK9 variants affecting LDL-C level in Chinese population

LDL-C (Low-Density Lipoprotein Cholesterol) level in bloodstream is a significant health factor. It has been well established that high LDL-C leads to increased cardiovascular risk. Given the measurement of LDL-C concentration, and the genotype data, we can perform association studies in Chinese cohorts to identify the relevant genes, and the corresponding genetic variants.

PCSK9 (proprotein convertase subtilisin/kexin type 9) is one of the relevant genes. The convertase binds to LDLR (Low-Density Lipoprotein Receptor) and causes its degradation. While LDLR removes LDL-C from circulation on the surface of liver, inactivation of the protein induces higher LDL-C levels and increased risk of cardiovascular events.

PCSK9 was intensively studied in Caucasian populations. Drugs blocking the protein have been developed to reduce LDL-C level in blood.

Given some Caucasian PCSK9 variants identified as associated with LDL-C are absent from or present at low frequencies in Chinese, we seek to identify additional PCSK9 variants associated with LDL-C levels in Chinese populations.

The genotype data set is the CKB genotyping phase1-2, QC stage 3.

Only autosome SNPs were used, leaving 636670 variations and 32205 subjects in the set.

Two types of LDL-C levels were used in this study. The indirect values were estimated using the Friedewald equation (?). The direct measurements were from enzymatic assay. (?)

The direct LDL-C measurements were from the file LDL-c\_biochem\_data.xlsx in directory K:\kadoorie\Groups\Genetics\PROJECTS\PCSK9.

The indirect LDL-C values and subjects ages were obtained using the CKB data request form snapshot 10.

Subject stratification was based on the file GWAS\_SNPdata\_samples.xlsx from directory K:\kadoorie\Groups\Genetics\Data Archive\Project Sample Lists\Lists\

1. PCA (Principal Component Analysis)

PCA components are often employed as covariates in association studies to account for population structure. A LD(Linkage Disequilibrium)-free subset of 120201 SNPs was selected. Before PCA we used Plink to estimate pairwise IBD (identity-by-descent) distances between all pairs of subjects. Pairs of subjects with IBD values higher than 0.05 were considered as related.

We used an iterative approach to remove the relatives. Given the network of related individuals, the degrees of connections were calculated. The most connected node (subject) was ejected. And the degrees were updated then. The process was iterated till no left subjects were related to each other.

6990 subjects were discharged. PCA was performed using the others and projected onto the discharged one later. Plink and eigensoft produced very similar results. We used the Plink components.

2. (stratification details here)

3. linear regression

We employed Plink for genome-wide linear regression. Genetic variants were regressed against raw LDL-C levels in each stratum. Covariates including the region code, sex, age and the 10 leading principal components were applied. For the smallest stratum (stratum 3), we used only sex, age and the 2 leading components.

4. RINT (Rank Inverse-Normal Transformed)

For each stratum, the raw LDL-C values were regressed against the region code, sex and age. We rank inverse-normal transformed the residuals to obtained phenotype sets following the standard normal distribution. In the linear regression of these RINT LDL-C values, we used only the 10 leading principal components as covariates (2 leading components only for stratum 3).

5. meta-analysis

Genome-wide meta-analysis were performed using Metal.

Because of the heterogeneity between the direct and indirect LDL-C values, two meta-analysis were executed.

First we combined the beta and the standard deviation of beta estimated in the linear regression of the first five strata, using the inverse-variance-weighted fixed-effects scheme. The result of the association of the direct LDL-C values were then combined with the linear regression results of stratum 6 applying the same scheme.

6. meta results

7. condition on SNPs.

Summary:

Given the measurements of LDL-C (Low-Density Lipoprotein Cholesterol) concentration, and the genotype data, we can perform association studies in Chinese cohorts to identify the relevant genes, and the corresponding genetic variants.

