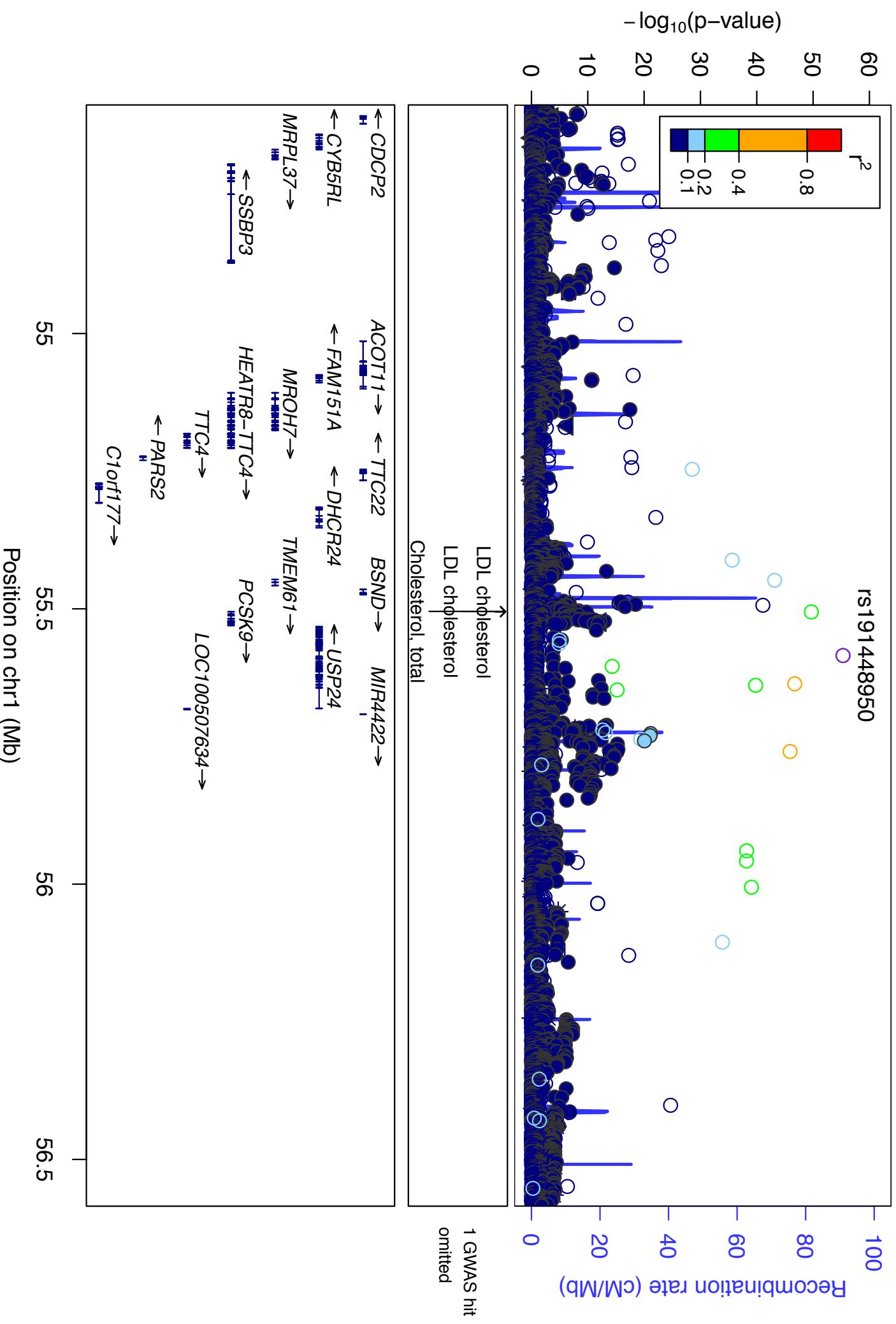
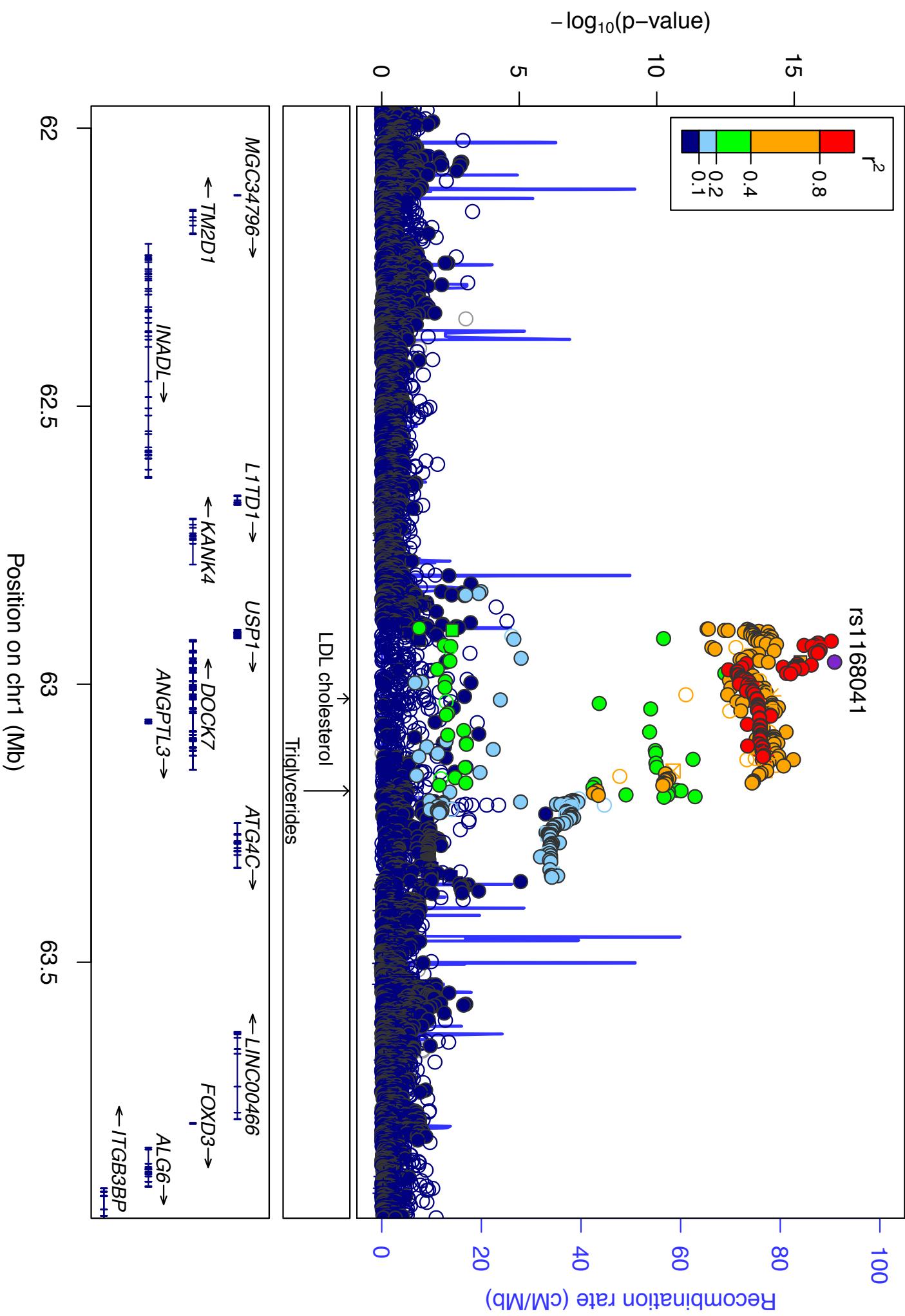


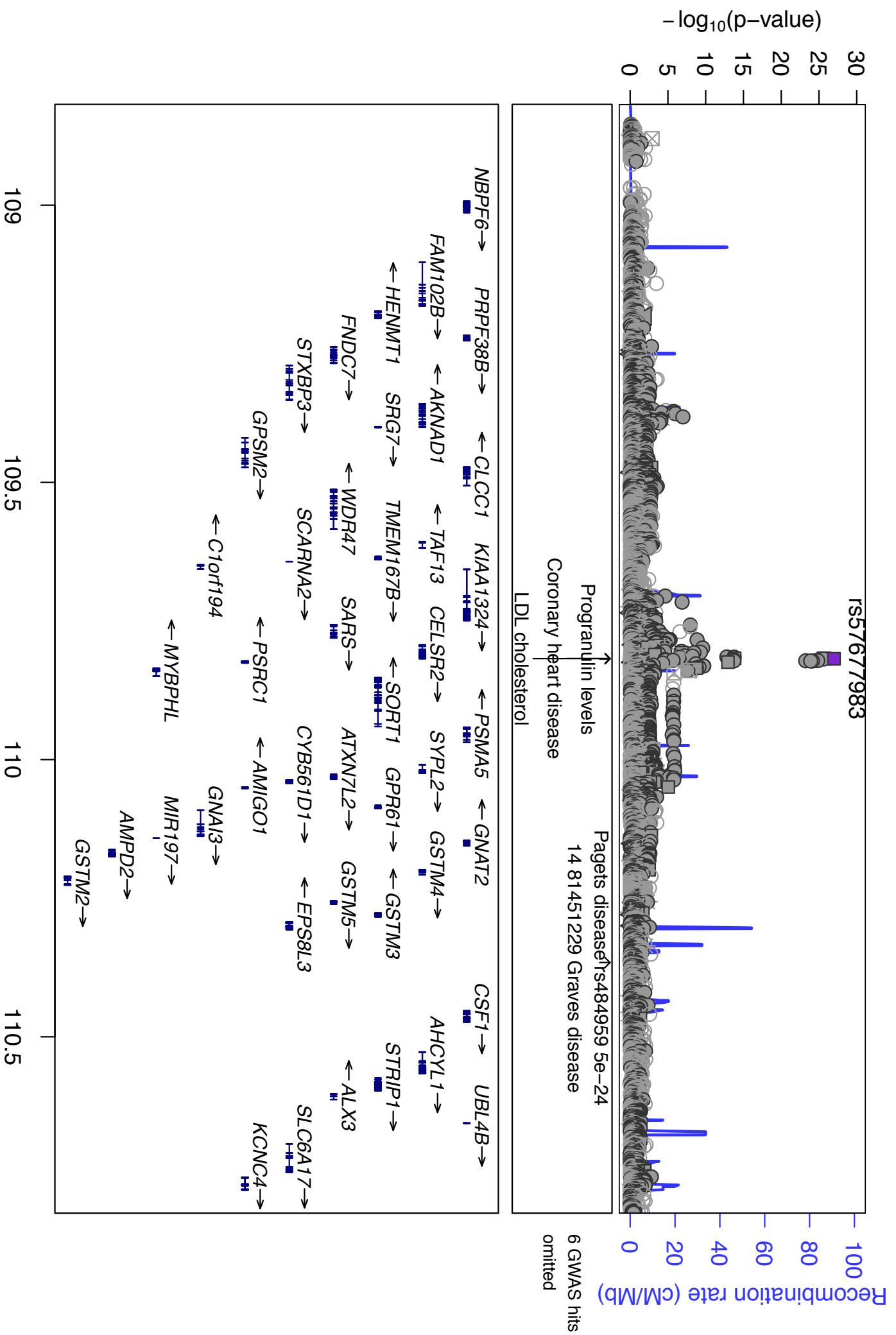
Supplementary Figure 1. Regional Manhattan plots for all 62 lead loci. Annotation key for the symbols is presented below. Correlation structure with the lead variant is presented as sliding color and key is present in all figures. Correlation structure could not be calculated for insertions/deletions and therefore those figures contain only grey symbols.

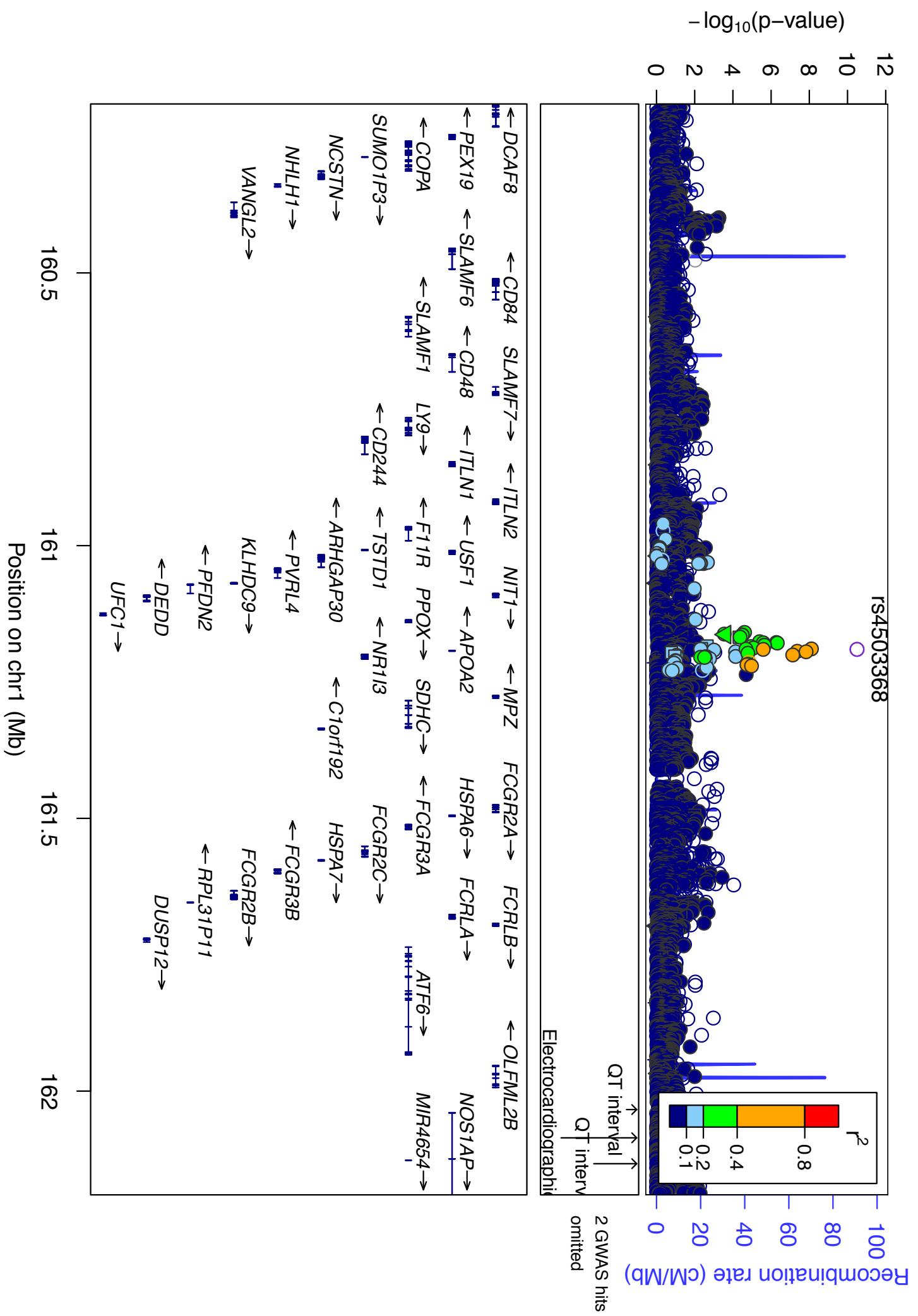
annotation key

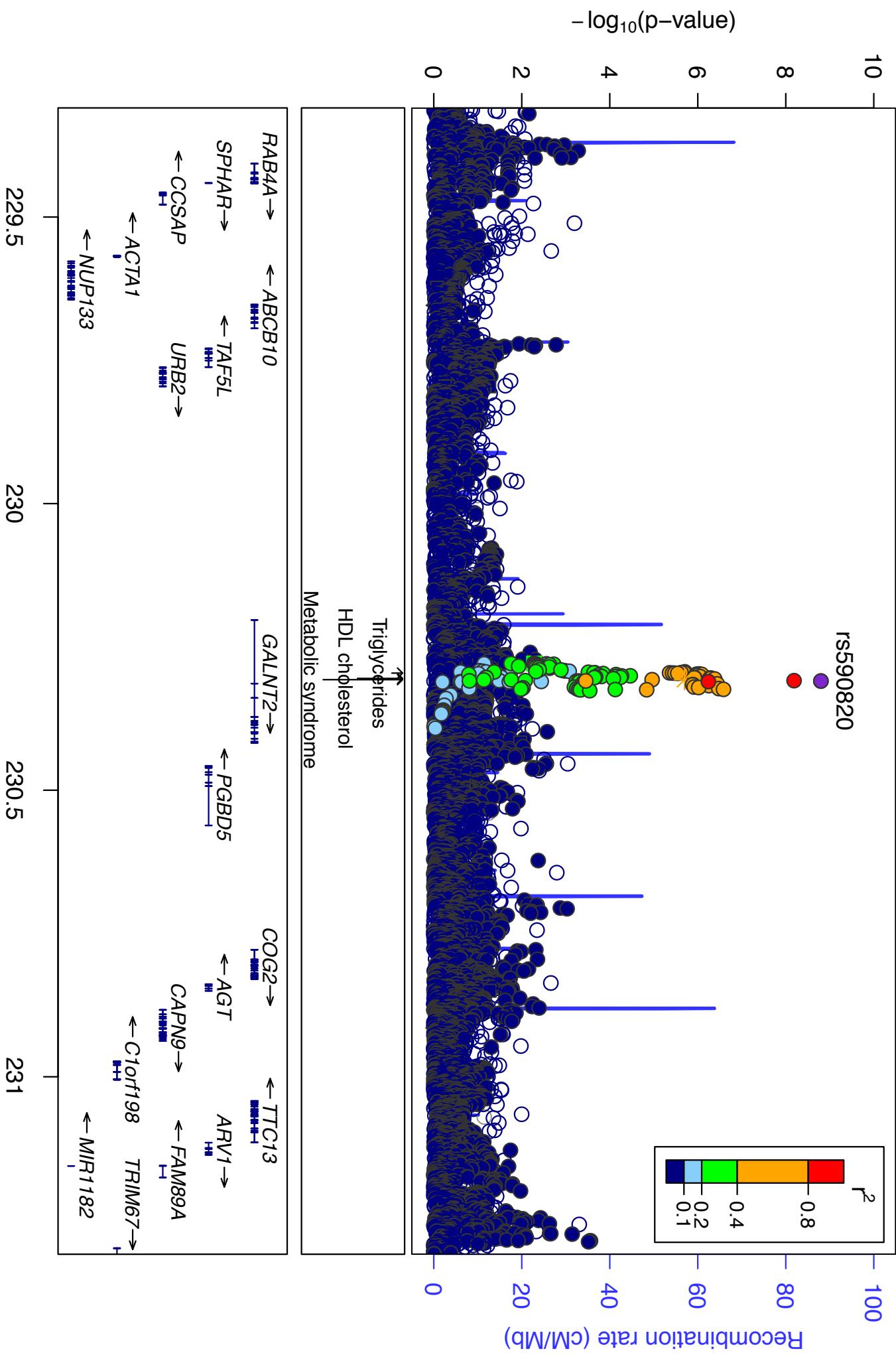
framestop	▲
splice	▲
nonsyn	▼
coding	■
utr	□
tfbscons	*
mcs44placental	☒
no annotation	●
none	○