PCSK9 (proprotein convertase subtilisin/kexin type 9) is one of the relevant genes. The protein binds to LDLR (Low-Density Lipoprotein receptor) on the surface of liver and causes it to degrade. Less active LDLR removing LDL-C from circulation leads to higher LDL-C concentration and increased risk of cardiovascular events. The gene was intensively studied with Caucasian populations. Drugs blocking the protein have been developed to reduce LDL-C concentration in blood.

Given some Caucasian PCSK9 variants identified as associated with LDL-C are absent from or present at low frequencies in Chinese, we seek to identify additional PCSK9 variants associated with LDL-C levels in Chinese populations.

The original genotype data is the stage3 set from /kuser/shared/data/GWASphase12 on nc2. Only autosome SNPs are used, leaving 636670 variants and 32205 people in the set.

The direct LDL-C measurements are from LDL-c\_biochem\_data.xlsx in K:\kadoorie\Groups\Genetics\PROJECTS\PCSK9.

The indirect LDL-C data are from the data request form CKB snapshot 10.

1. PCA

A LD (Linkage Disequilibrium) - free subset of 120201 SNPs was used for pairwise identity-by-descent (IBD) estimates for all pairs of subjects. Pairs of subjects with IBD values higher than 0.05 were considered as related. In the network of related subjects, the most connected subject would be removed. Then the connection degrees were updated. The process was iterated till no subjects were related to each other. 6990 subjects were ejected during the iterations. Principle component analysis (PCA) was performed using plink on the rest and projected to the removed subjects. Eigensoft produced very similar results.

1. Sample stratification and linear regression

Given GWAS\_SNPdata\_samples.xlsx from K:\kadoorie\Groups\Genetics\Data Archive\Project Sample Lists\Lists\, we grouped the subjects into 6 strata.

We have direct measurements of LDL-C concentration for the first 5 strata, and indirect measures for the 6th.

We performed genome-wide linear regression of the genetic variants against the LDL-C measurements with covariates including the region code, sex, age and the 10 leading principal components using plink. For the smallest stratum (stratum 3), only sex, age and the 2 leading components were employed as covariates.

For each stratum, raw LDL-C values were regressed against region, sex and age. We rank inverse-normal transformed (RINT) the residuals to obtained phenotype sets following the standard normal distribution. In the linear regression of these RINT LDL-C values, we used only the 10 leading principal components as covariates (2 leading components only for the stratum 3).

The inflation factors (Lambda) of the GWAS are acceptable, range between 0.99 and 1.03.

|  |
| --- |
| Stratum 1, raw LDL-C |
| Stratum 2, raw LDL-C |
| Stratum 3, raw LDL-C |

|  |
| --- |
| Stratum 4, raw LDL-C |
| Stratum 5, raw LDL-C |
| Stratum 6, raw LDL-C (indirect) |

|  |
| --- |
| Stratum 1, RINT LDL-C |
| Stratum 2, RINT LDL-C |
| Stratum 3, RINT LDL-C |

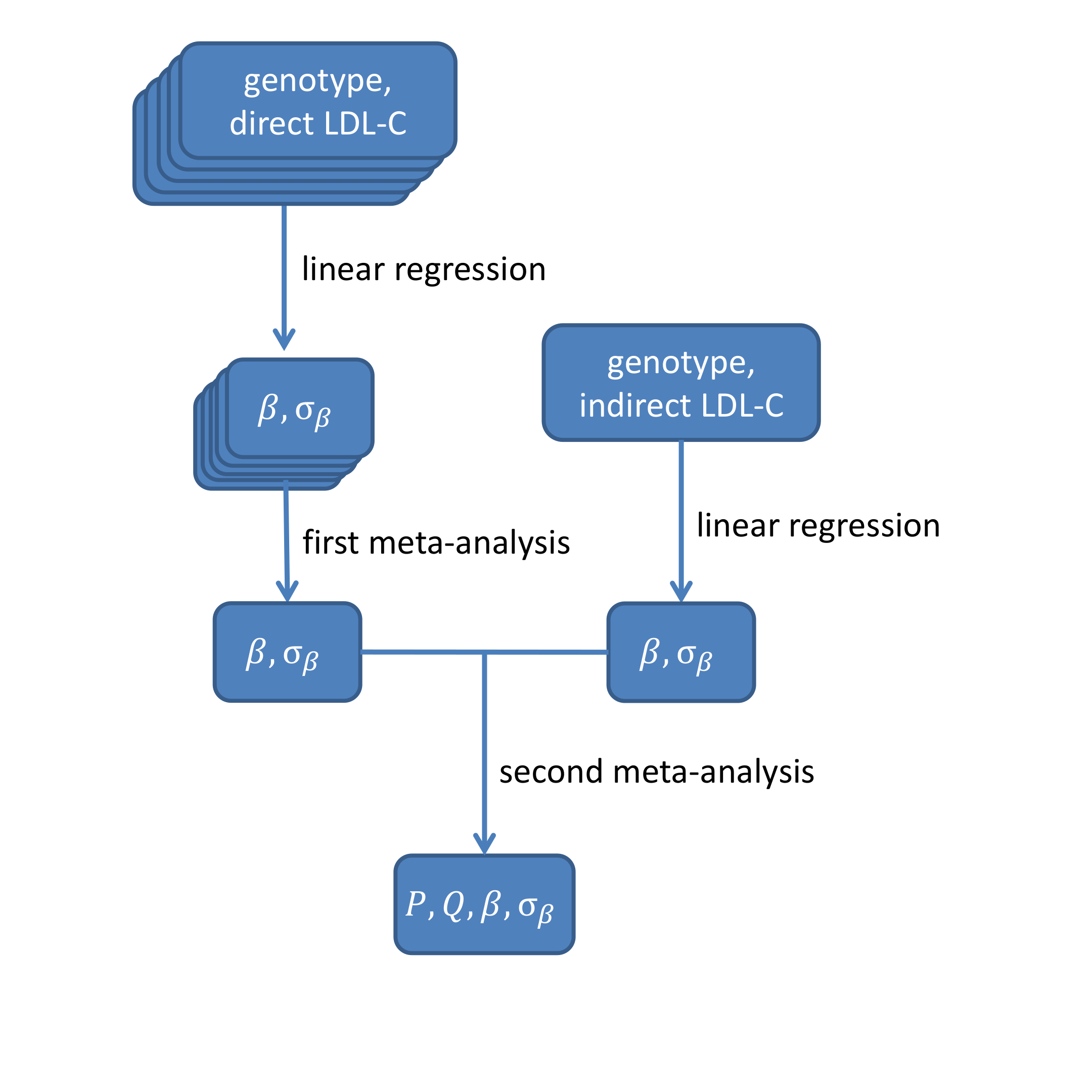
|  |
| --- |
| Stratum 4, RINT LDL-C |
| Stratum 5, RINT LDL-C |
| Stratum 6, RINT LDL-C |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **stratum** | **Ascertainment** | **size** | **LDL-C type** | **Lambda\_raw** | **Lambda\_RINT** |
| stratum\_1 | ICH | 4762 | direct | 1.007485 | 1.019263 |
| stratum\_2 | IS | 5210 | direct | 1.008424 | 1.010303 |
| stratum\_3 | SAH | 167 | direct | 0.9920914 | 0.9990672 |
| stratum\_4 | MI/IHD | 1265 | direct | 0.9907005 | 0.998135 |
| stratum\_5 | Control | 6687 | direct | 1.024951 | 1.026377 |
| stratum\_6 | Control/resurvey 2 | 4177 | indirect | 1.019736 | 1.02971 |

1. Meta analysis

We used Metal for meta-analysis. The beta and standard errors estimated in the linear regression of the first five strata were combined using the inverse-variance-weighted fixed-effects scheme.