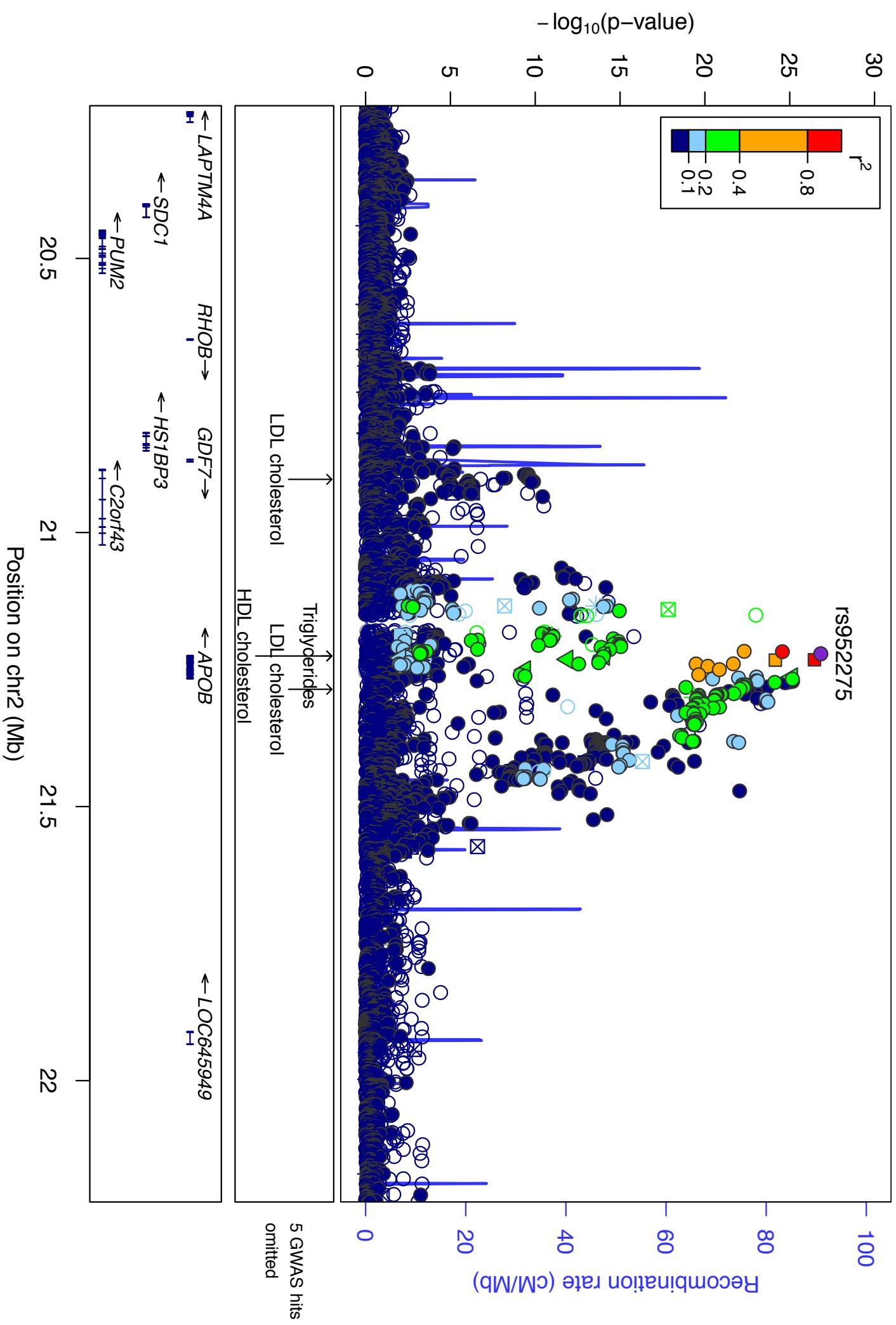


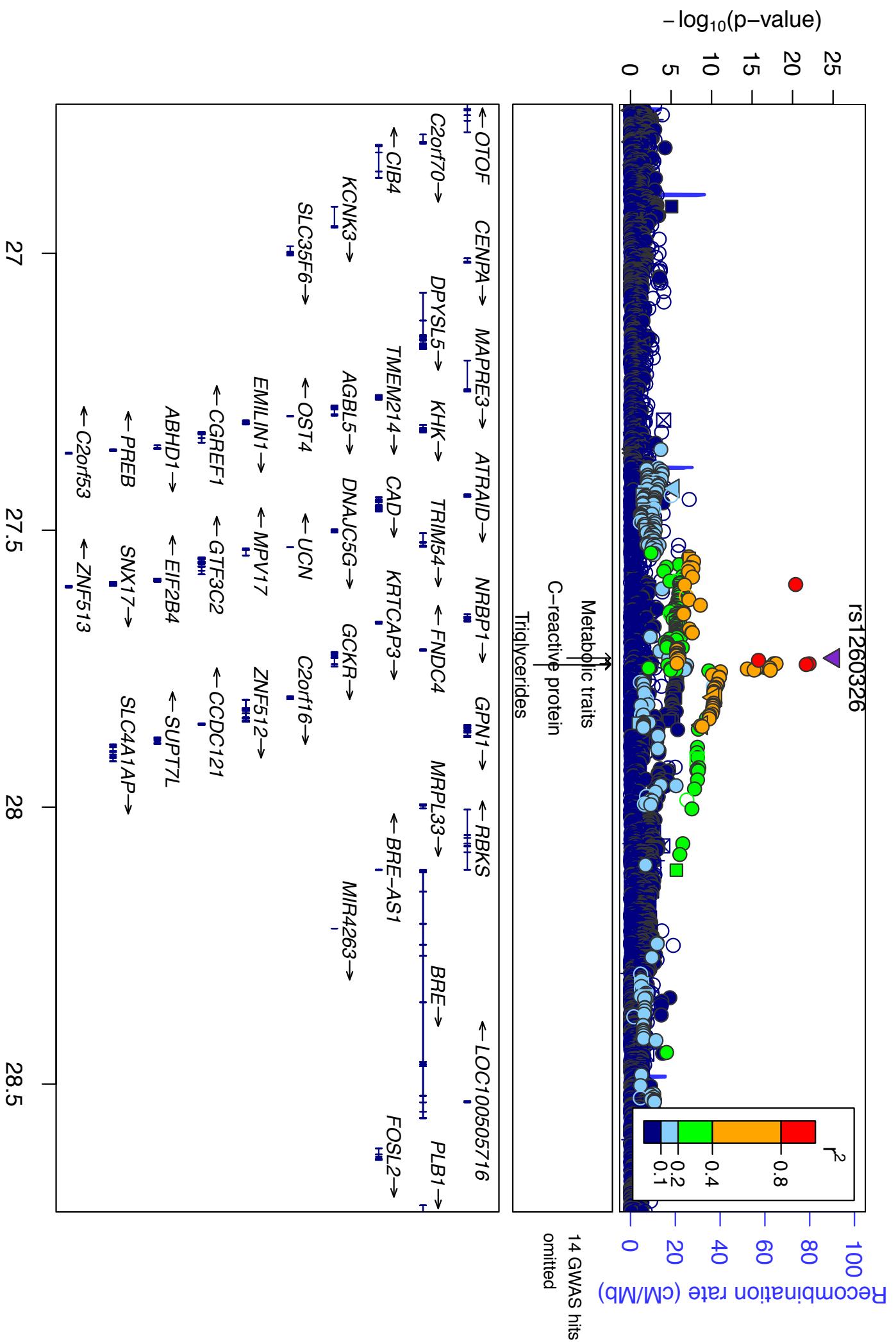












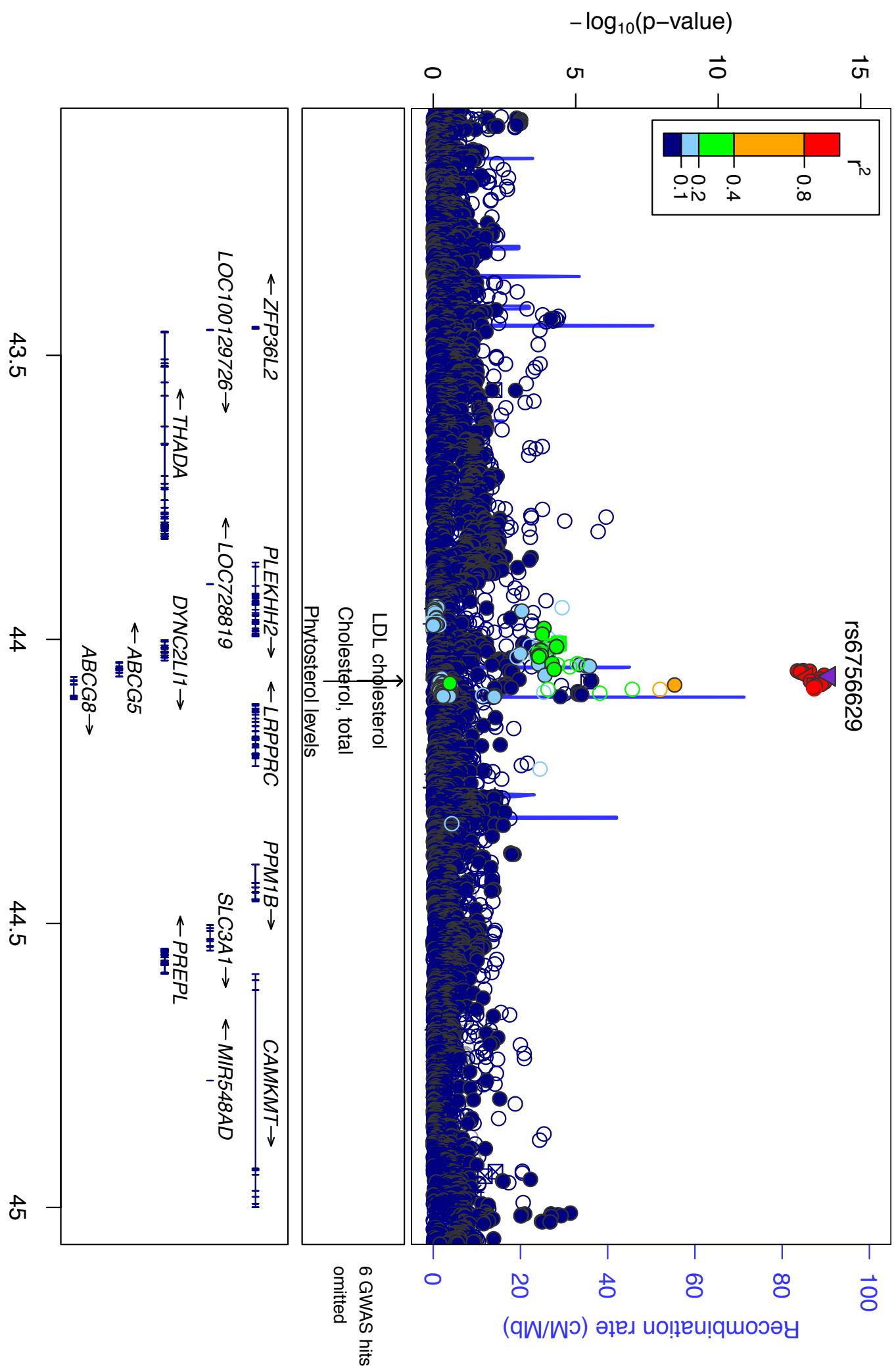
27

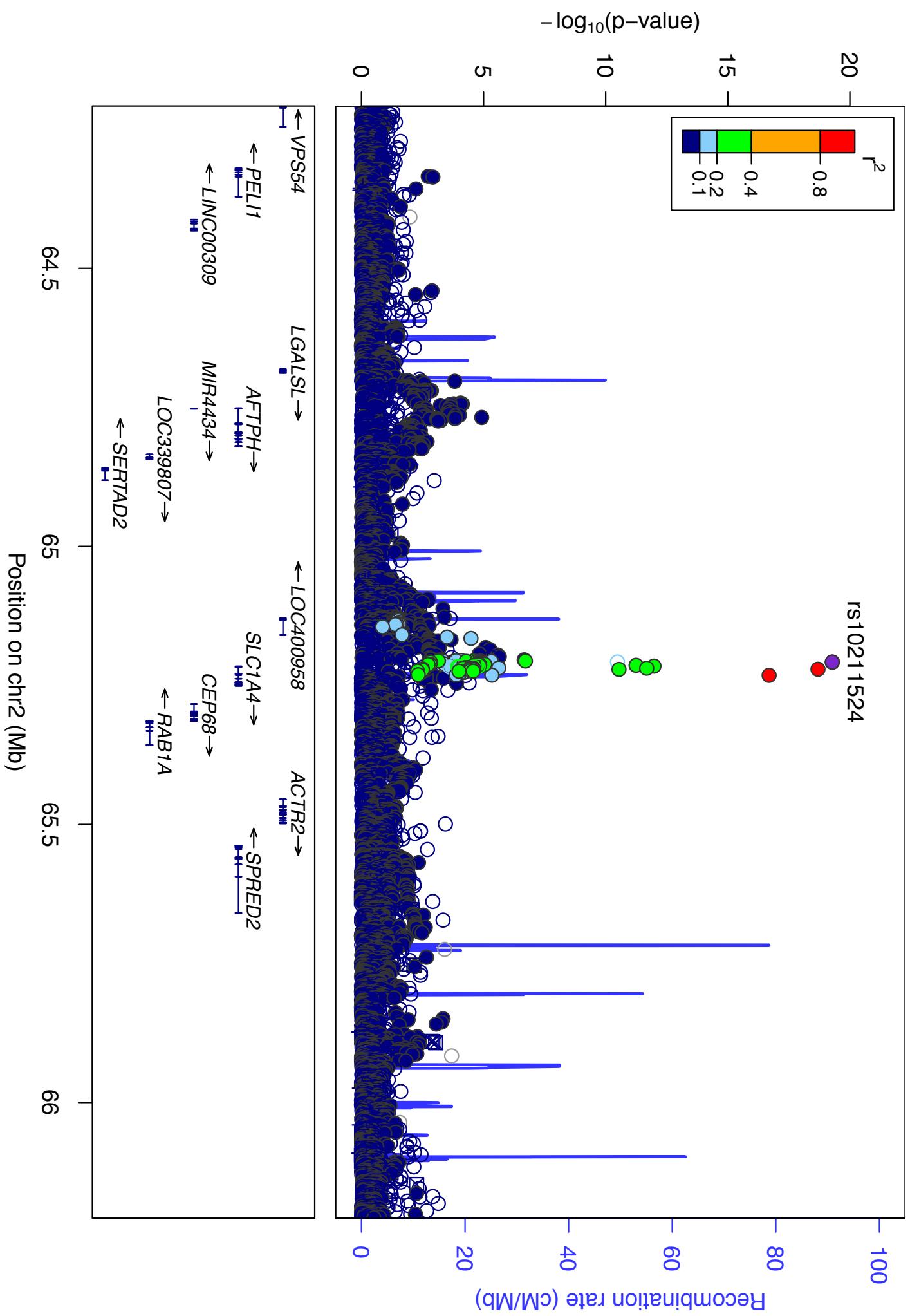
27.5

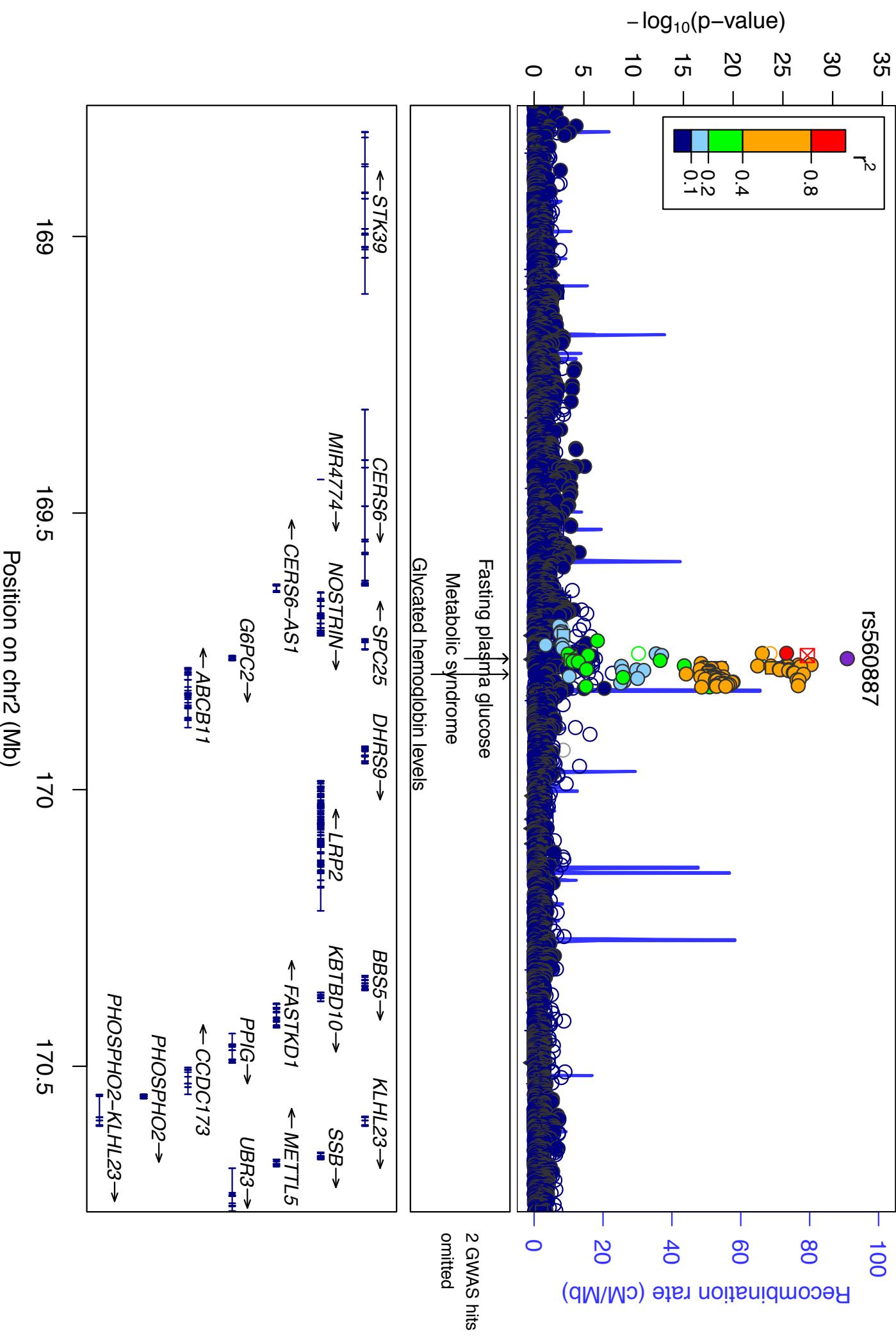
28

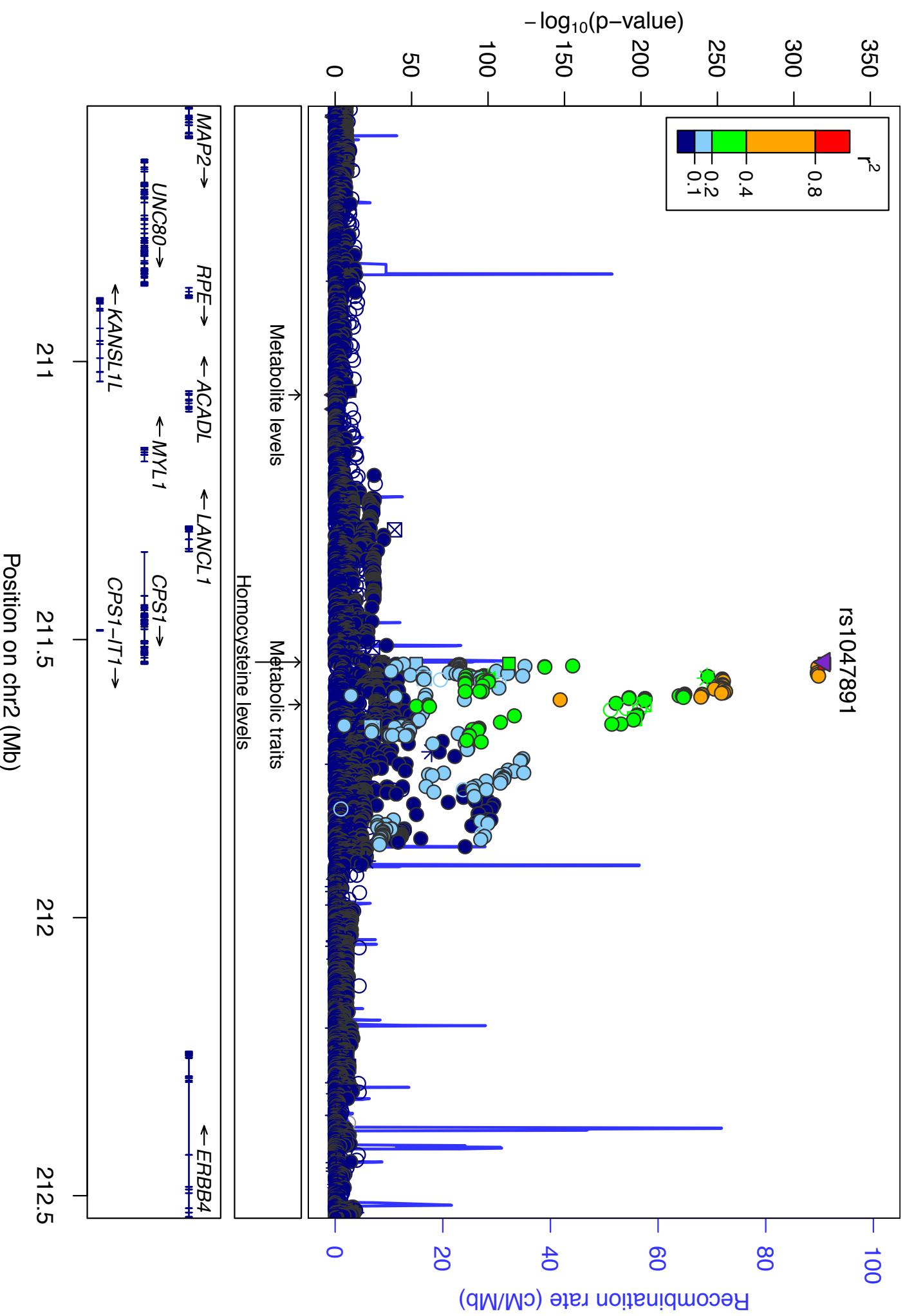
28.5

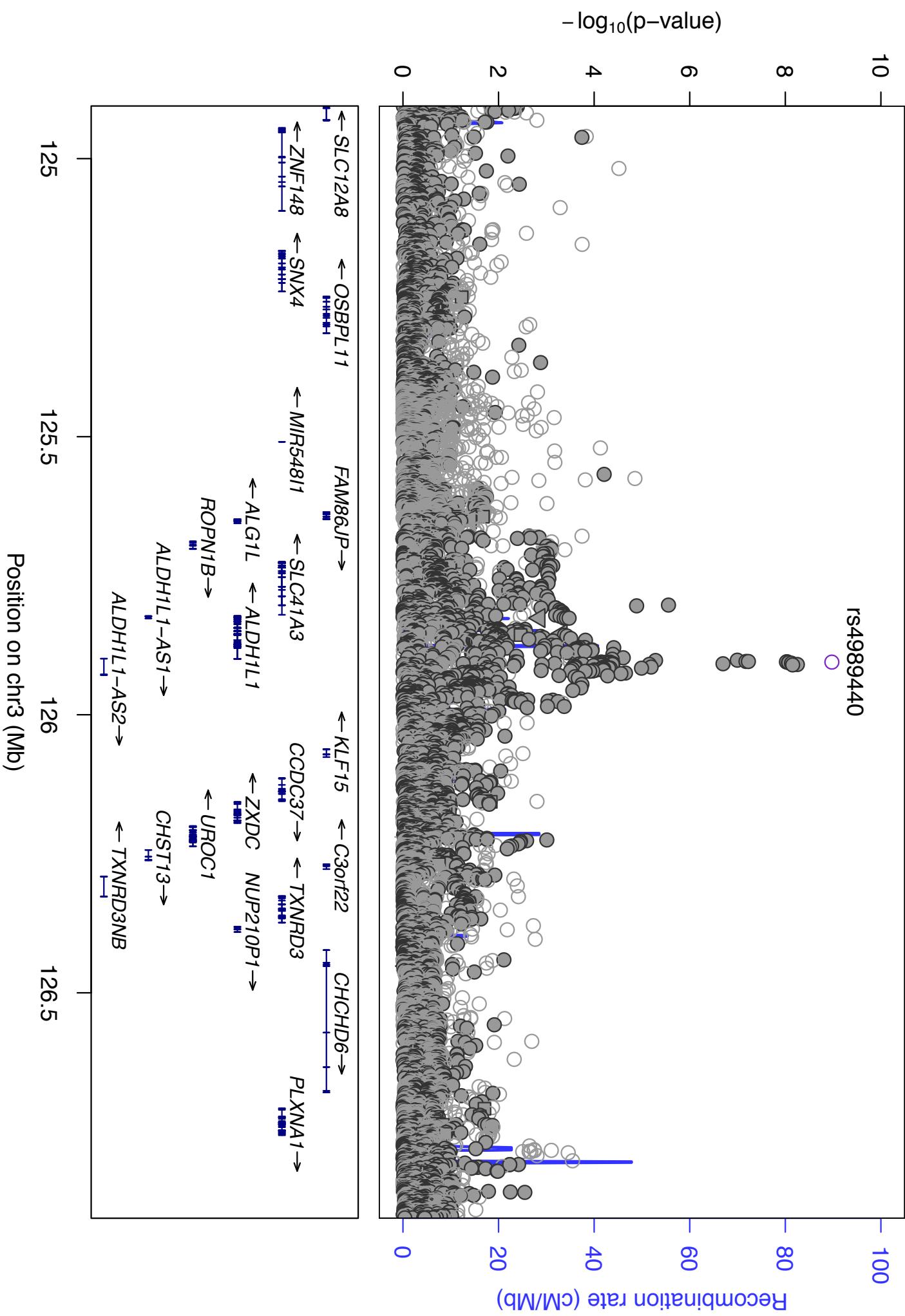
Position on chr2 (Mb)