The stratum 6 is the only set with indirect LDL-C measurements. Its linear regression result was included with the output of the first meta-analysis in the second round.



……

The meta-analysis results:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SNP** | **CHR** | **BP** | **MAF** | **A1** | **A2** | **Beta** | **SE\_Beta** | **Q** | **direction** | **P** |
| AX-83389438 | 1 | 55509585 | 1.36E-02 | T | C | -3.86E-01 | 0.0302 | 0.938 | ----- | 1.45E-37 |
| AX-11576926 | 1 | 55630151 | 0.1165 | T | C | 0.0535 | 0.011 | 0.6708 | +++++ | 1.06E-06 |
| AX-39912161 | 1 | 55513061 | 1.95E-01 | T | C | -4.01E-02 | 0.0088 | 0.7789 | ----- | 4.82E-06 |
| AX-31642001 | 1 | 55517883 | 0.2873 | C | G | -0.035 | 0.0077 | 0.2967 | ----- | 5.61E-06 |
| AX-31642169 | 1 | 55521242 | 0.2477 | T | G | -0.036 | 0.0081 | 0.8867 | ----- | 8.52E-06 |
| AX-31641677 | 1 | 55509939 | 0.06613 | T | C | 0.0539 | 0.0141 | 0.8634 | ++-++ | 0.0001352 |
| AX-11541856 | 1 | 55529187 | 0.06007 | A | G | -0.0525 | 0.0145 | 0.4883 | ----- | 0.0002852 |
| AX-39912159 | 1 | 55512995 | 0.07861 | A | G | -0.0465 | 0.013 | 0.8858 | --+-- | 0.0003583 |
| AX-11629248 | 1 | 55719166 | 0.2107 | T | G | 0.0293 | 0.0086 | 0.7209 | +++++ | 0.0006577 |
| AX-31641243 | 1 | 55498982 | 0.1261 | A | G | -0.0354 | 0.0106 | 0.1539 | ----- | 0.000828 |
| AX-51209582 | 1 | 55498949 | 0.3514 | T | C | 0.0237 | 0.0073 | 0.8529 | +++++ | 0.0012 |

1. Conditional analysis

To find the independently-associated SNPs, we iteratively performed the linear regression and the meta-analysis. After each iteration, the SNP with the lowest P value found in the previous meta-analysis was to be added to the list of SNP hits. The linear regression would then be performed conditioned on the genotypes of these SNPs. We stopped when no additional SNP hits with P value less than 10-4 can be found or conditioned upon.

|  |
| --- |
| The original meta-analysis of direct LDL-C. The vertical lines are 10-4 and 0.05 accordingly. |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SNP | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p | | AX-83389438 | 1 | 55509585 | 1.36E-02 | T | C | -3.86E-01 | 0.0302 | 0.938 | ----- | 1.45E-37 | | AX-11576926 | 1 | 55630151 | 0.1165 | T | C | 0.0535 | 0.011 | 0.6708 | +++++ | 1.06E-06 | | AX-39912161 | 1 | 55513061 | 1.95E-01 | T | C | -4.01E-02 | 0.0088 | 0.7789 | ----- | 4.82E-06 | | AX-31642001 | 1 | 55517883 | 0.2873 | C | G | -0.035 | 0.0077 | 0.2967 | ----- | 5.61E-06 | | AX-31642169 | 1 | 55521242 | 0.2477 | T | G | -0.036 | 0.0081 | 0.8867 | ----- | 8.52E-06 | | AX-31641677 | 1 | 55509939 | 0.06613 | T | C | 0.0539 | 0.0141 | 0.8634 | ++-++ | 0.0001352 | | AX-11541856 | 1 | 55529187 | 0.06007 | A | G | -0.0525 | 0.0145 | 0.4883 | ----- | 0.0002852 | | AX-39912159 | 1 | 55512995 | 0.07861 | A | G | -0.0465 | 0.013 | 0.8858 | --+-- | 0.0003583 | | AX-11629248 | 1 | 55719166 | 0.2107 | T | G | 0.0293 | 0.0086 | 0.7209 | +++++ | 0.0006577 | | AX-31641243 | 1 | 55498982 | 0.1261 | A | G | -0.0354 | 0.0106 | 0.1539 | ----- | 0.000828 | |

After the initial round, the linear regression would be performed condition on the top SNP AX-83389438.

|  |
| --- |
| Direct LDL-C, conditioned on AX-83389438 |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SNP | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p | | AX-31642001 | 1 | 55517883 | 0.2873 | C | G | -0.0425 | 0.0077 | 0.2819 | ----- | 3.41E-08 | | AX-11576926 | 1 | 55630151 | 0.1165 | T | C | 0.0587 | 0.0109 | 0.7909 | +++++ | 7.84E-08 | | AX-39912161 | 1 | 55513061 | 0.1947 | T | C | -0.0466 | 0.0087 | 0.8085 | ----- | 1.01E-07 | | AX-31642169 | 1 | 55521242 | 0.2477 | T | G | -0.0428 | 0.0081 | 0.9458 | ----- | 1.19E-07 | | AX-51209582 | 1 | 55498949 | 0.3514 | T | C | 0.031 | 0.0073 | 0.8995 | +++++ | 2.19E-05 | | AX-31641677 | 1 | 55509939 | 0.06613 | T | C | 0.0588 | 0.0141 | 0.8474 | ++-++ | 2.96E-05 | | AX-11629248 | 1 | 55719166 | 0.2107 | T | G | 0.0344 | 0.0086 | 0.8304 | +++++ | 5.90E-05 | | AX-39912159 | 1 | 55512995 | 0.07861 | A | G | -0.0519 | 0.013 | 0.9019 | --+-- | 6.30E-05 | | AX-31641243 | 1 | 55498982 | 0.1261 | A | G | -0.0407 | 0.0106 | 0.199 | ----- | 0.0001156 | |

The top SNP hit AX-31642001 was added to the list of SNPs to be conditioned upon.

|  |
| --- |
| Direct LDL-C, conditioned on AX-83389438 and AX-31642001 |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SNP | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p | | AX-11541856 | 1 | 55529187 | 0.06007 | A | G | -0.0674 | 0.0148 | 0.5517 | ----- | 5.07E-06 | | AX-39912161 | 1 | 55513061 | 0.1947 | T | C | -0.0327 | 0.0096 | 0.6193 | ----- | 0.0006279 | | AX-11576926 | 1 | 55630151 | 0.1165 | T | C | 0.0396 | 0.0121 | 0.7732 | +++++ | 0.001094 | | AX-39912661 | 1 | 55529828 | 0.1115 | T | C | -0.0363 | 0.0117 | 0.1707 | ----- | 0.001845 | | AX-12930867 | 1 | 55486064 | 0.4125 | A | G | -0.0214 | 0.0071 | 0.08162 | +-++- | 0.002652 | | AX-64101281 | 1 | 55837600 | 0.01583 | T | C | 0.0833 | 0.0282 | 0.4171 | +++++ | 0.00313 | | AX-39911557 | 1 | 55484810 | 0.4638 | T | C | -0.0197 | 0.0071 | 0.06822 | +-+-- | 0.005147 | | AX-50958112 | 1 | 55496556 | 0.4706 | T | C | 0.019 | 0.0071 | 0.04406 | -+-++ | 0.007468 | | AX-51209582 | 1 | 55498949 | 0.3514 | T | C | 0.0205 | 0.0077 | 0.8707 | +++++ | 0.00773 | | AX-11532730 | 1 | 55541174 | 0.106 | A | G | 0.0317 | 0.0119 | 0.4849 | +++++ | 0.007954 | |