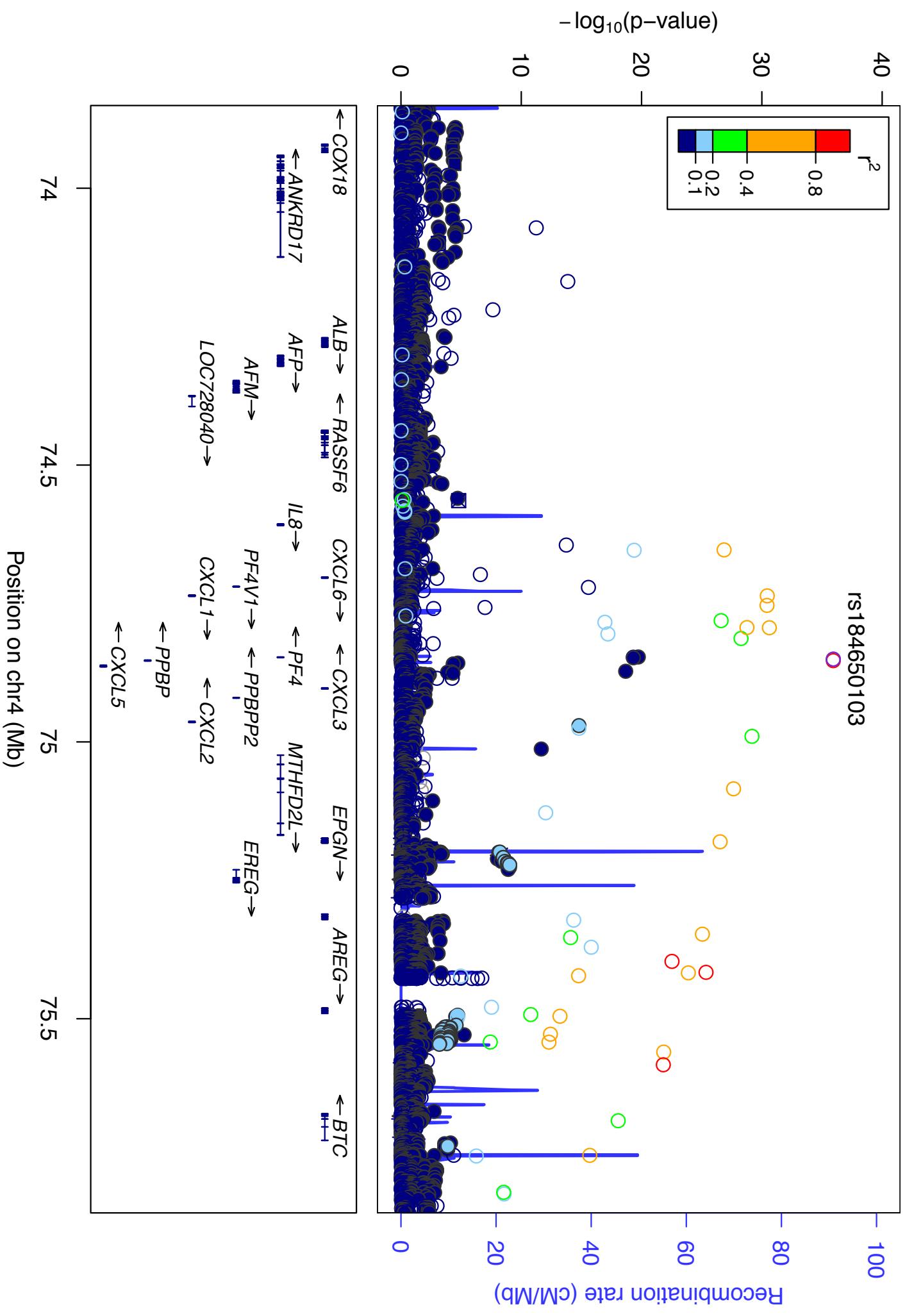


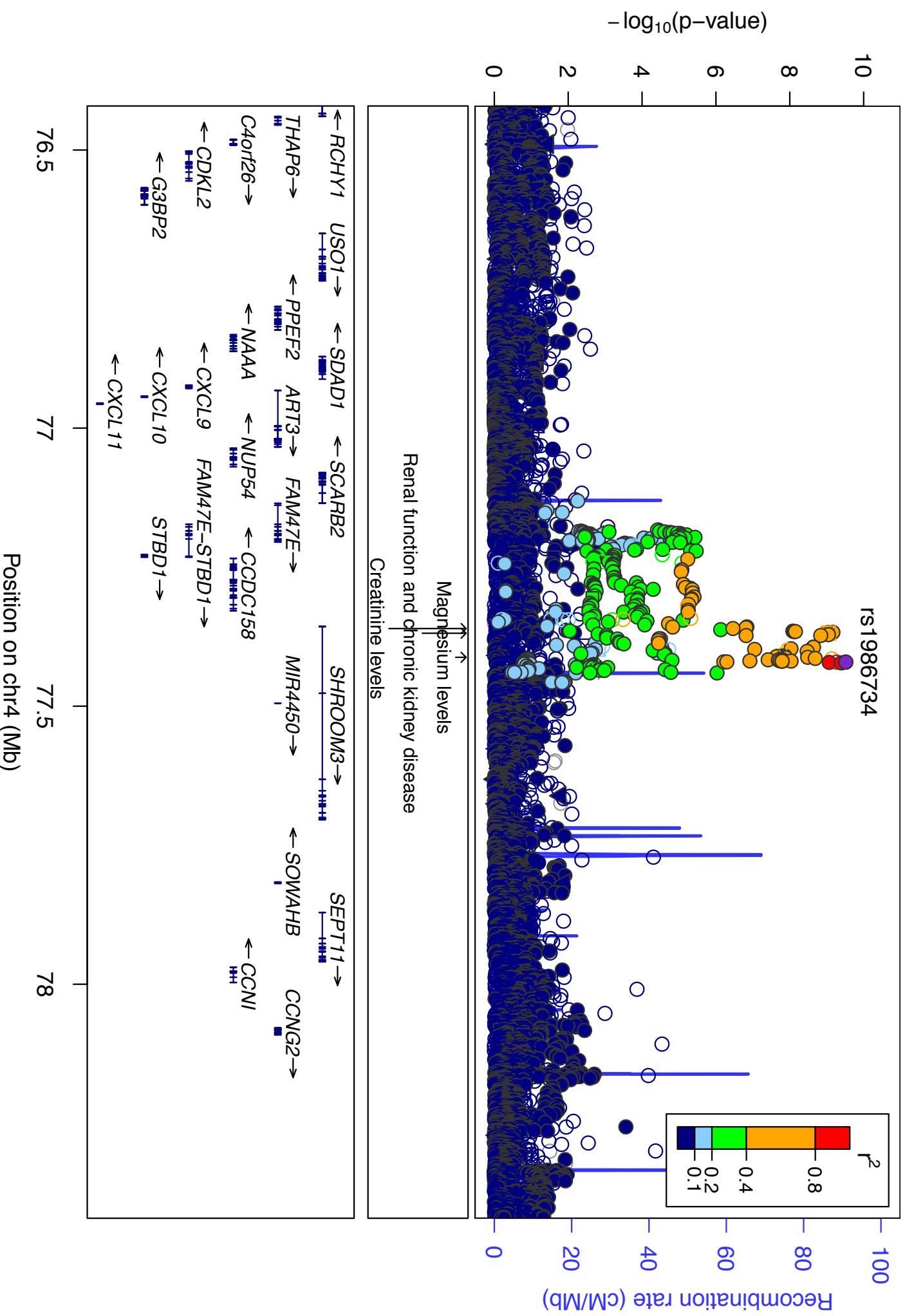


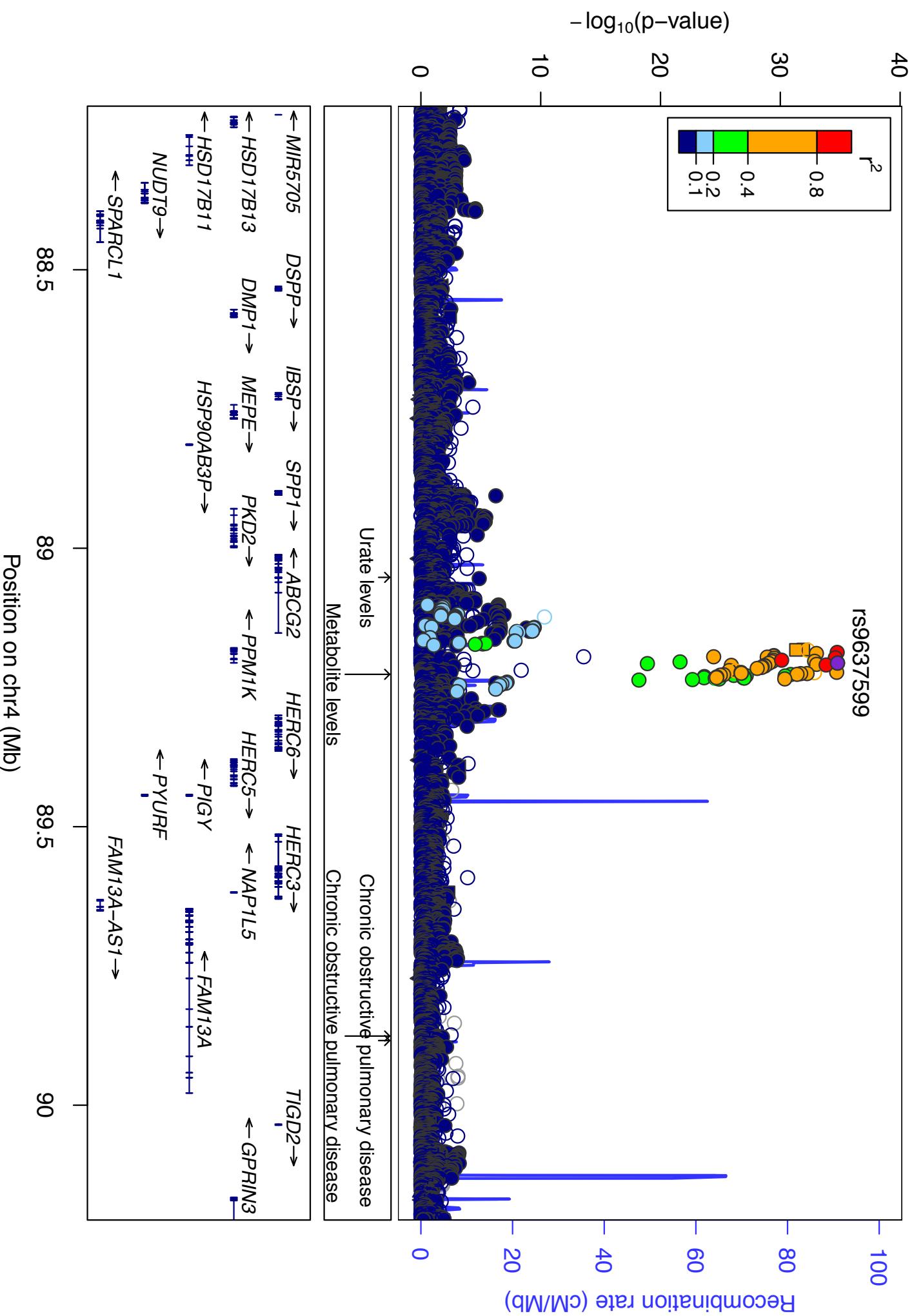


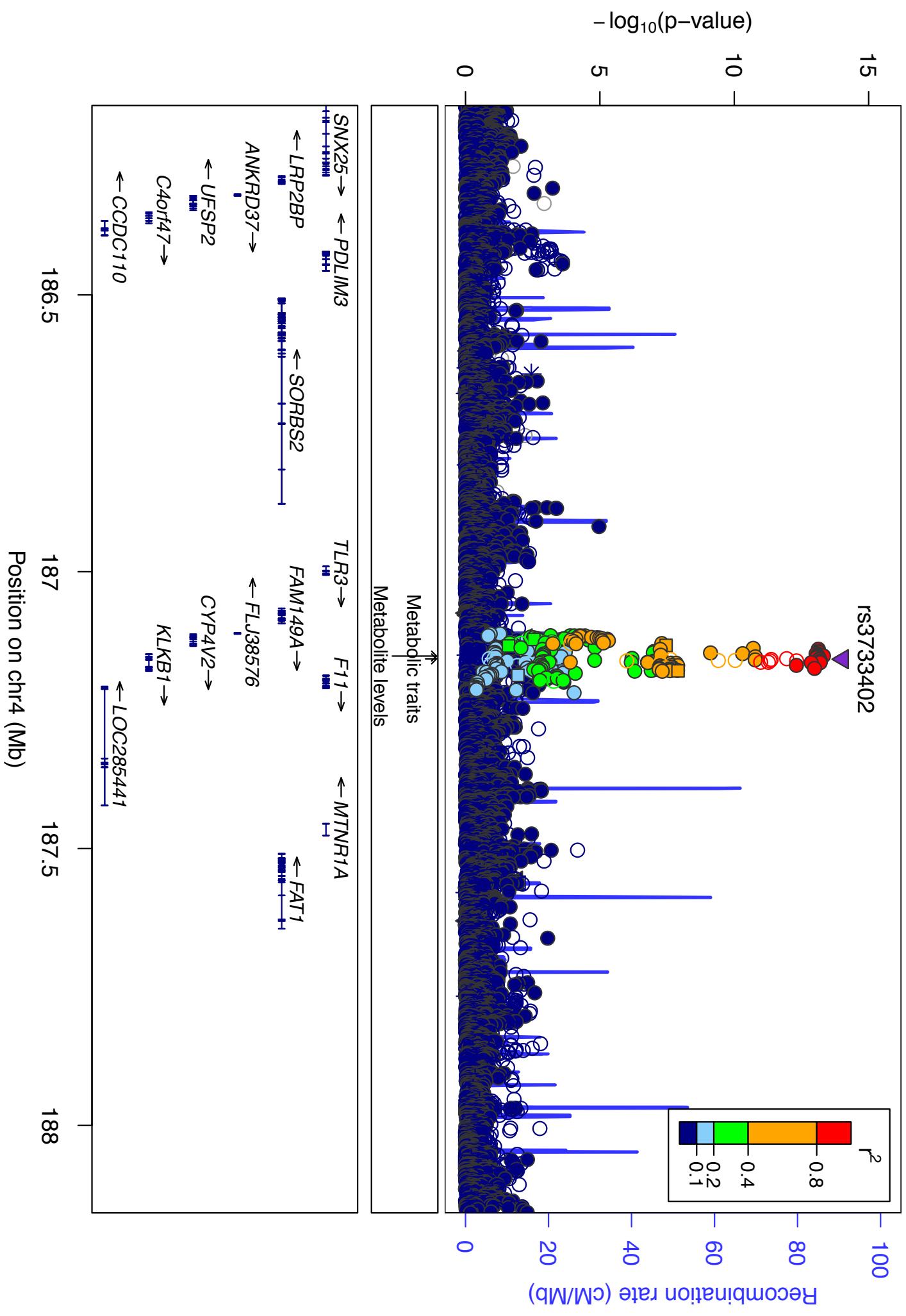


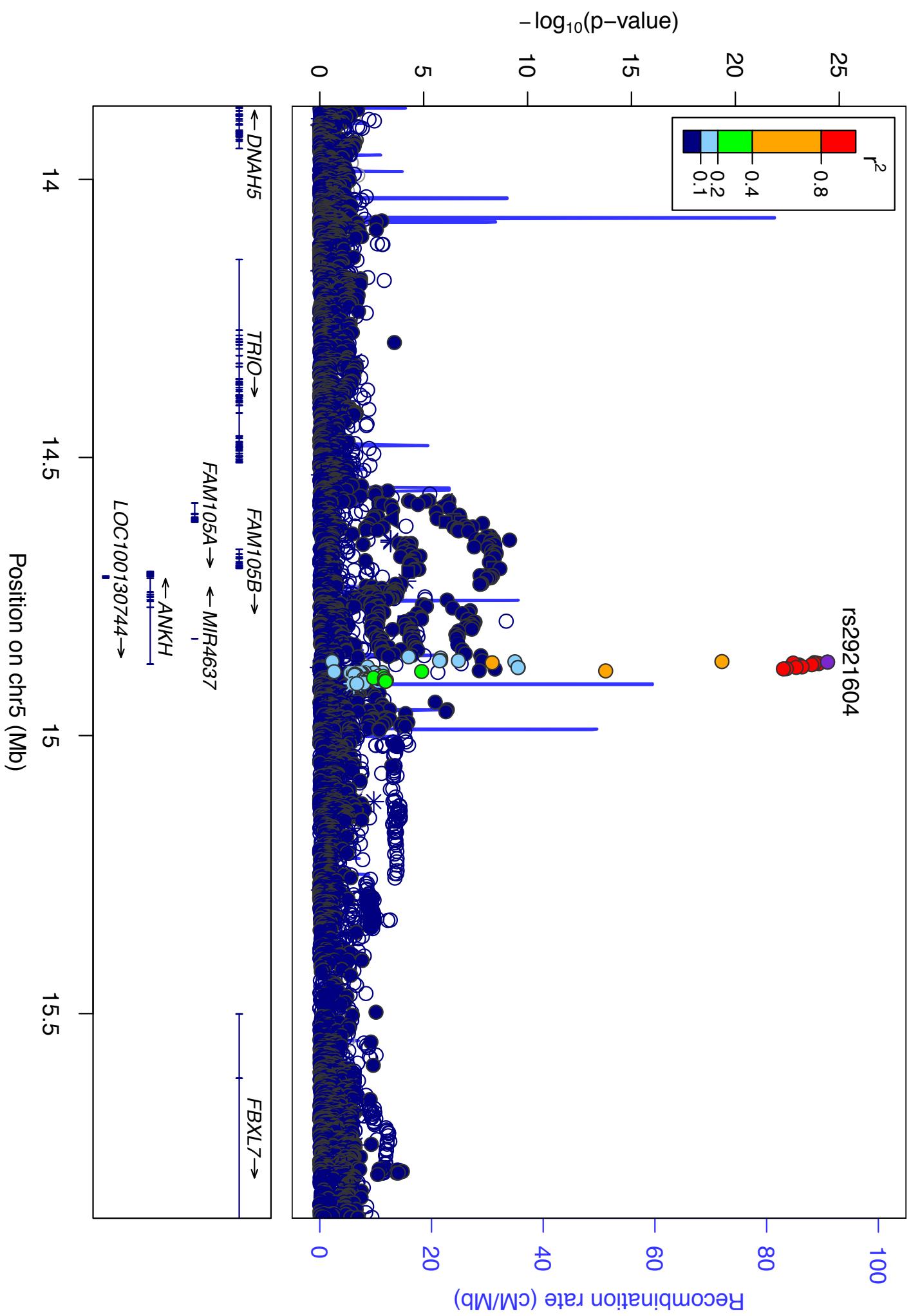


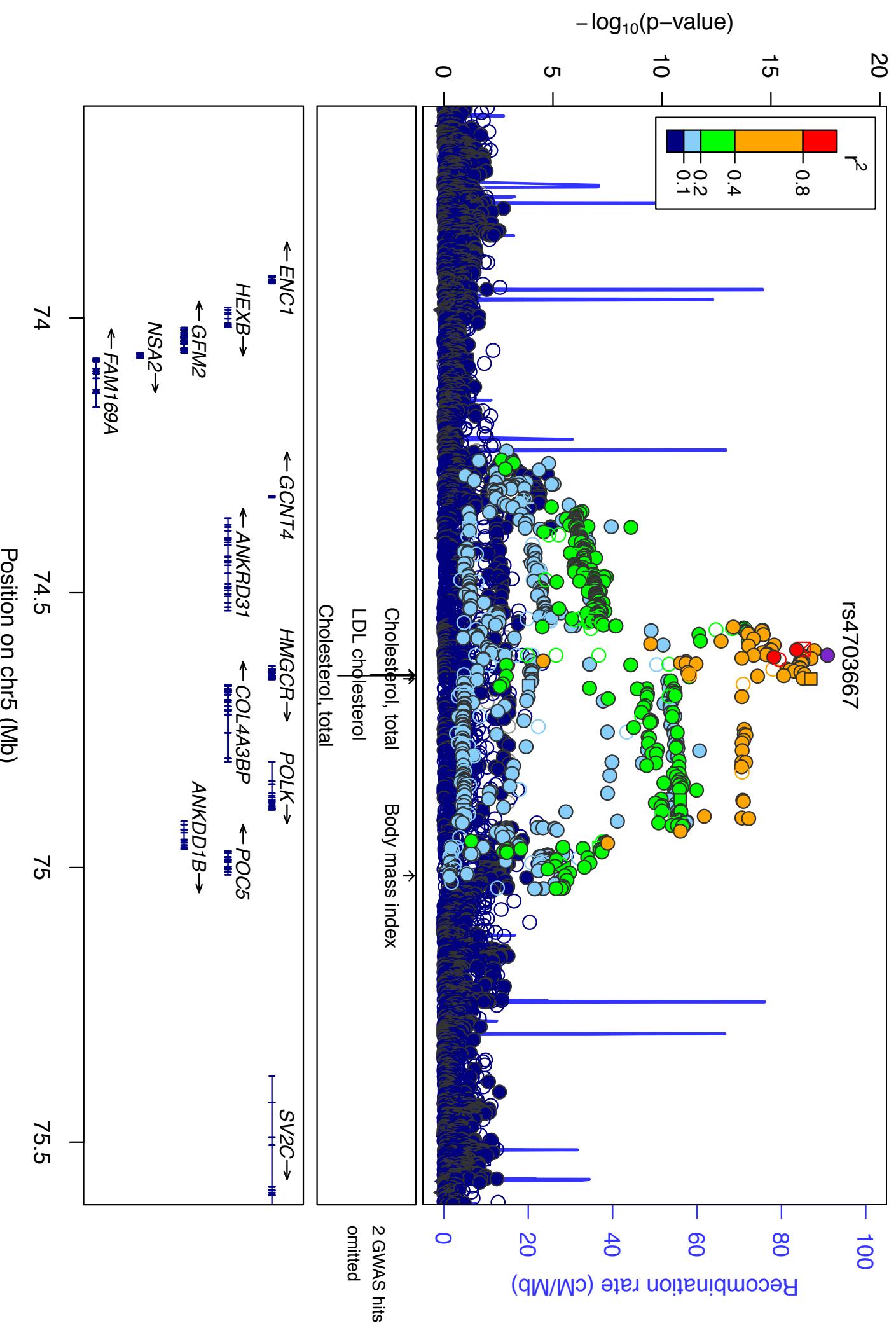