|  |
| --- |
| Direct LDL-C, conditioned on AX-83389438, AX-31642001 and AX-11541856 |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SNP | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p | | AX-31642169 | 1 | 55521242 | 0.2477 | T | G | -0.047 | 0.0126 | 0.31 | ----- | 0.0001938 | | AX-12930867 | 1 | 55486064 | 0.4125 | A | G | -0.0217 | 0.0071 | 0.09975 | +-++- | 0.002412 | | AX-39912161 | 1 | 55513061 | 0.1947 | T | C | -0.0282 | 0.0096 | 0.5948 | ----- | 0.003415 | | AX-51209582 | 1 | 55498949 | 0.3514 | T | C | 0.0222 | 0.0077 | 0.855 | +++++ | 0.00387 | | AX-39911557 | 1 | 55484810 | 0.4638 | T | C | -0.0197 | 0.0071 | 0.07759 | +-+-- | 0.005333 | | AX-11447635 | 1 | 55766150 | 0.1156 | T | C | -0.0308 | 0.0113 | 0.619 | --+-- | 0.006578 | | AX-50958112 | 1 | 55496556 | 0.4706 | T | C | 0.0187 | 0.0071 | 0.05692 | -+-++ | 0.00862 | | AX-39912909 | 1 | 55541050 | 0.04024 | T | C | 0.0449 | 0.0174 | 0.7185 | ++-++ | 0.009886 | |

|  |
| --- |
| RINT direct LDL-C, original meta-analysis |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SNP | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p | | AX-83389438 | 1 | 55509585 | 0.01357 | T | C | -0.621 | 0.0455 | 0.6727 | ----- | 1.91E-42 | | AX-11576926 | 1 | 55630151 | 0.1165 | T | C | 0.0811 | 0.0166 | 0.7462 | +++++ | 9.82E-07 | | AX-39912161 | 1 | 55513061 | 0.1947 | T | C | -0.0638 | 0.0132 | 0.9448 | ----- | 1.44E-06 | | AX-31642001 | 1 | 55517883 | 0.2873 | C | G | -0.0507 | 0.0116 | 0.3767 | ----- | 1.34E-05 | | AX-31642169 | 1 | 55521242 | 0.2477 | T | G | -0.053 | 0.0122 | 0.9855 | ----- | 1.48E-05 | | AX-31641677 | 1 | 55509939 | 0.06613 | T | C | 0.0881 | 0.0213 | 0.9513 | ++-++ | 3.66E-05 | | AX-11541856 | 1 | 55529187 | 0.06007 | A | G | -0.0855 | 0.0219 | 0.412 | ----- | 9.14E-05 | | AX-39912159 | 1 | 55512995 | 0.07861 | A | G | -0.0749 | 0.0197 | 0.9052 | --+-- | 0.0001397 | |

The same process was applied on RINT LDL-C measurements.

|  |
| --- |
| RINT direct LDL-C, conditioned on AX-83389438 |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SNP | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p | | AX-39912161 | 1 | 55513061 | 0.1947 | T | C | -0.0742 | 0.0132 | 0.9572 | ----- | 1.91E-08 | | AX-11576926 | 1 | 55630151 | 0.1165 | T | C | 0.0892 | 0.0165 | 0.8511 | +++++ | 6.49E-08 | | AX-31642001 | 1 | 55517883 | 0.2873 | C | G | -0.0626 | 0.0116 | 0.3141 | ----- | 7.25E-08 | | AX-31642169 | 1 | 55521242 | 0.2477 | T | G | -0.0636 | 0.0122 | 0.9856 | ----- | 1.81E-07 | | AX-31641677 | 1 | 55509939 | 0.06613 | T | C | 0.0957 | 0.0212 | 0.9184 | ++-++ | 6.54E-06 | | AX-51209582 | 1 | 55498949 | 0.3514 | T | C | 0.0492 | 0.011 | 0.9208 | +++++ | 8.20E-06 | | AX-39912159 | 1 | 55512995 | 0.07861 | A | G | -0.0835 | 0.0196 | 0.9109 | --+-- | 1.98E-05 | | AX-11629248 | 1 | 55719166 | 0.2107 | T | G | 0.0546 | 0.0129 | 0.8881 | +++++ | 2.43E-05 | | AX-31641243 | 1 | 55498982 | 0.1261 | A | G | -0.0637 | 0.0159 | 0.299 | ----- | 6.40E-05 | | AX-11541856 | 1 | 55529187 | 0.06007 | A | G | -0.0762 | 0.0218 | 0.4945 | ----- | 0.0004647 | |