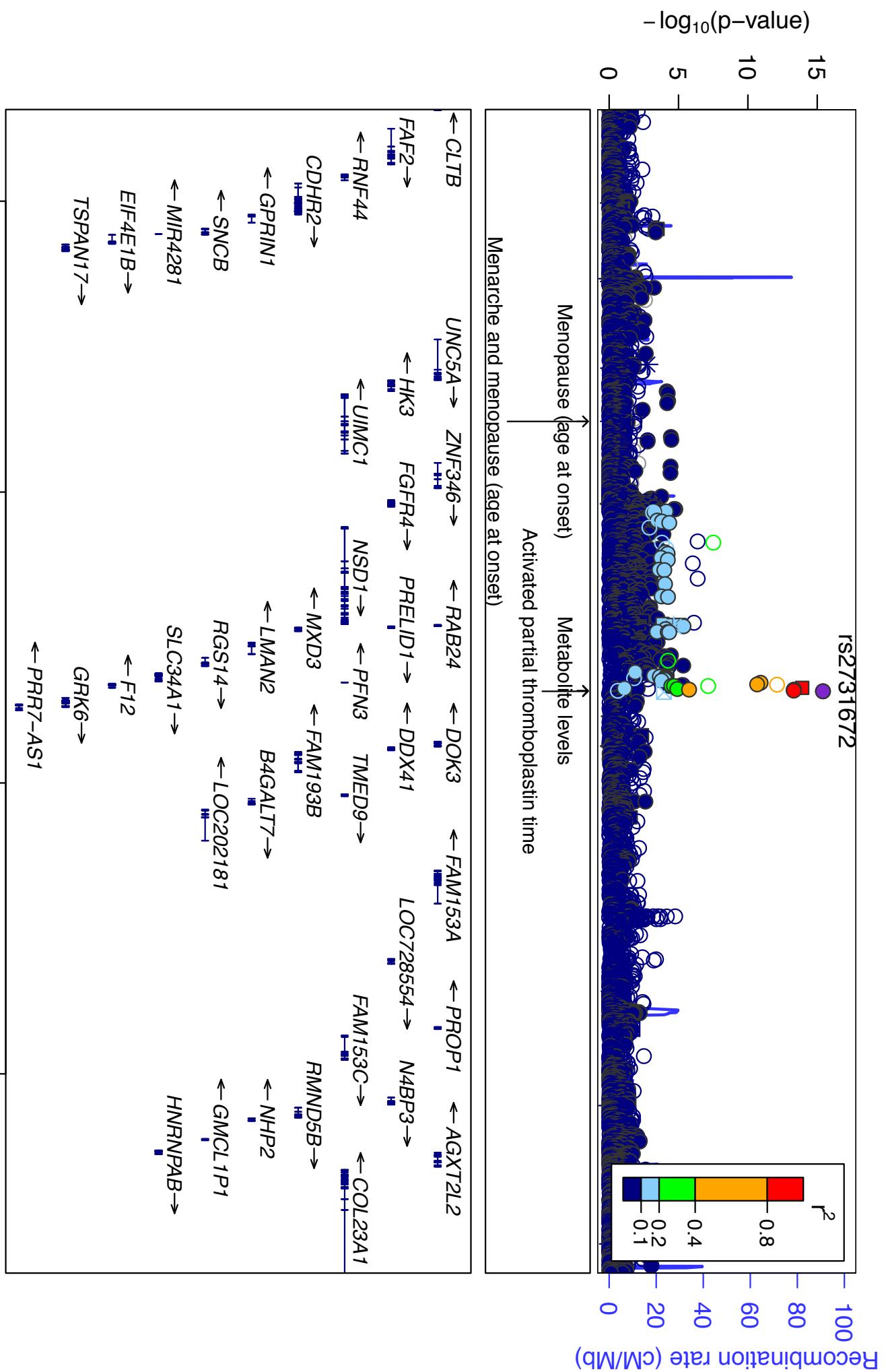


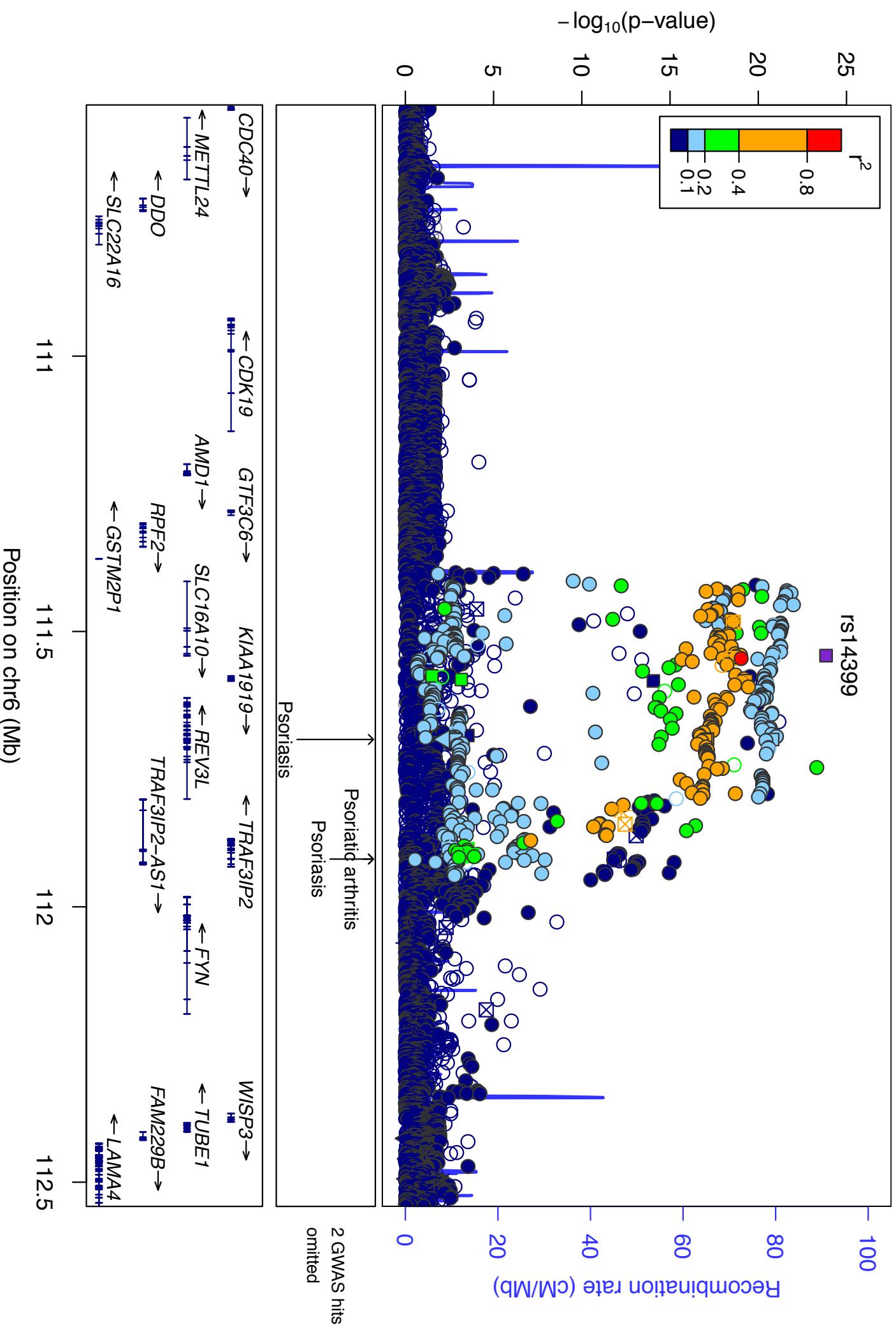


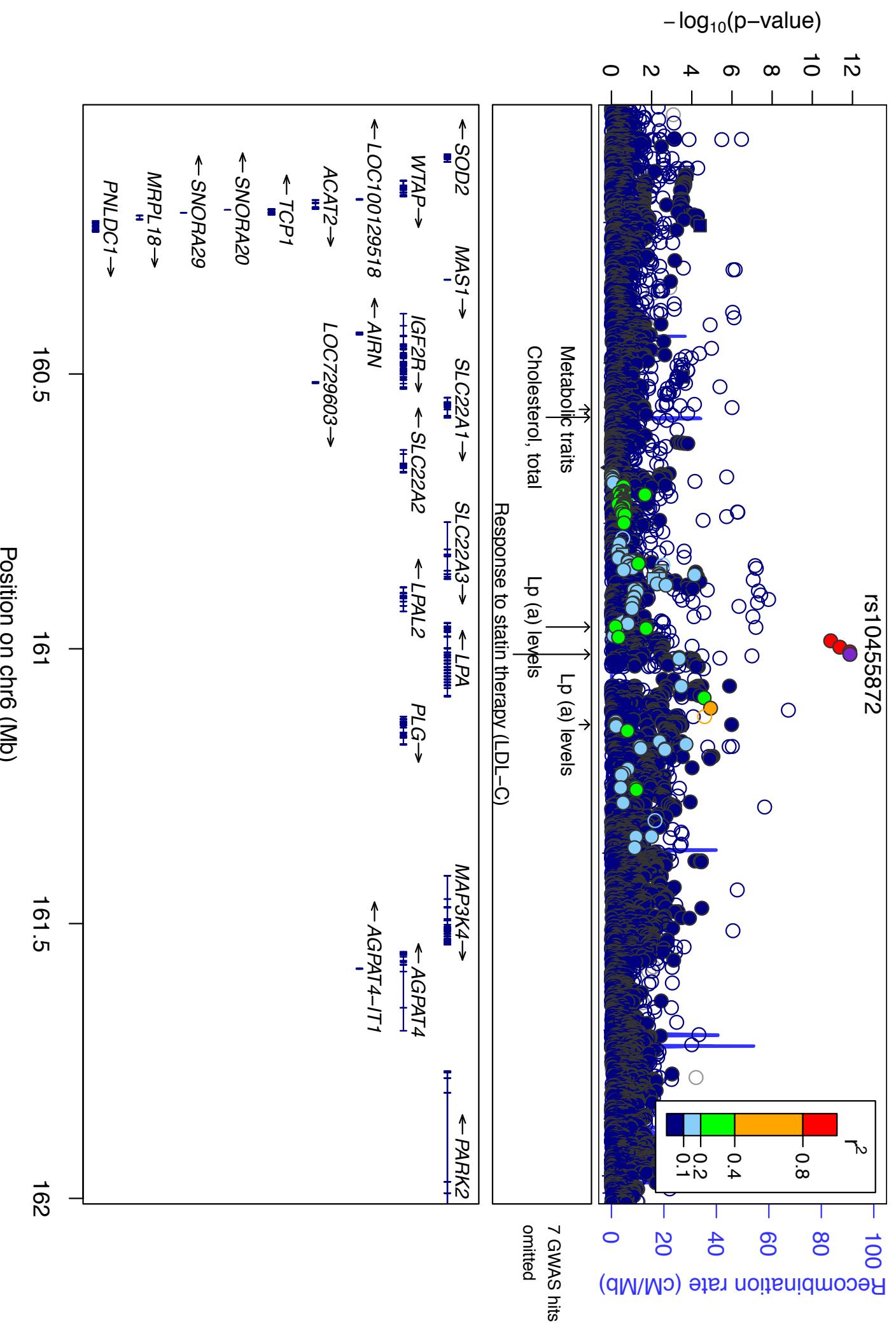


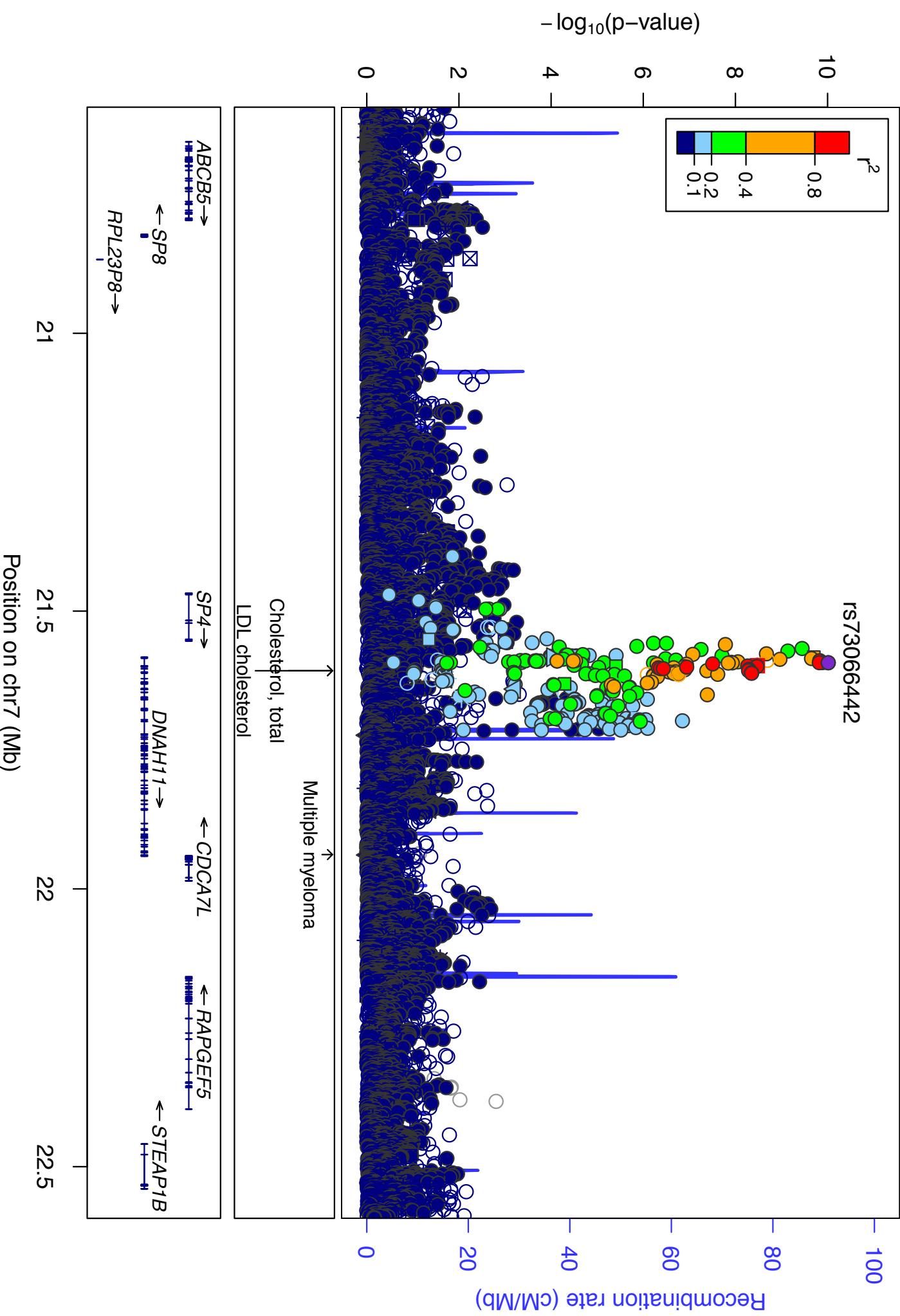


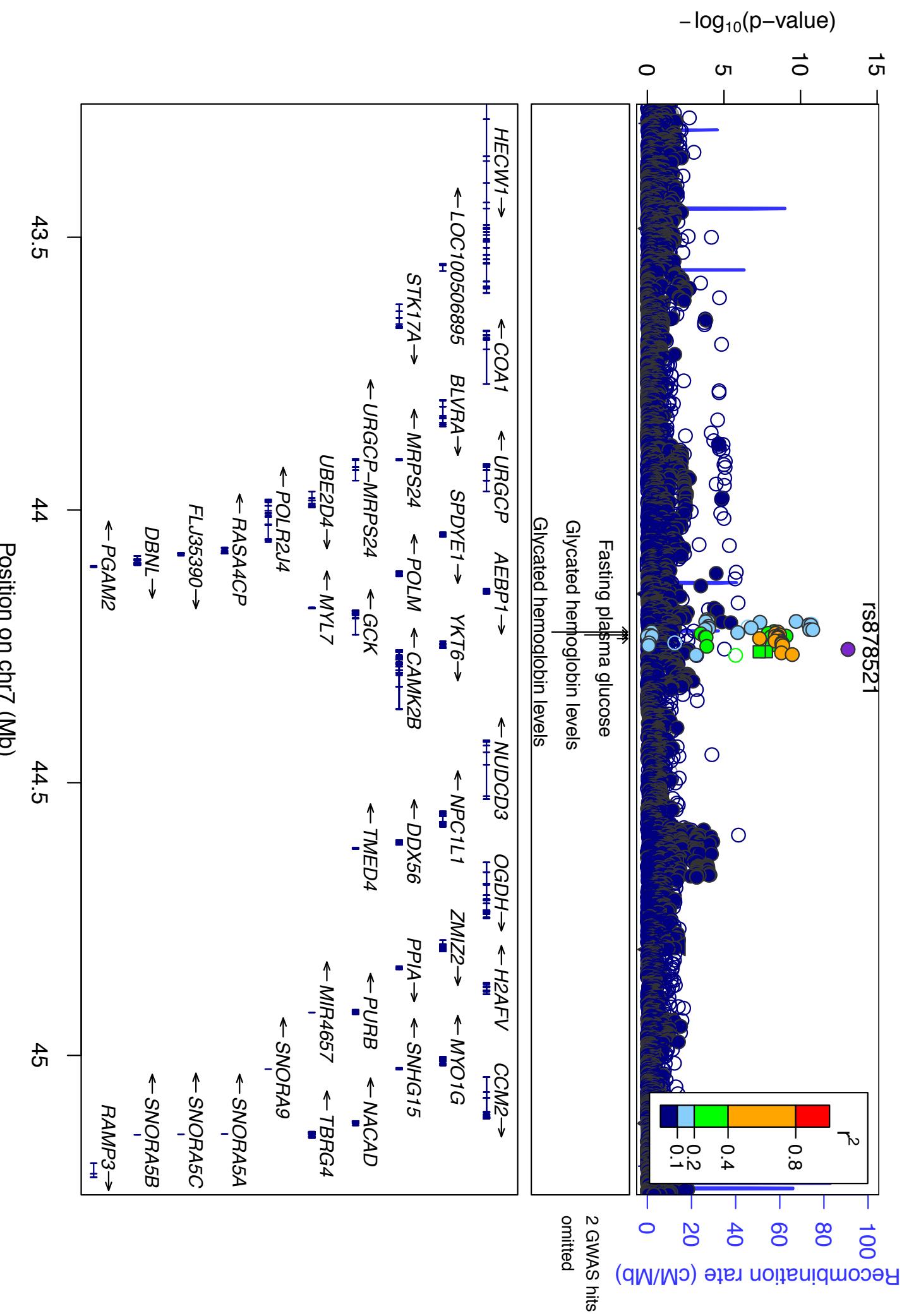


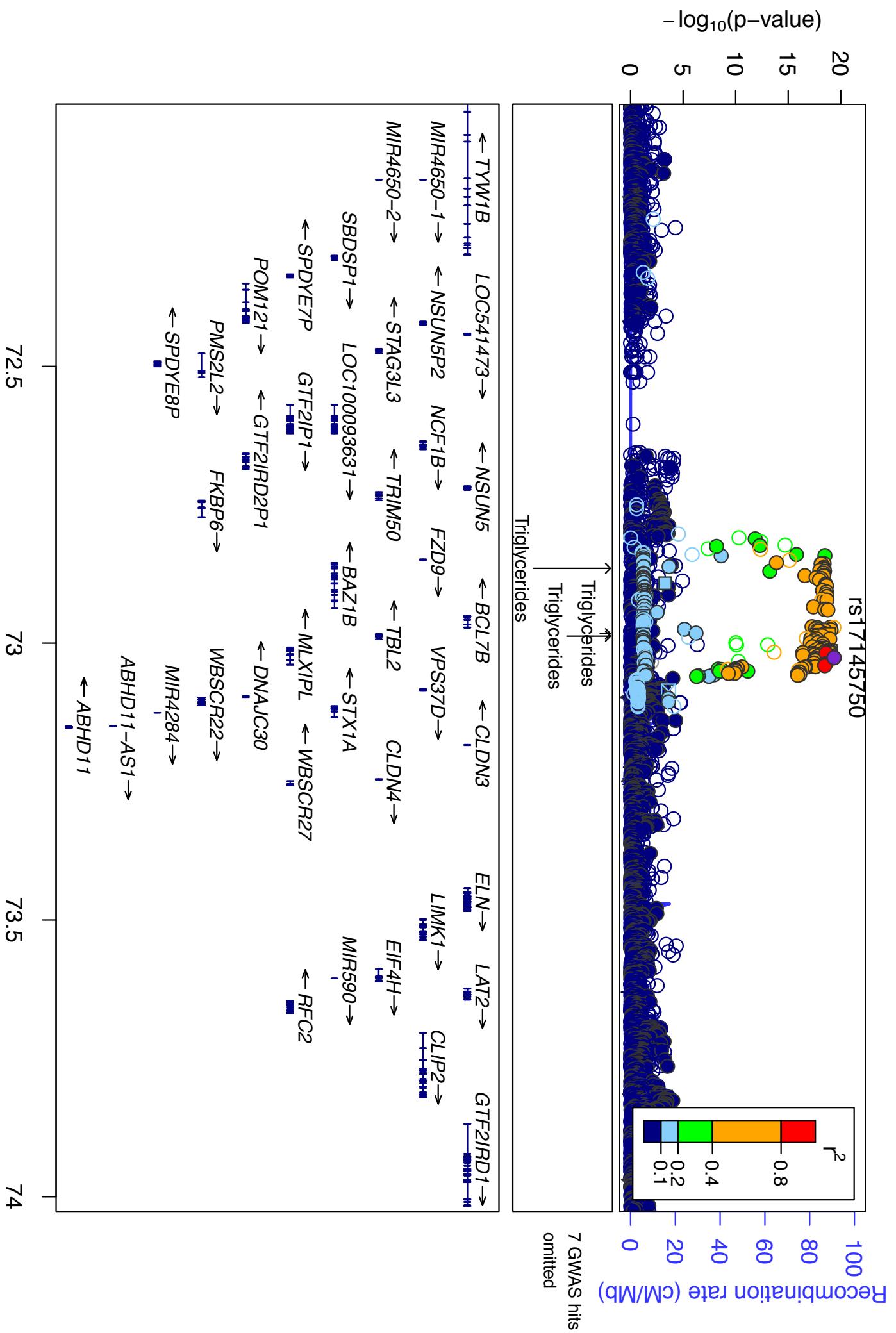


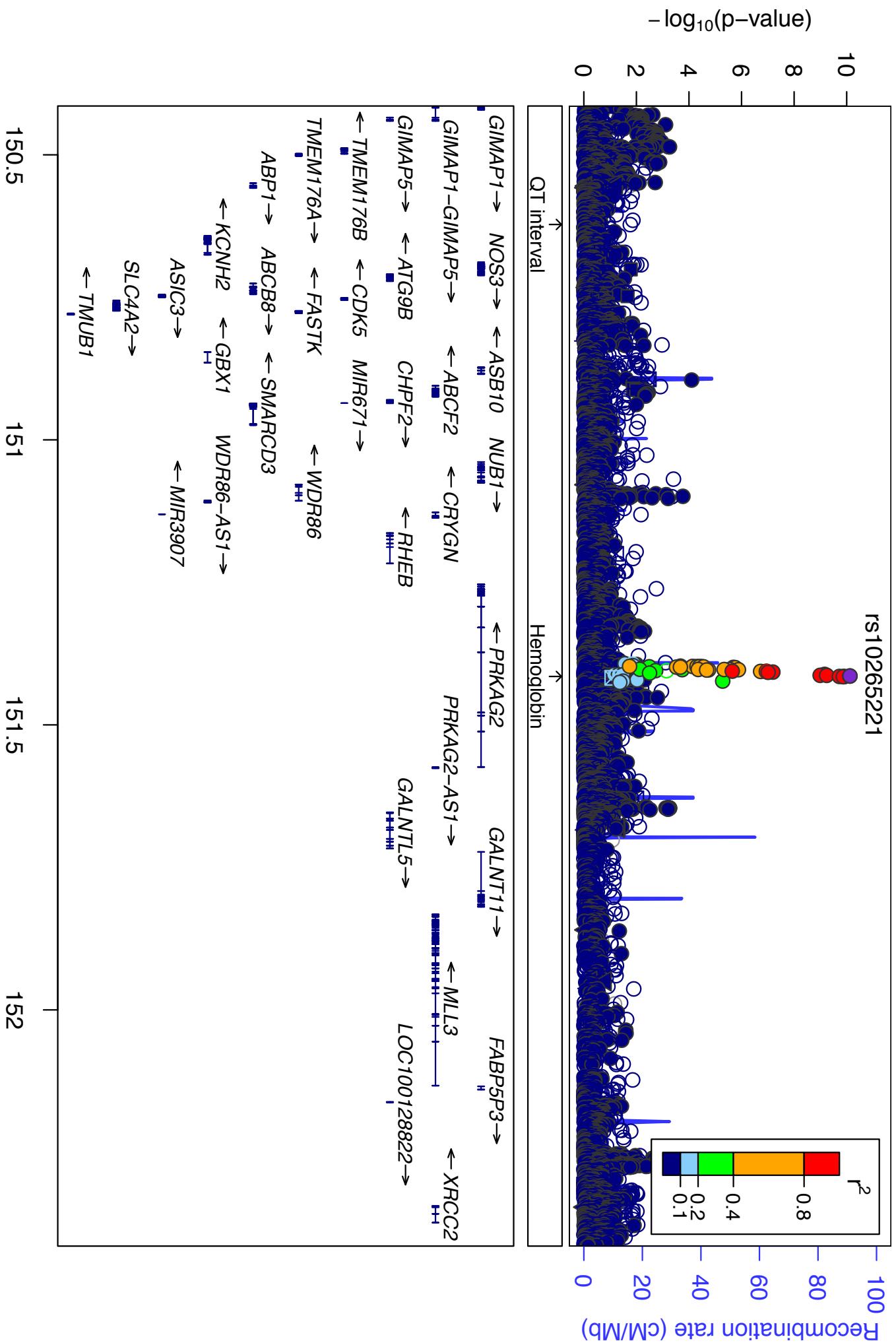