|  |
| --- |
| RINT direct LDL-C, conditioned on AX-83389438 and AX-39912161 |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SNP | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p | | AX-11576926 | 1 | 55630151 | 0.1165 | T | C | 0.0723 | 0.0169 | 0.8426 | +++++ | 1.93E-05 | | AX-11541856 | 1 | 55529187 | 0.06007 | A | G | -0.0762 | 0.0217 | 0.4009 | ----- | 0.0004606 | | AX-31642001 | 1 | 55517883 | 0.2873 | C | G | -0.0432 | 0.0127 | 0.1764 | -++-- | 0.0006774 | | AX-11629248 | 1 | 55719166 | 0.2107 | T | G | 0.0429 | 0.0131 | 0.866 | +++++ | 0.001108 | | AX-31642169 | 1 | 55521242 | 0.2477 | T | G | -0.0432 | 0.0133 | 0.956 | --+-- | 0.001169 | | AX-64101281 | 1 | 55837600 | 0.01583 | T | C | 0.131 | 0.0423 | 0.2685 | +++++ | 0.00197 | | AX-12930867 | 1 | 55486064 | 0.4125 | A | G | -0.031 | 0.0107 | 0.1735 | --++- | 0.003827 | |

|  |
| --- |
| RINT direct LDL-C, conditioned on AX-83389438, AX-39912161 and AX-11576926 |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SNP | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p | | AX-64101281 | 1 | 55837600 | 0.01583 | T | C | 0.1295 | 0.0424 | 0.3135 | +++++ | 0.002235 | | AX-11541856 | 1 | 55529187 | 0.06007 | A | G | -0.0659 | 0.0219 | 0.3565 | ----- | 0.002579 | | AX-12930867 | 1 | 55486064 | 0.4125 | A | G | -0.0306 | 0.0107 | 0.1651 | --++- | 0.004382 | | AX-31639191 | 1 | 55447704 | 0.06497 | T | G | -0.0552 | 0.0214 | 0.315 | --+-- | 0.009942 | | AX-39911557 | 1 | 55484810 | 0.4638 | T | C | -0.0267 | 0.0107 | 0.1962 | +-++- | 0.01233 | |

SNPs selected using the two sets of LDL-C measurements:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SNP | rs\_id | BP | MAF | Function | raw\_beta | raw\_P | RINT\_beta | RINT\_P | direction | org\_raw\_p |
| AX-83389438 | rs151193009 | 55509585 | 0.0136 | R93→C | -0.3864 | 1.45E-37 | -0.6210 | 1.91E-42 | ----- | 1.45E-37 |
| AX-11576926 | rs6663252 | 55630151 | 0.1165 | intron |  |  | 0.0723 | 1.93E-05 | +++++ | 1.06E-06 |
| AX-39912161 | rs10888897 | 55513061 | 0.1947 | intron |  |  | -0.0742 | 1.91E-08 | ----- | 4.82E-06 |
| AX-31642001 | rs624612 | 55517883 | 0.2873 | intron | -0.0425 | 3.41E-08 |  |  | ----- | 5.61E-06 |
| AX-31642169 | rs471705 | 55521242 | 0.2477 | intron | -0.0470 | 1.94E-04 |  |  | ----- | 8.52E-06 |
| AX-31641677 | rs4275490 | 55509939 | 0.0661 |  |  |  |  |  | ++-++ | 1.35E-04 |
| AX-11541856 | rs505151 | 55529187 | 0.0601 | G670→E | -0.0674 | 5.07E-06 |  |  | ----- | 2.85E-04 |
| AX-39912159 | rs2495478 | 55512995 | 0.0786 |  |  |  |  |  | --+-- | 3.58E-04 |

1. Some side notes

The leading SNP hit AX-83389438 (rs151193009) was found only in the CHB, CHS, JPT and KHV cohorts in the 1000 Genomes Project (phase 3) data set, with the MAF ranges between 1-2%.