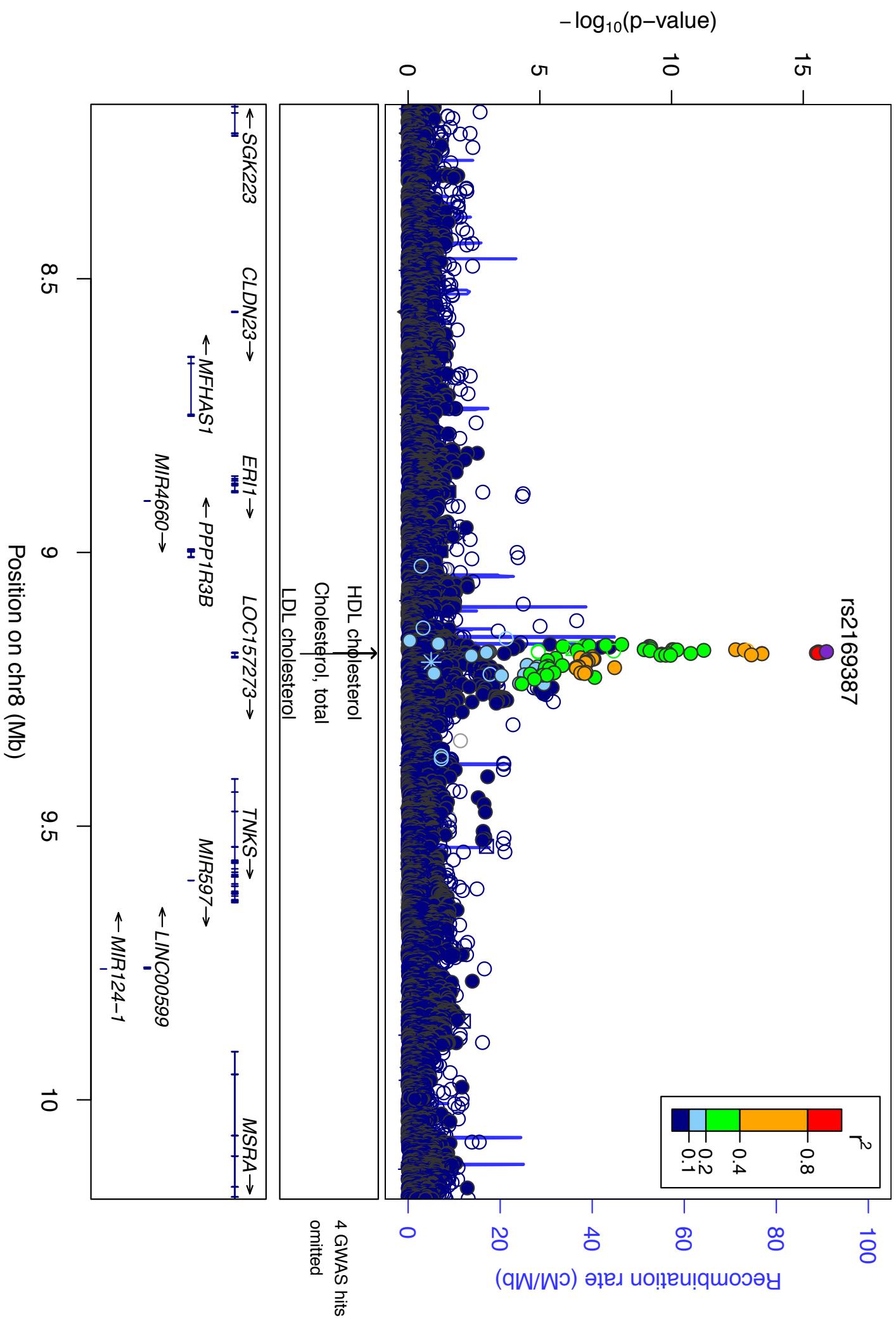


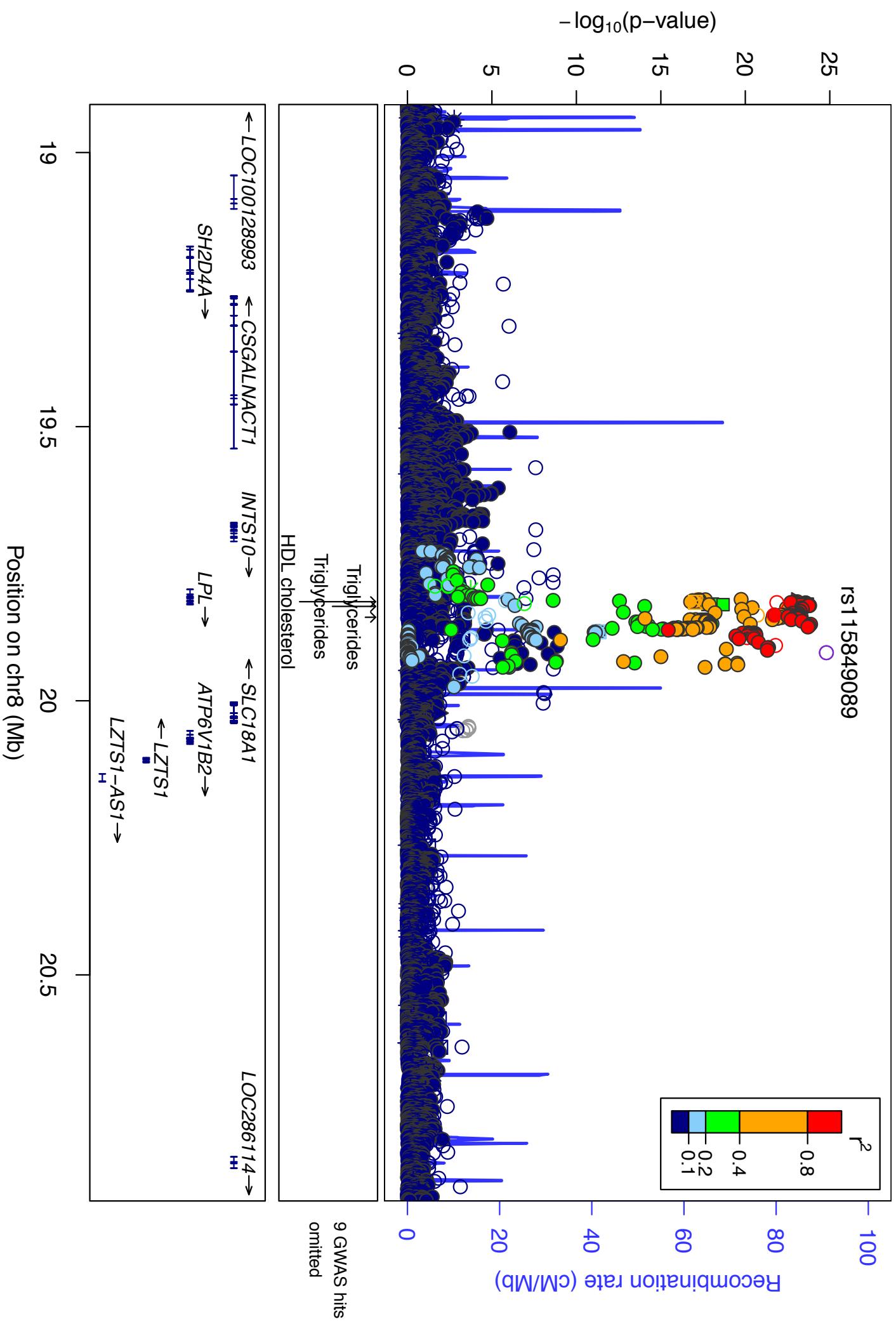


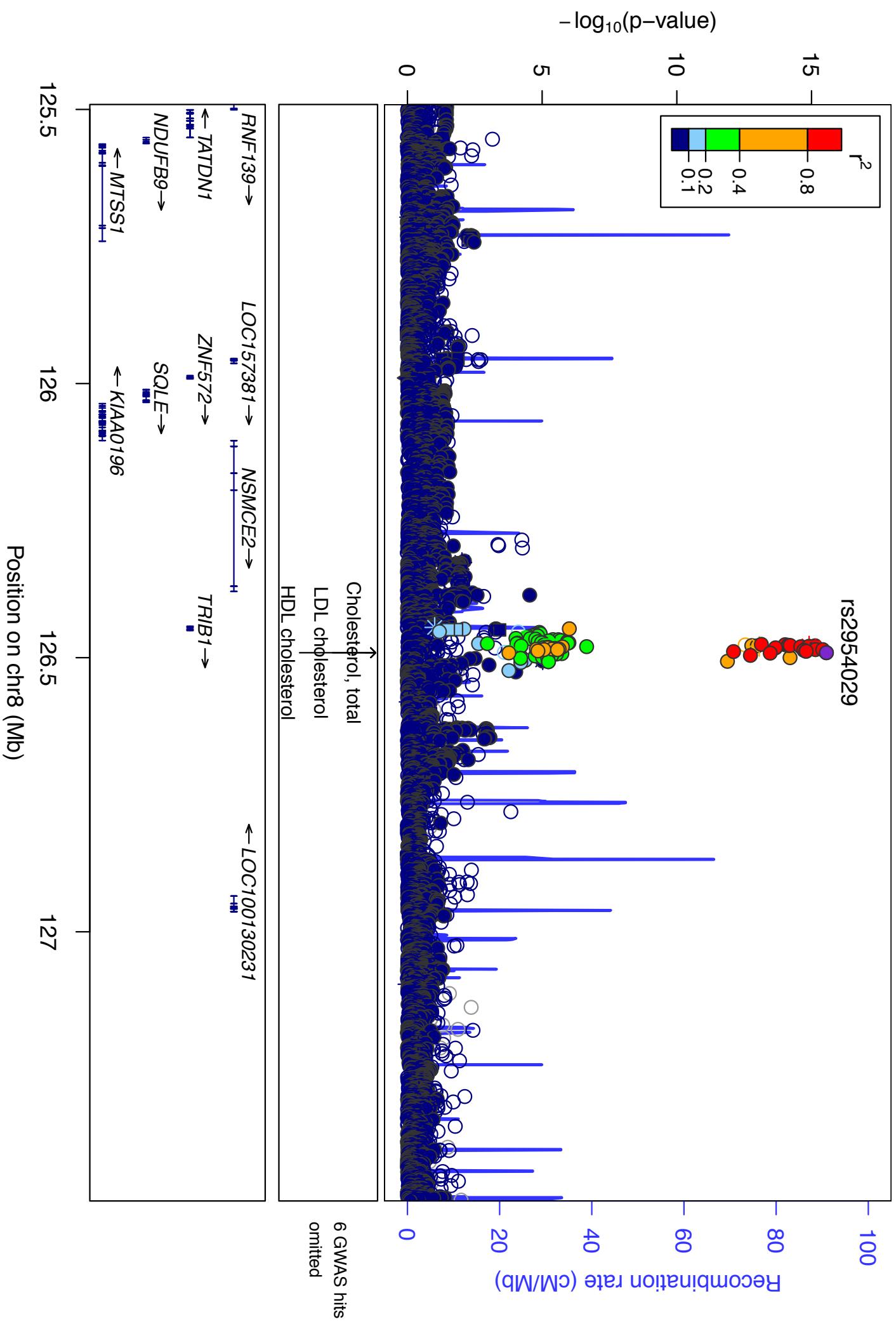


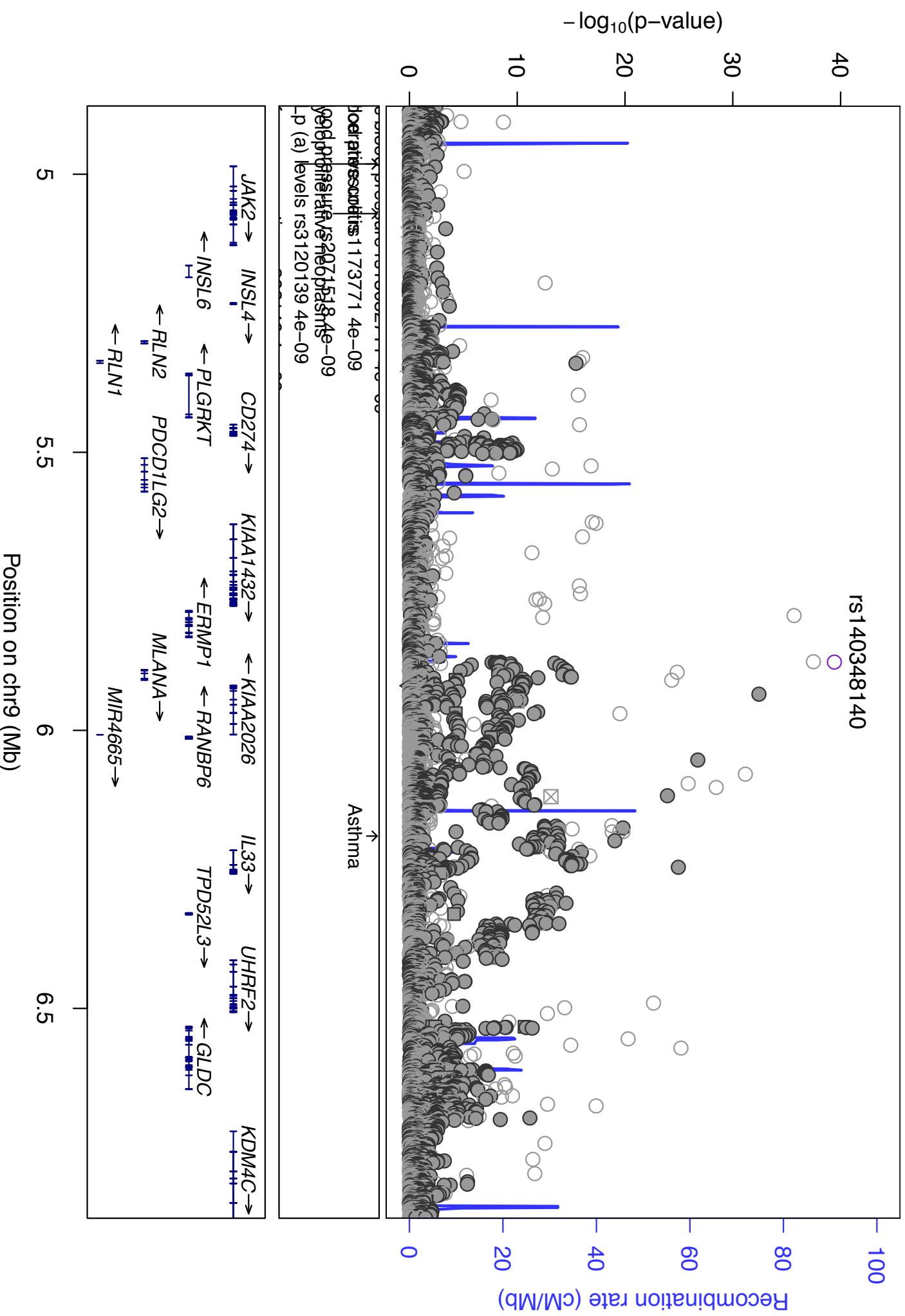


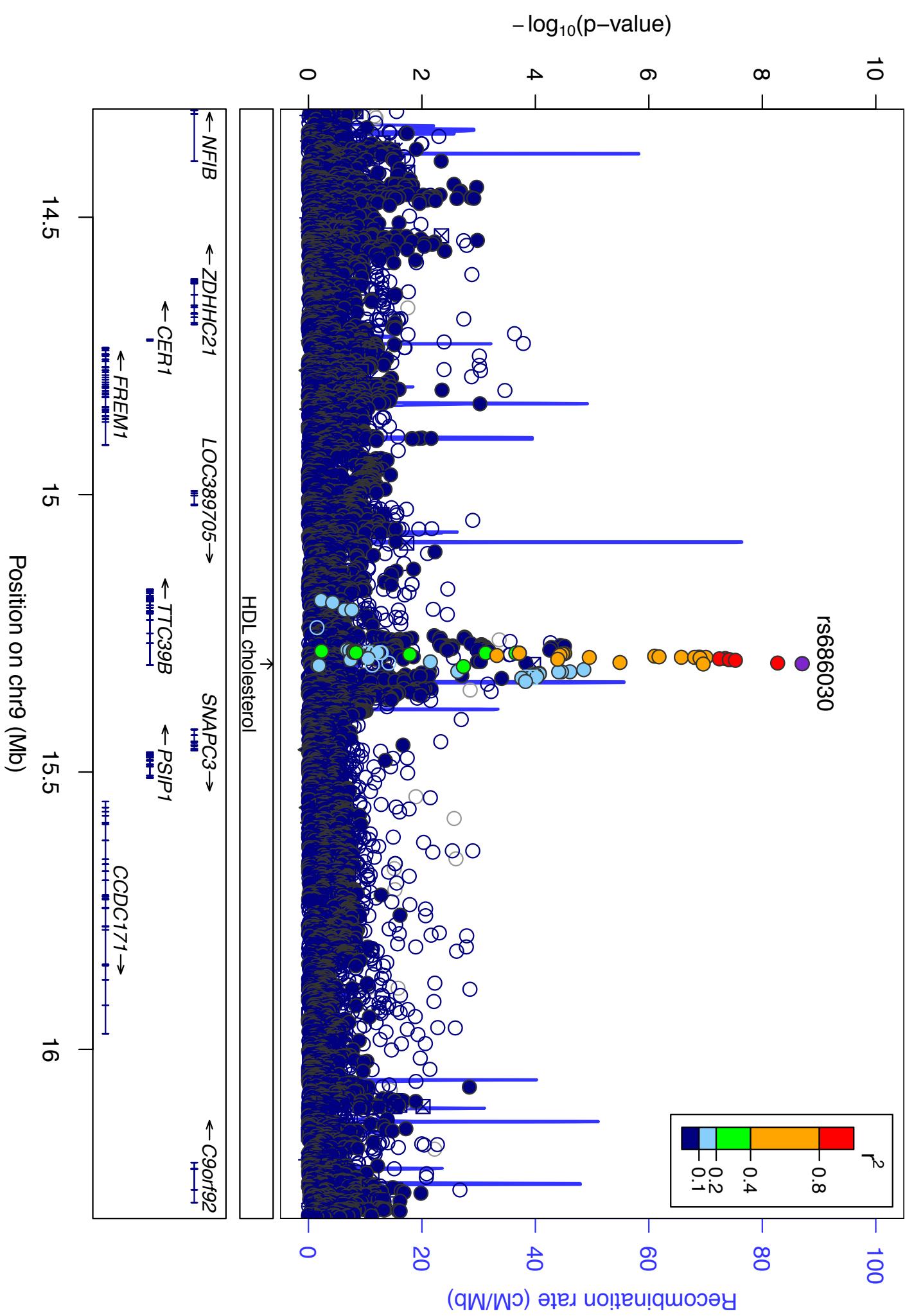


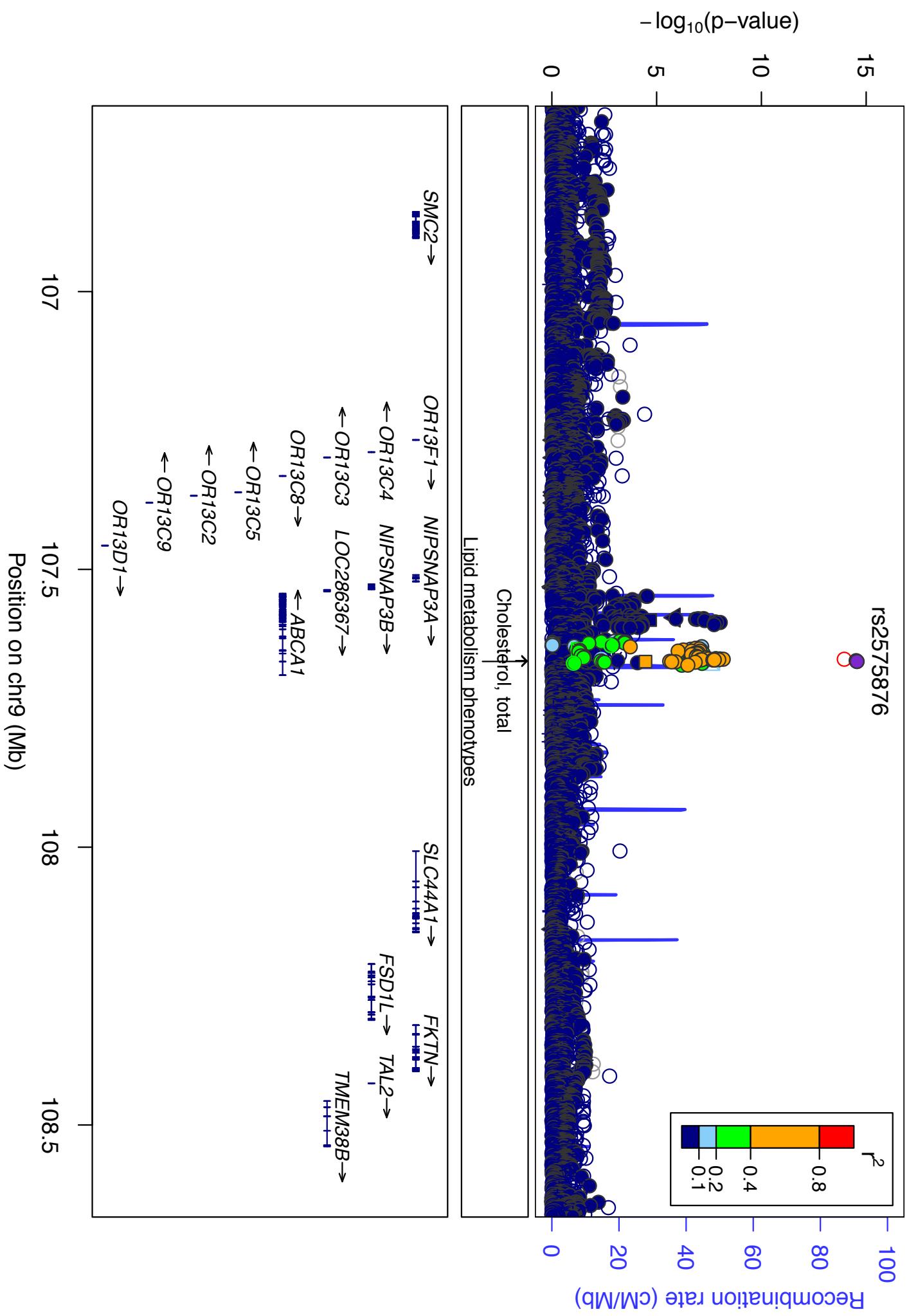


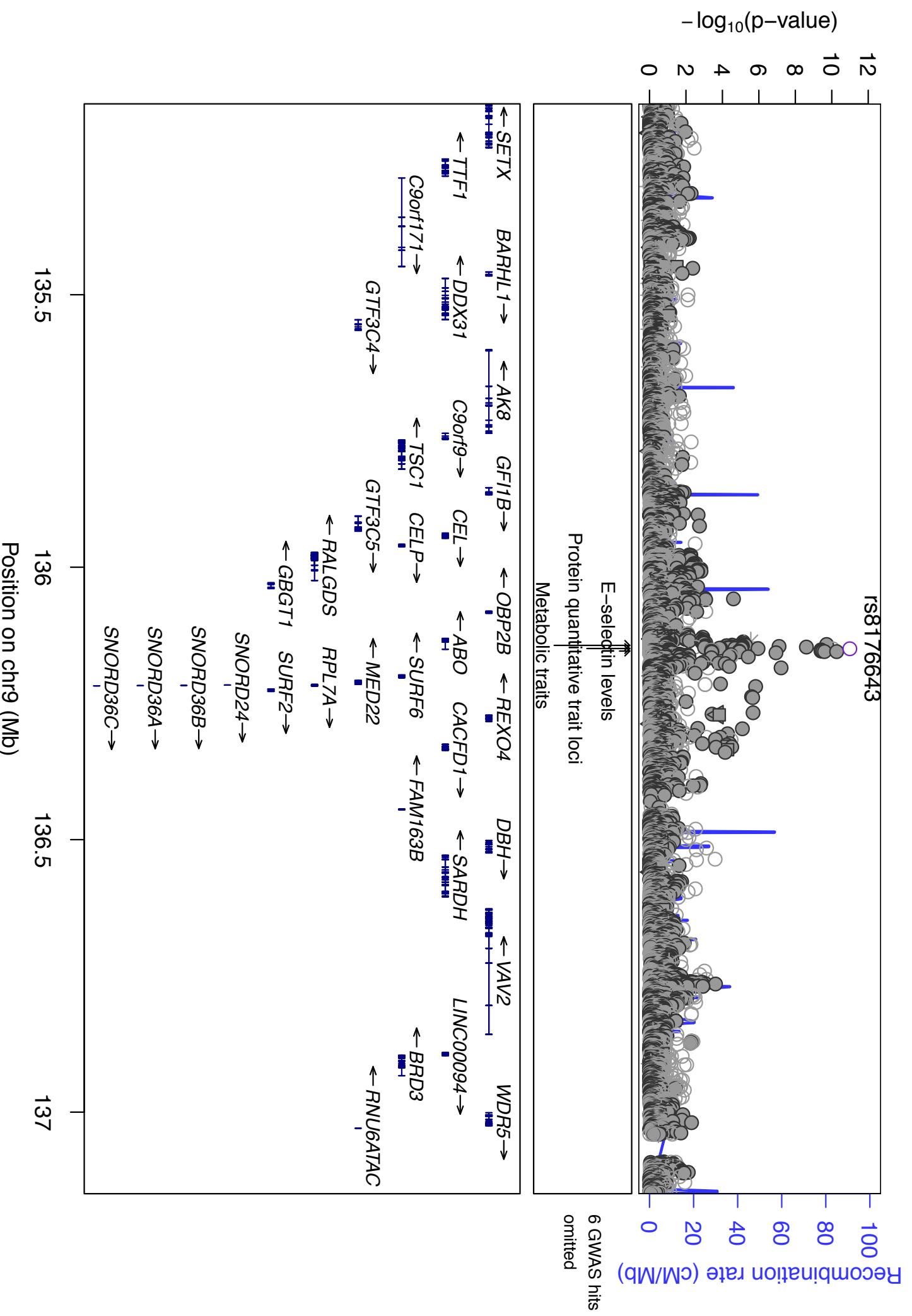


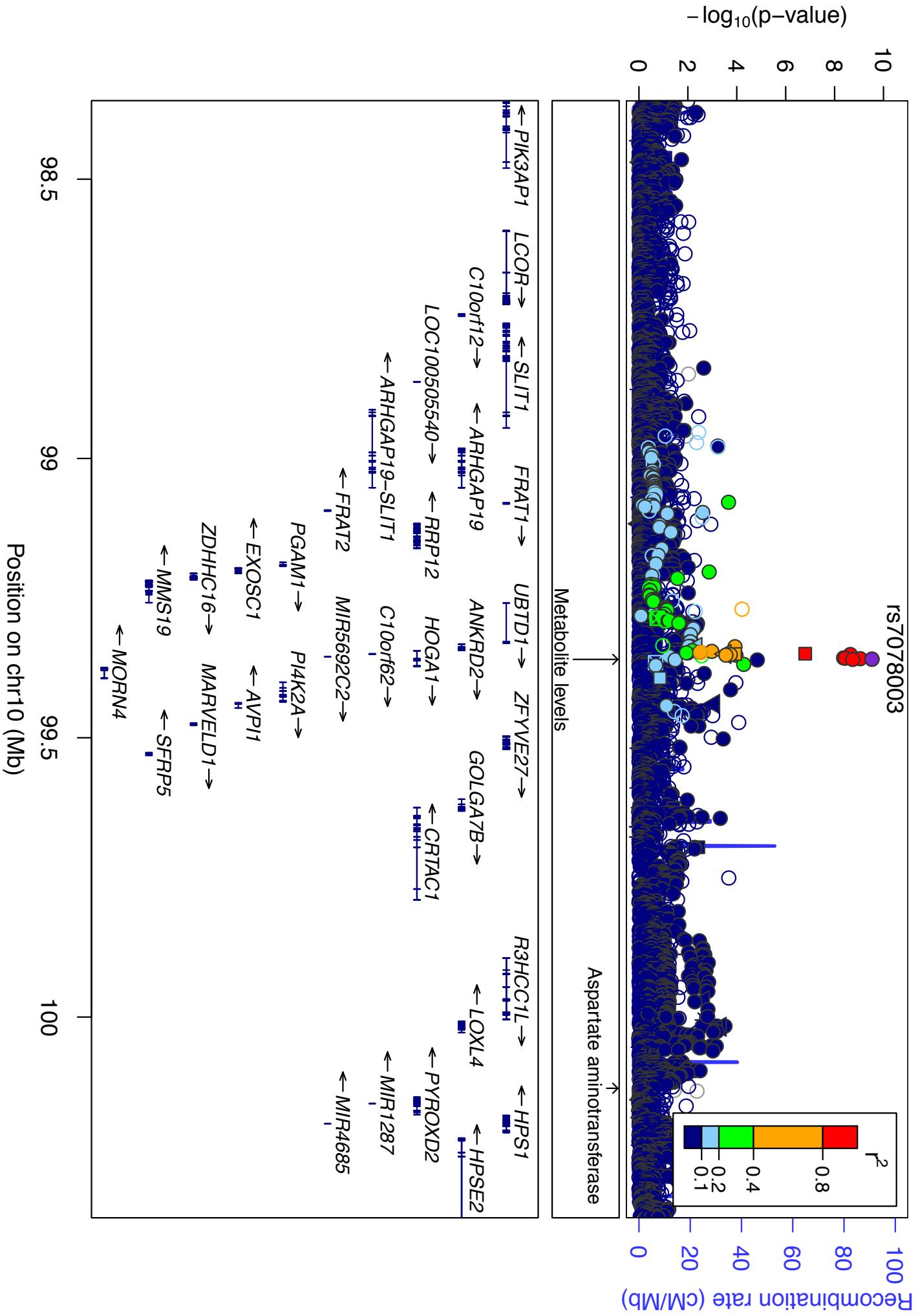


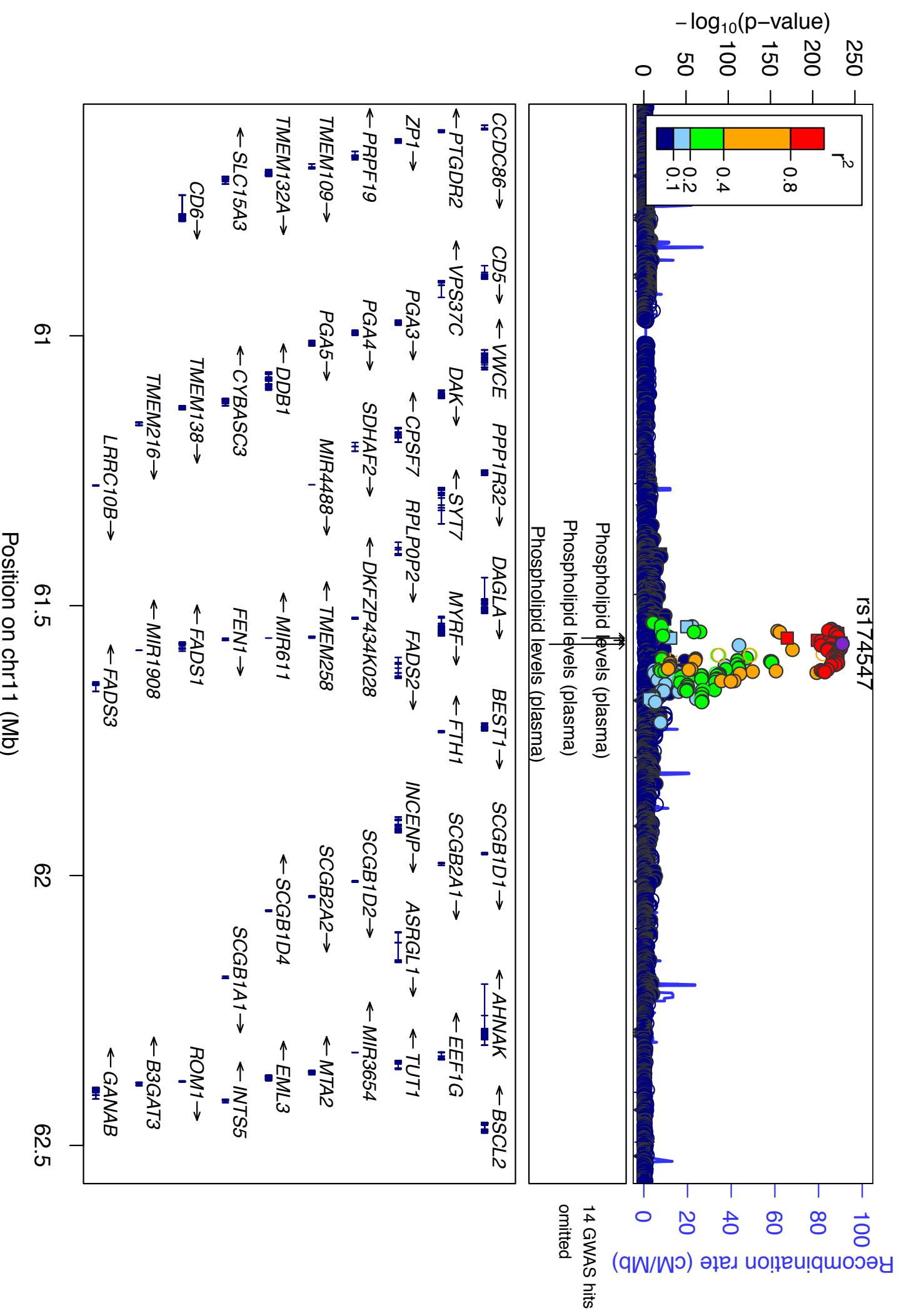


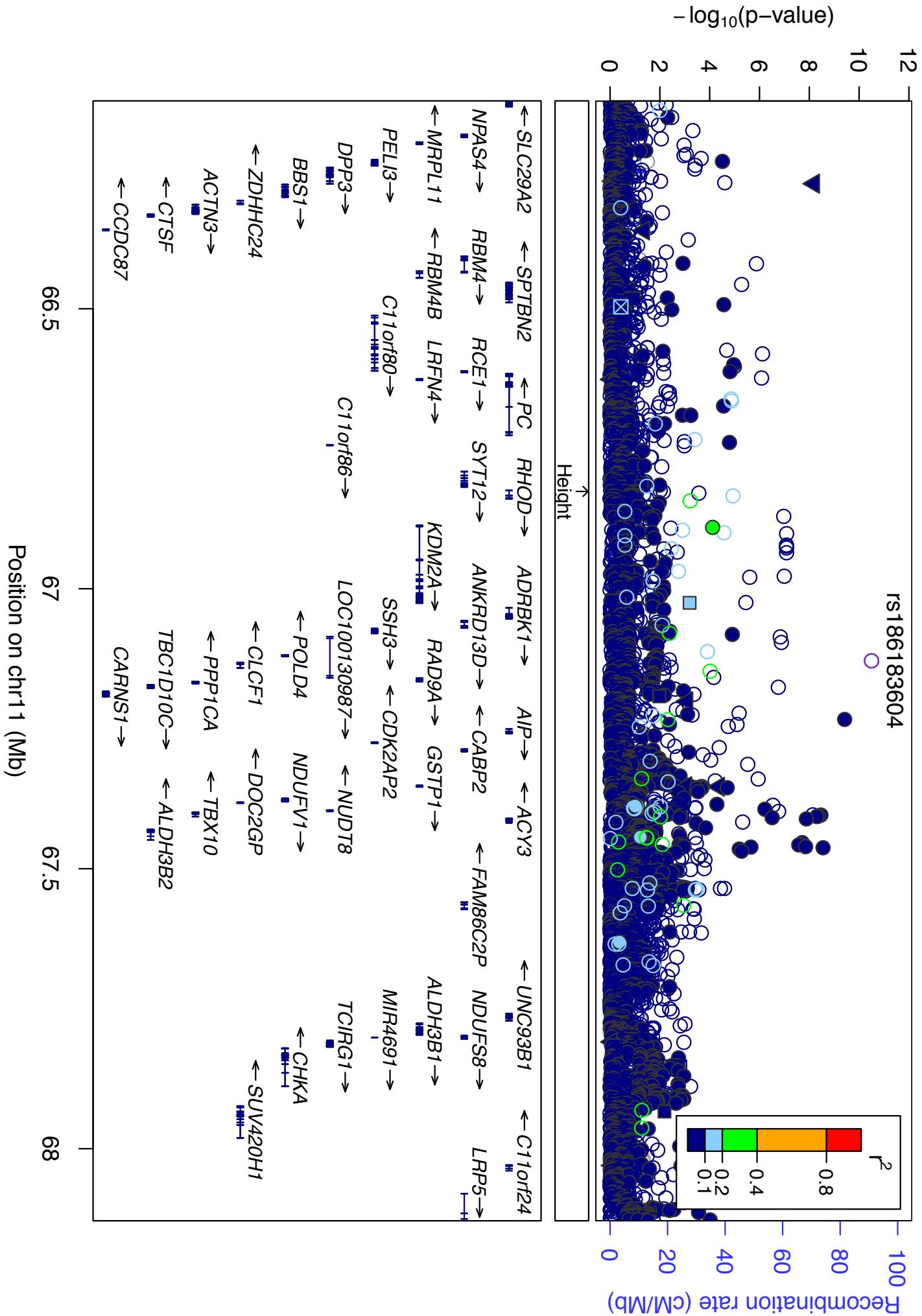


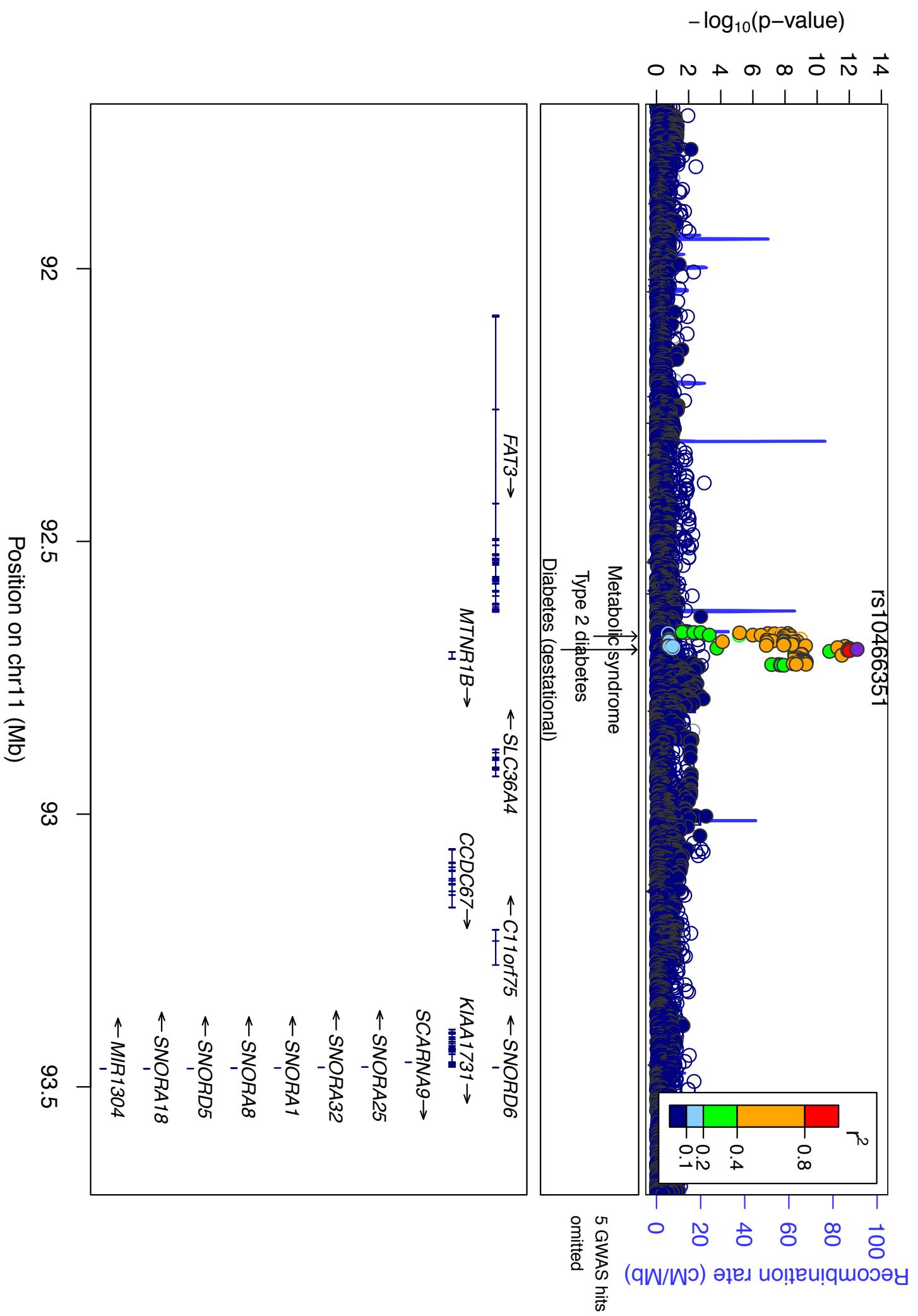


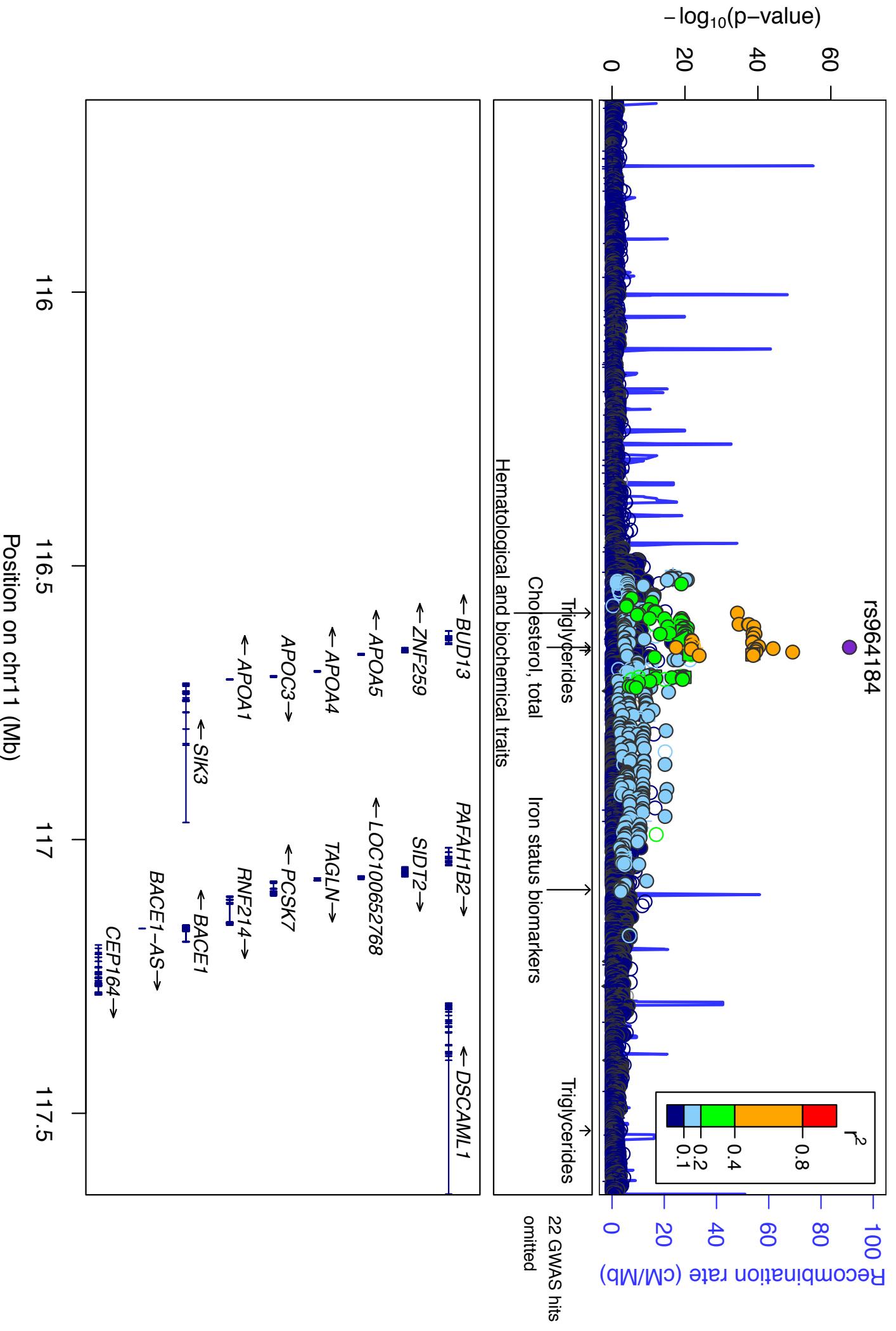