Two SNPs in this region are also available in the SNP 384 panel.

AX-11150762 (rs11206510)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| META | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p |
| raw\_dir\_ldl\_meta\_0 | 1 | 55496039 | 0.0564 | T | C | 0.0308 | 0.0151 | 0.4555 | +++-+ | 0.04063 |
| raw\_dir\_ldl\_meta\_1 | 1 | 55496039 | 0.0564 | T | C | 0.0351 | 0.015 | 0.4839 | +++-+ | 0.01933 |
| raw\_dir\_ldl\_meta\_2 | 1 | 55496039 | 0.0564 | T | C | 0.0287 | 0.0151 | 0.4496 | +++-+ | 0.05688 |
| raw\_dir\_ldl\_meta\_3 | 1 | 55496039 | 0.0564 | T | C | 0.0245 | 0.0151 | 0.4445 | +++-+ | 0.1048 |
| Rint\_dir\_ldl\_meta\_0 | 1 | 55496039 | 0.0564 | T | C | 0.0444 | 0.0228 | 0.5506 | +++-+ | 0.05113 |
| rint\_dir \_ldl\_meta\_1 | 1 | 55496039 | 0.0564 | T | C | 0.0514 | 0.0227 | 0.5647 | +++-+ | 0.02332 |
| rint\_dir \_ldl\_meta\_2 | 1 | 55496039 | 0.0564 | T | C | 0.0132 | 0.0238 | 0.5117 | +-+-+ | 0.579 |
| rint\_dir \_ldl\_meta\_3 | 1 | 55496039 | 0.0564 | T | C | 0.0166 | 0.0238 | 0.5042 | +-+-+ | 0.4857 |

AX-39911995 (rs2479409)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| META | CHR | BP | MAF | a1 | a2 | meta\_beta | meta\_se | meta\_q | meta\_dir | meta\_p |
| raw\_dir\_ldl\_meta\_0 | 1 | 55504650 | 0.3181 | A | G | -0.0073 | 0.0075 | 0.5959 | ---+- | 0.3301 |
| raw\_dir\_ldl\_meta\_1 | 1 | 55504650 | 0.3181 | A | G | -0.0156 | 0.0075 | 0.5641 | ---+- | 0.0378 |
| raw\_dir\_ldl\_meta\_2 | 1 | 55504650 | 0.3181 | A | G | -0.0124 | 0.0076 | 0.5271 | ---+- | 0.1011 |
| raw\_dir\_ldl\_meta\_3 | 1 | 55504650 | 0.3181 | A | G | -0.0148 | 0.0076 | 0.4978 | ---+- | 0.05061 |
| rint\_dir \_ldl\_meta\_0 | 1 | 55504650 | 0.3181 | A | G | -0.0117 | 0.0114 | 0.7847 | ---+- | 0.3039 |
| rint\_dir \_ldl\_meta\_1 | 1 | 55504650 | 0.3181 | A | G | -0.0249 | 0.0113 | 0.7438 | ---+- | 0.02813 |
| rint\_dir \_ldl\_meta\_2 | 1 | 55504650 | 0.3181 | A | G | -0.0096 | 0.0117 | 0.7401 | ---+- | 0.4134 |
| rint\_dir \_ldl\_meta\_3 | 1 | 55504650 | 0.3181 | A | G | -0.0118 | 0.0117 | 0.7811 | --++- | 0.3156 |