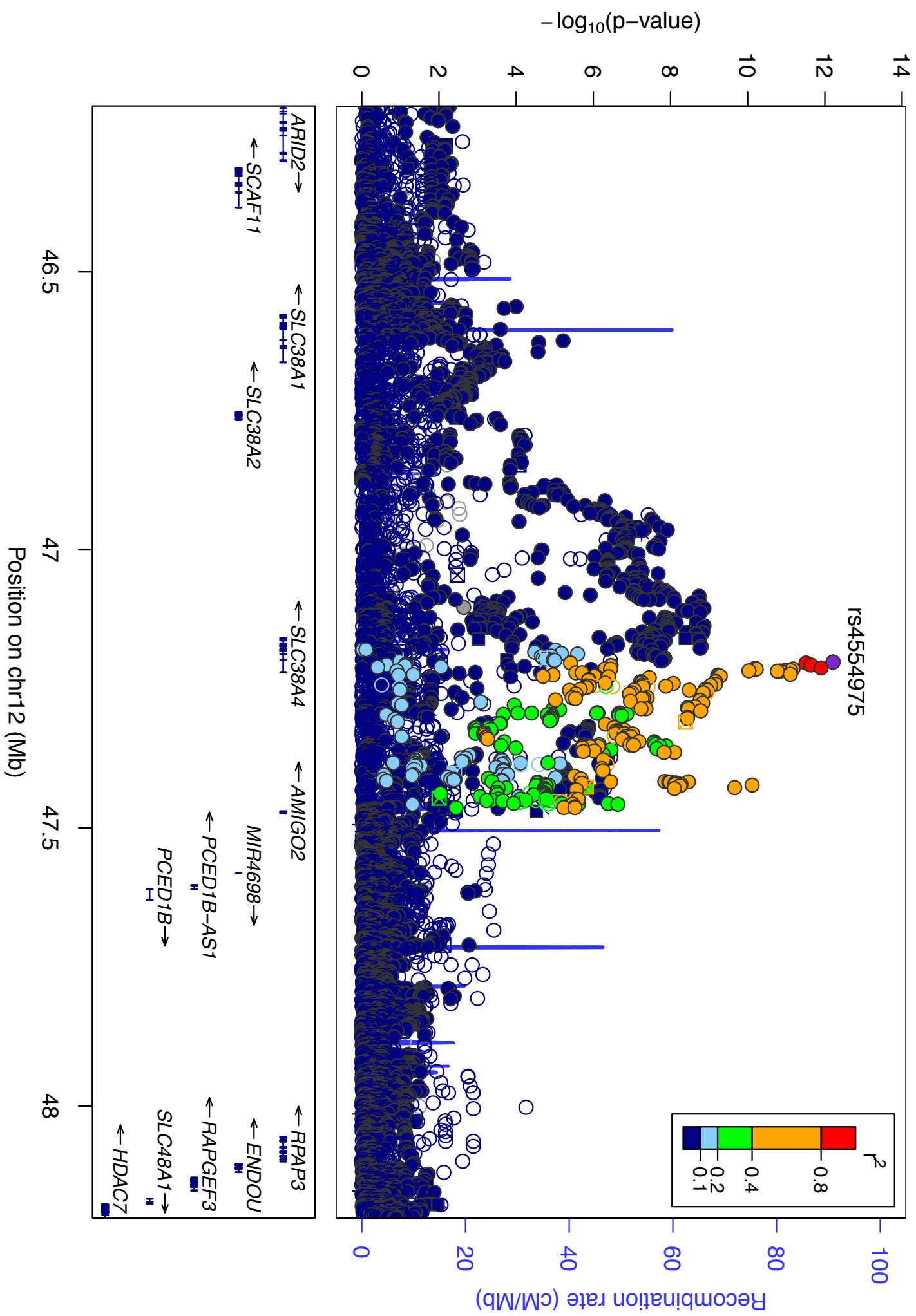


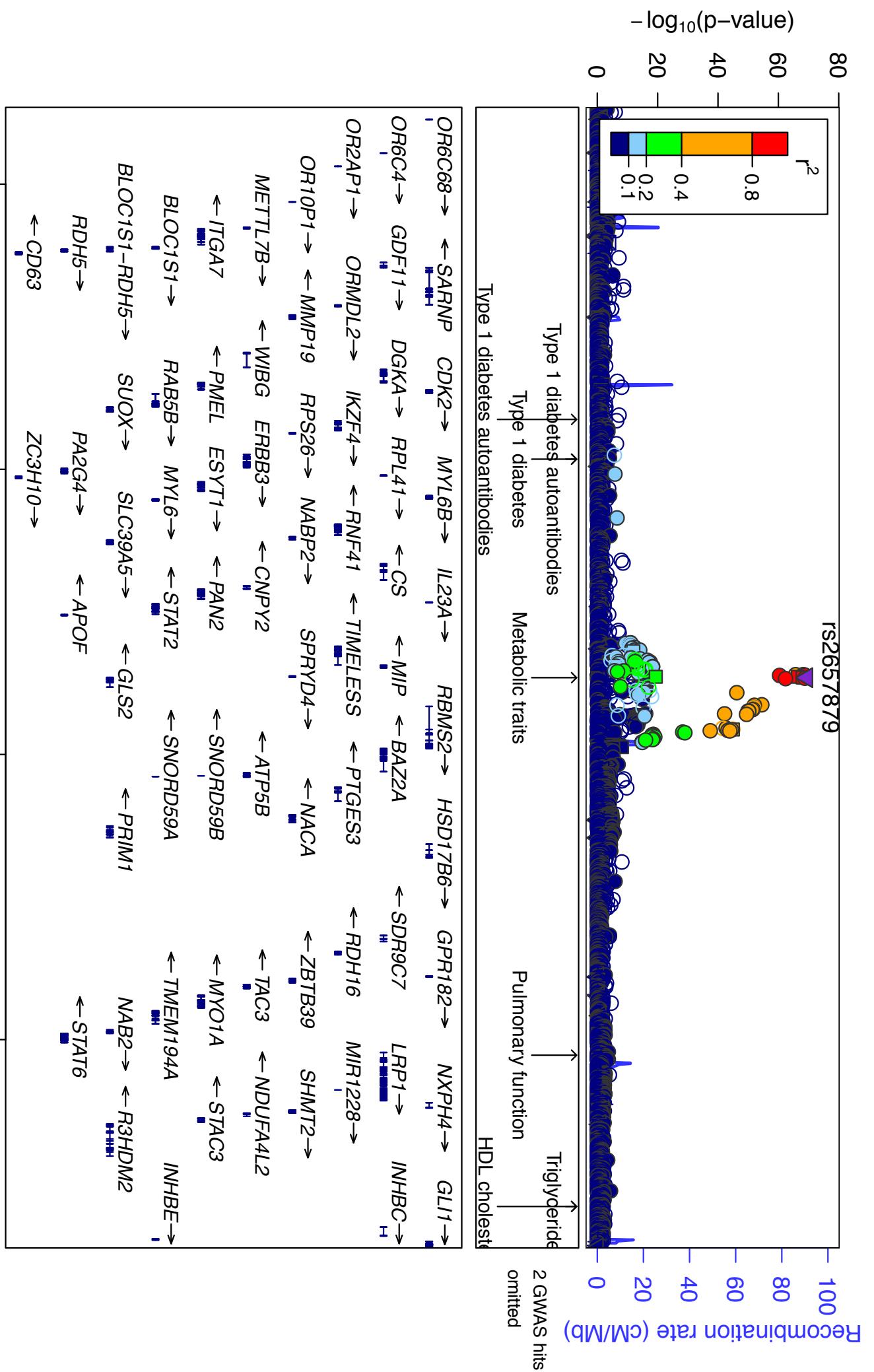


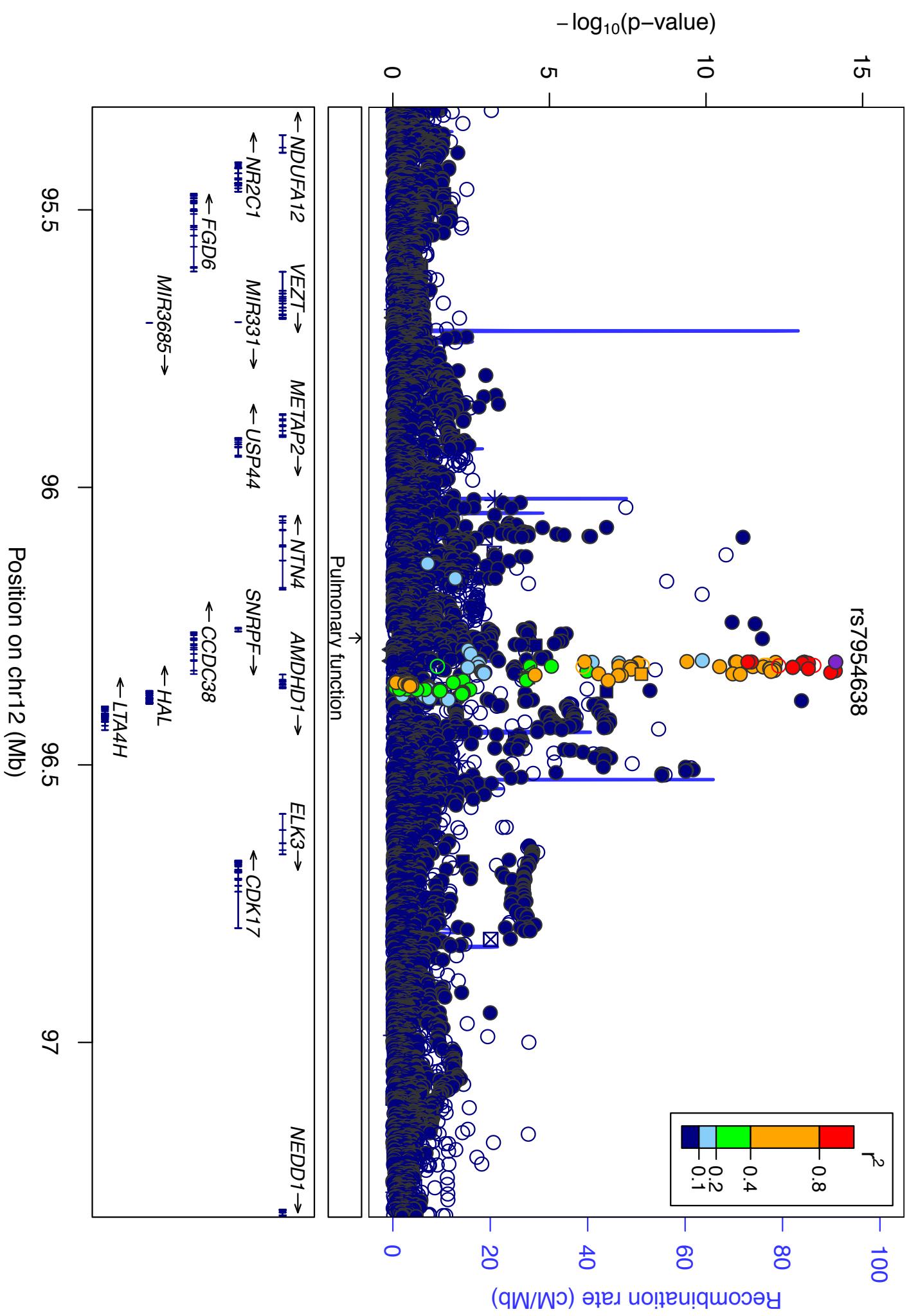


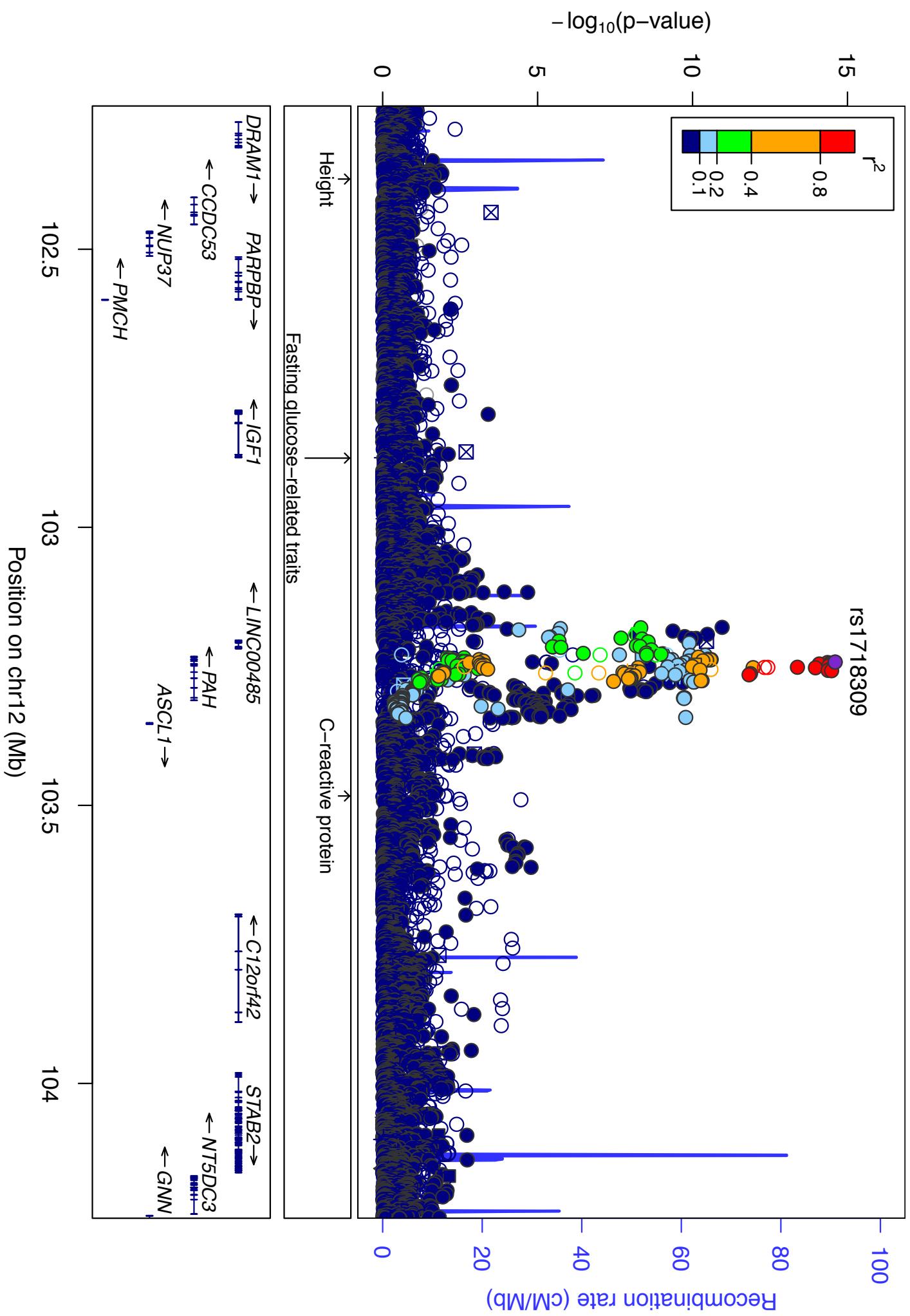


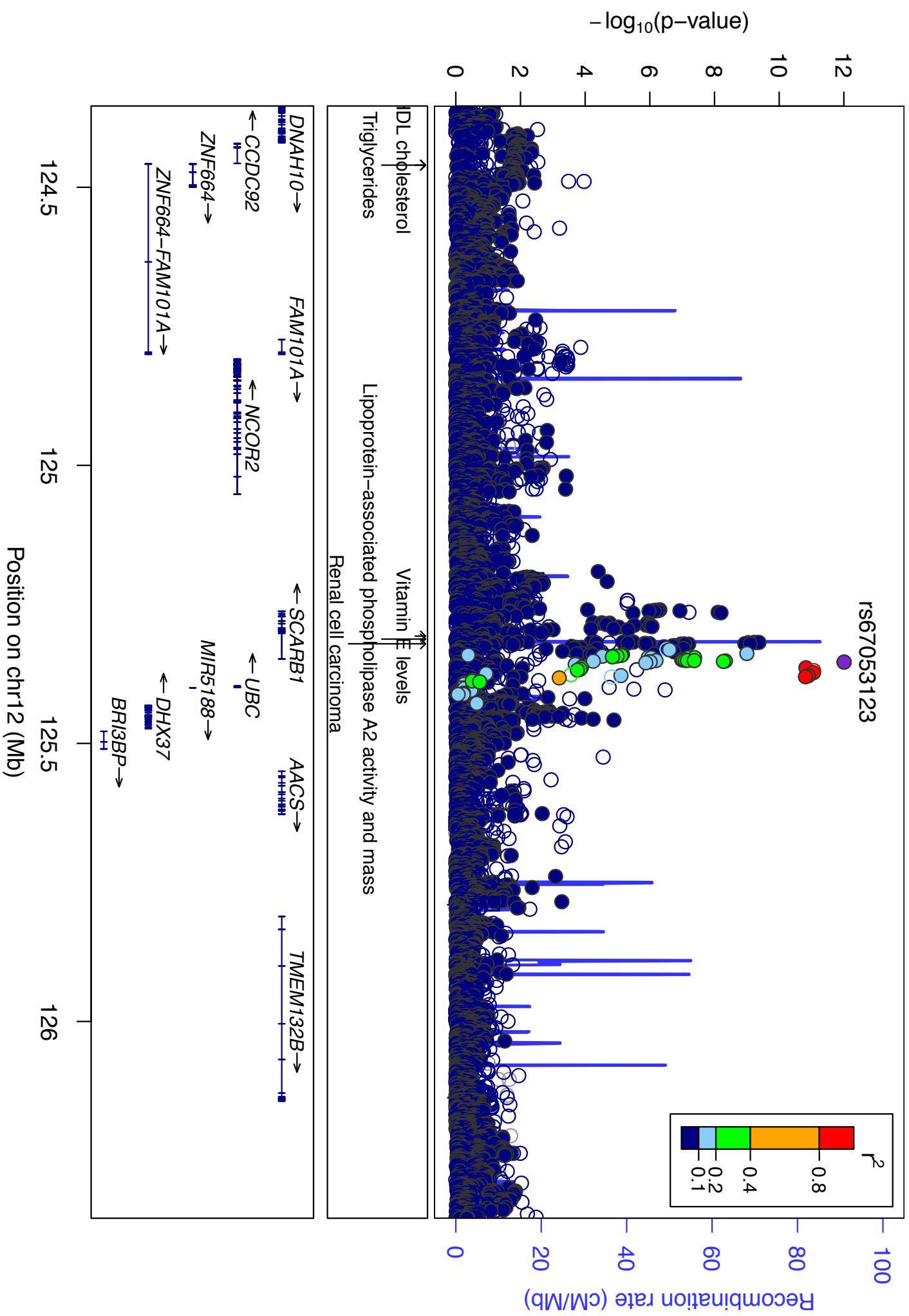


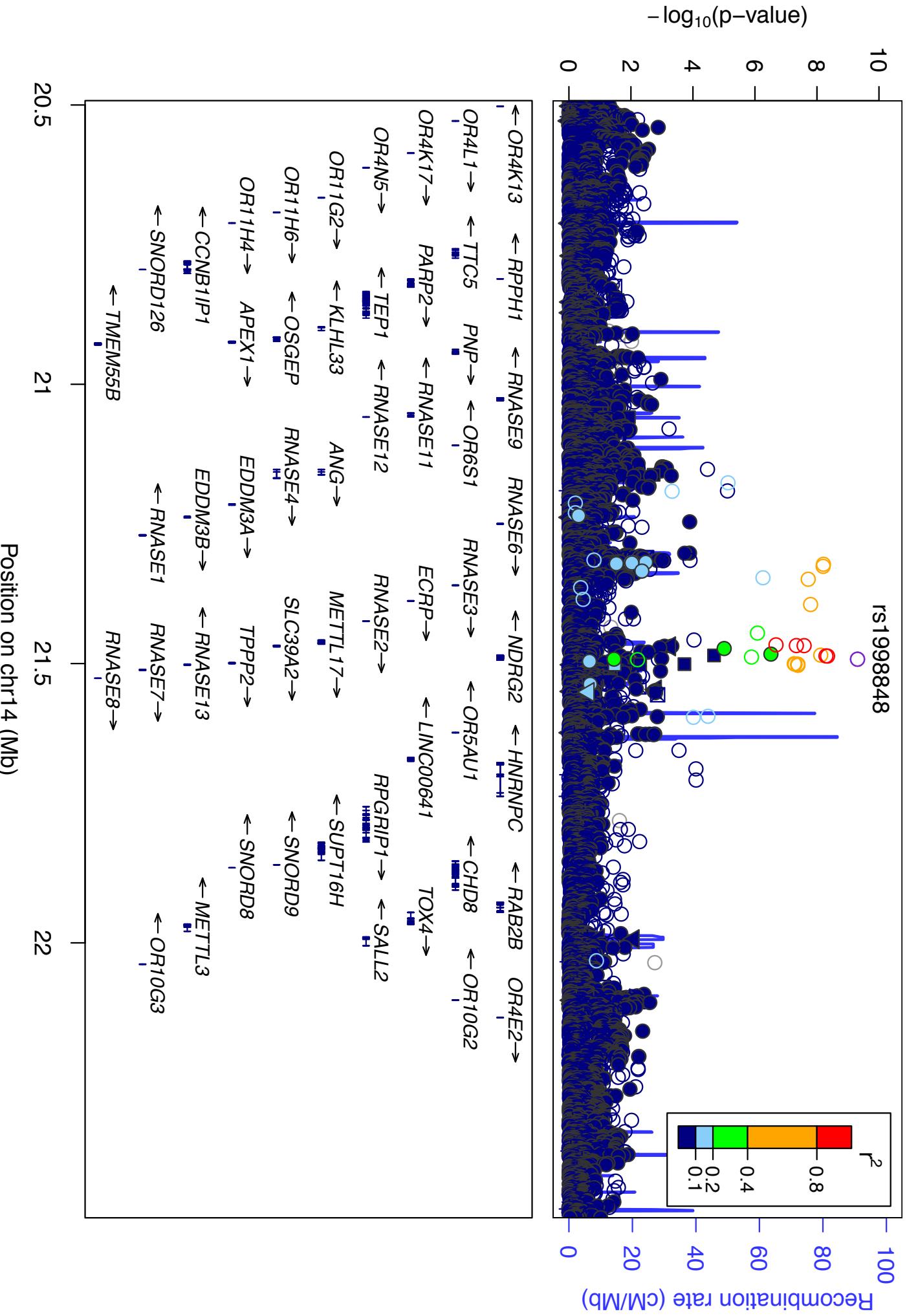


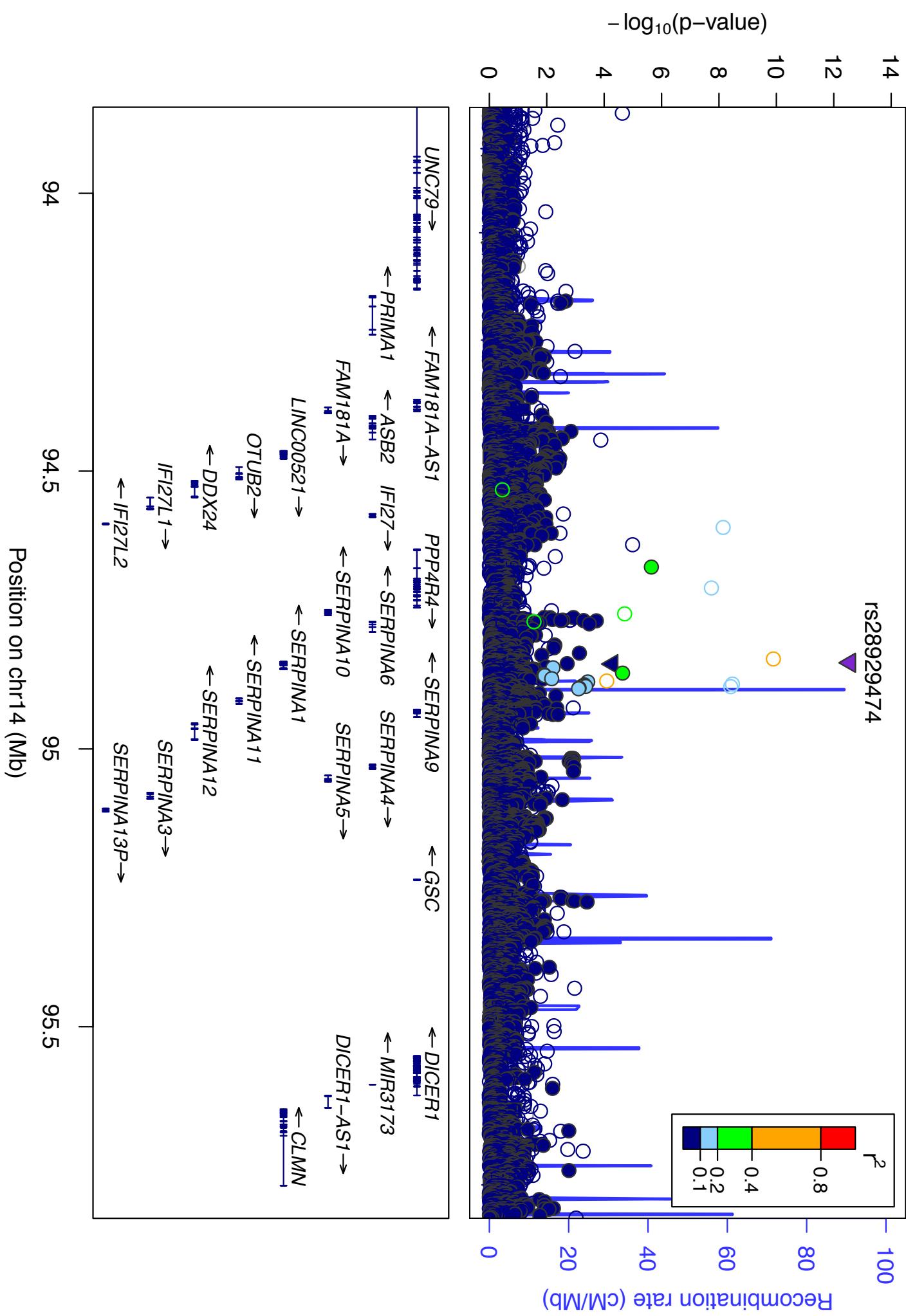


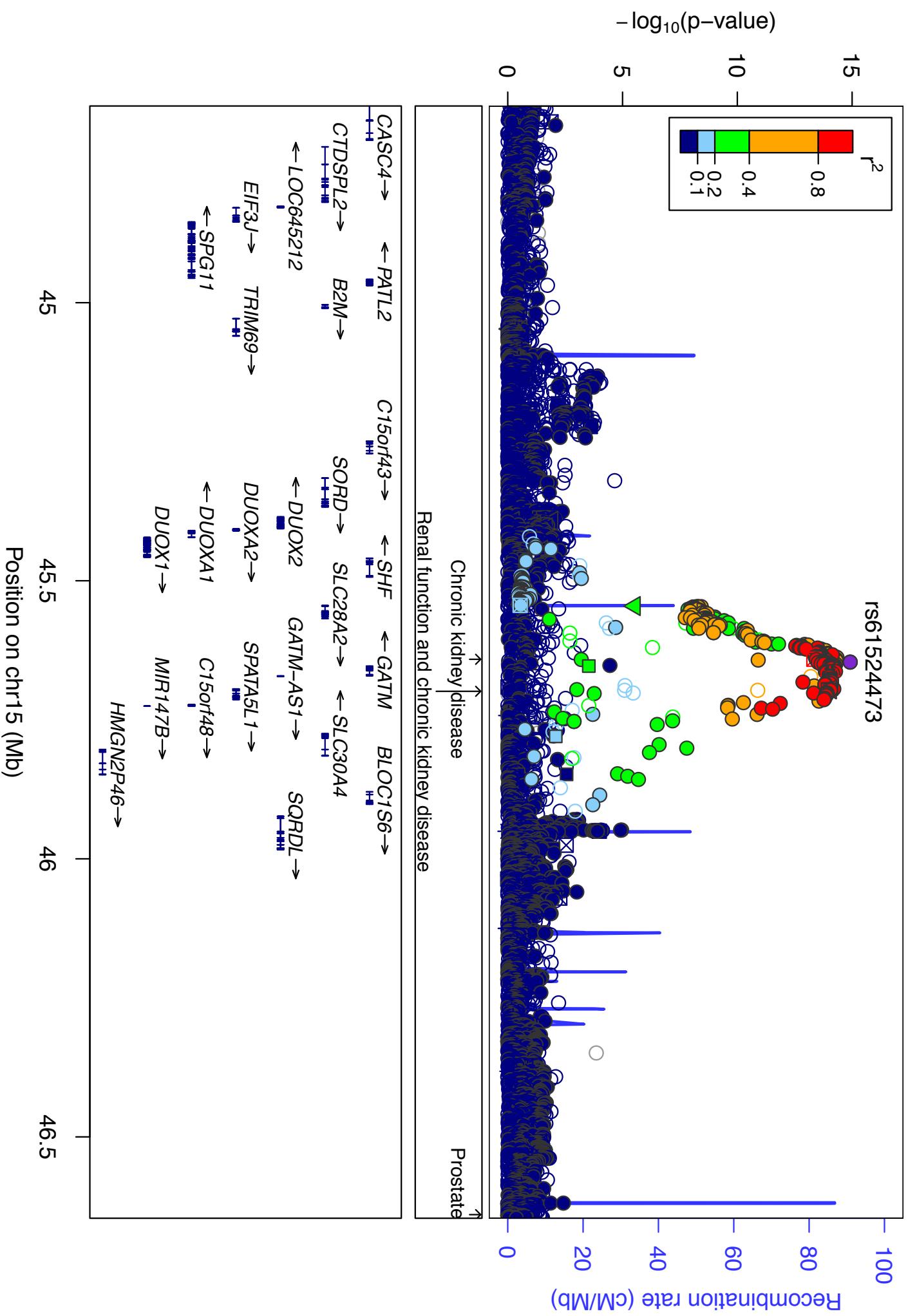


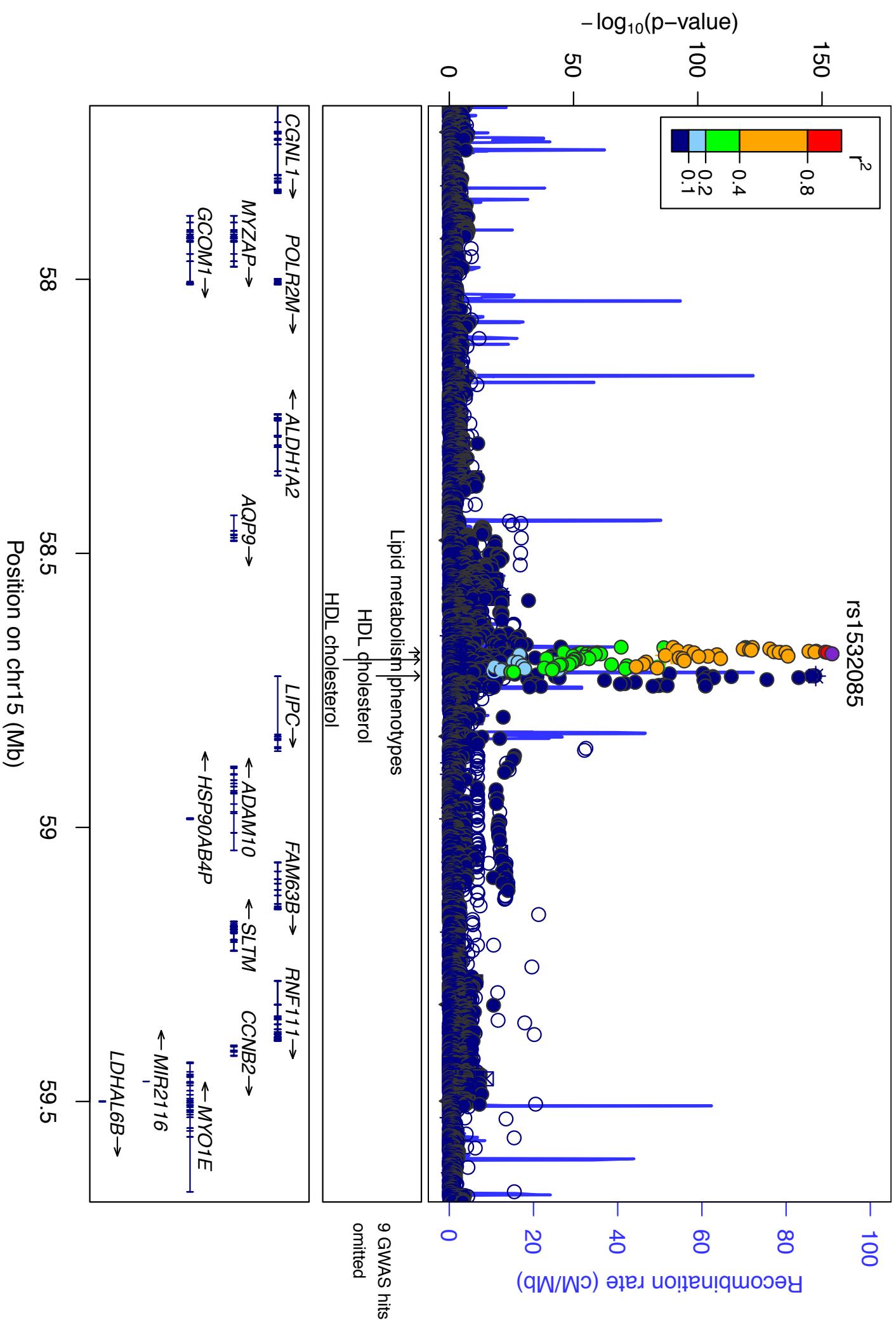


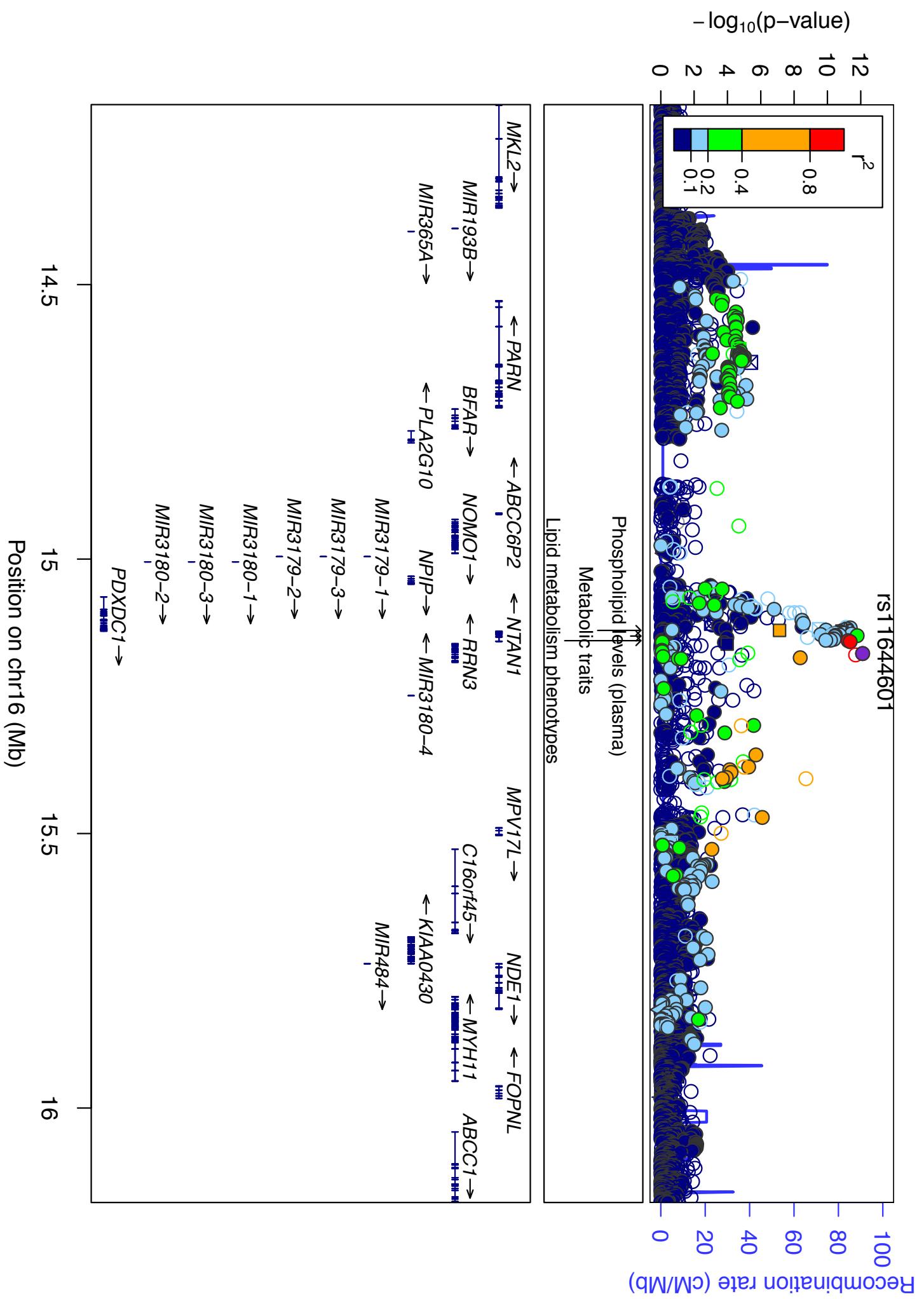


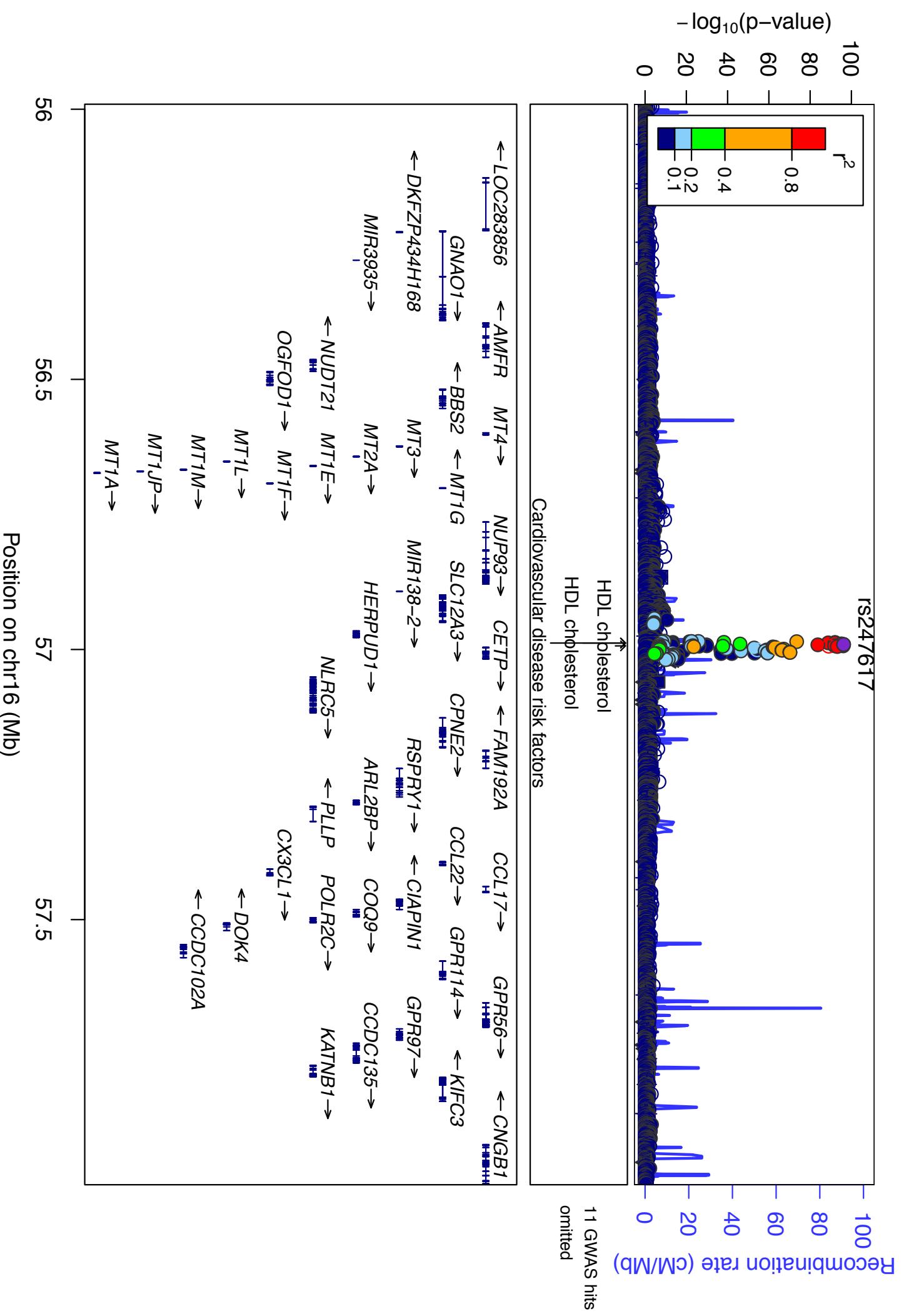


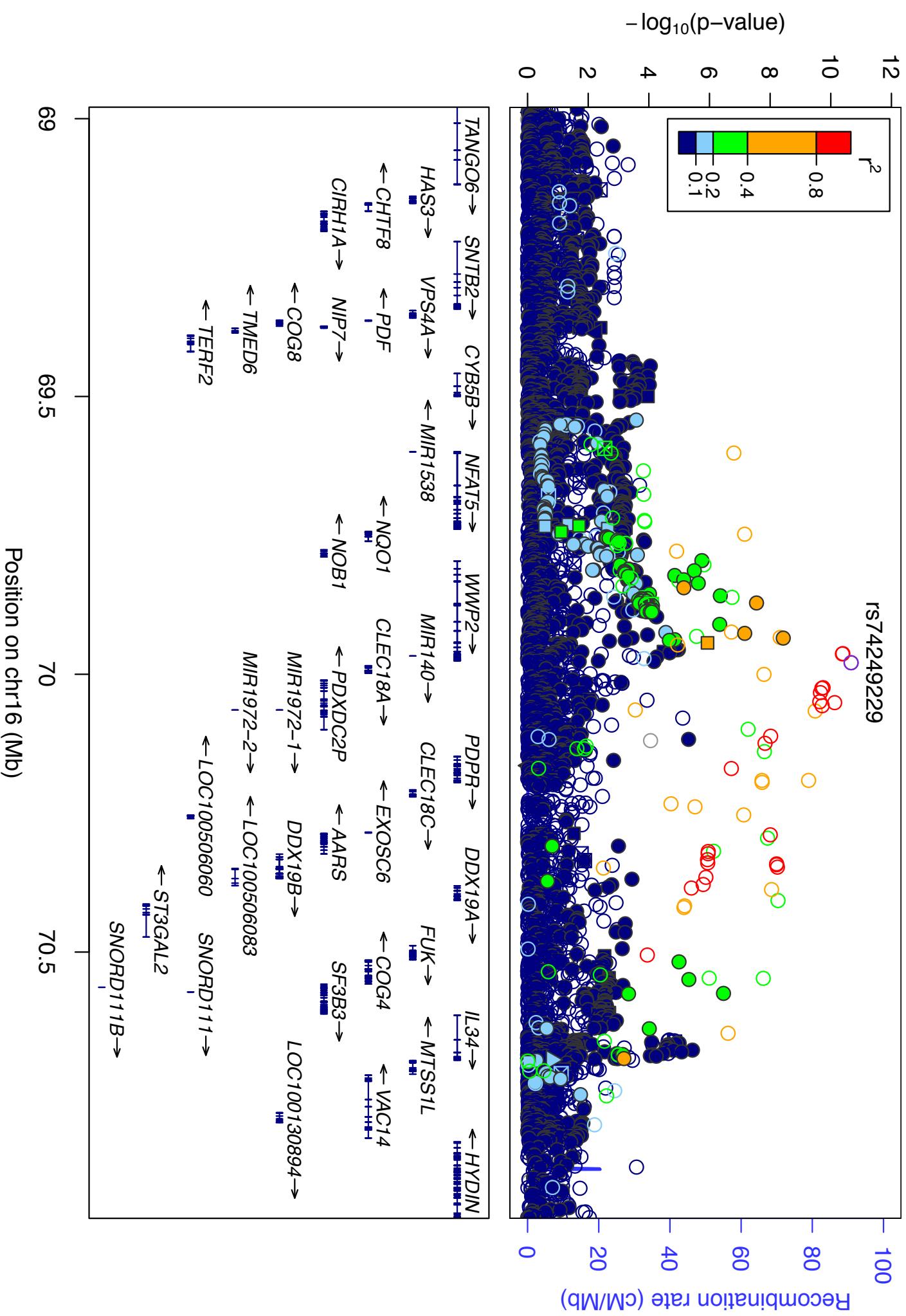


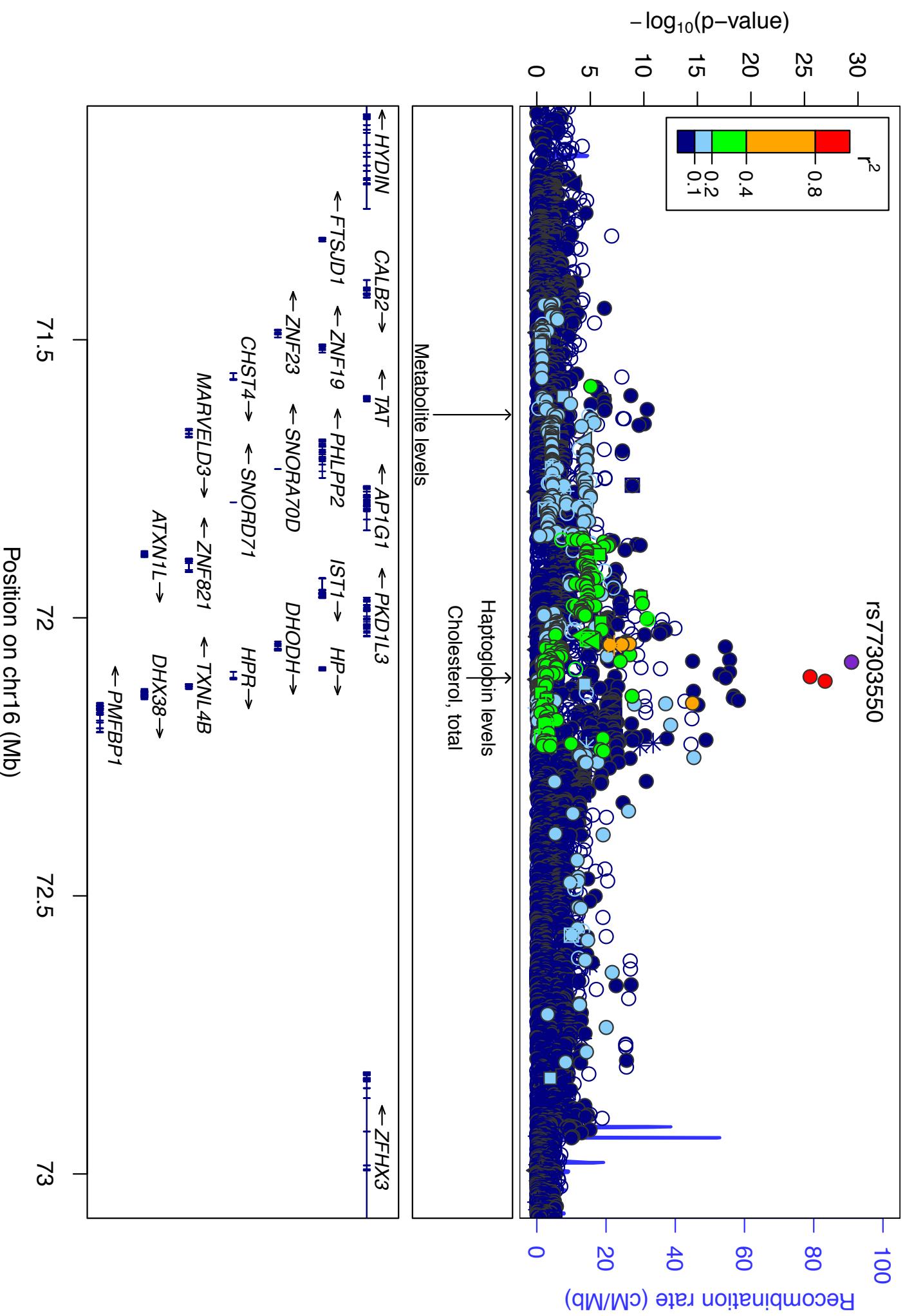


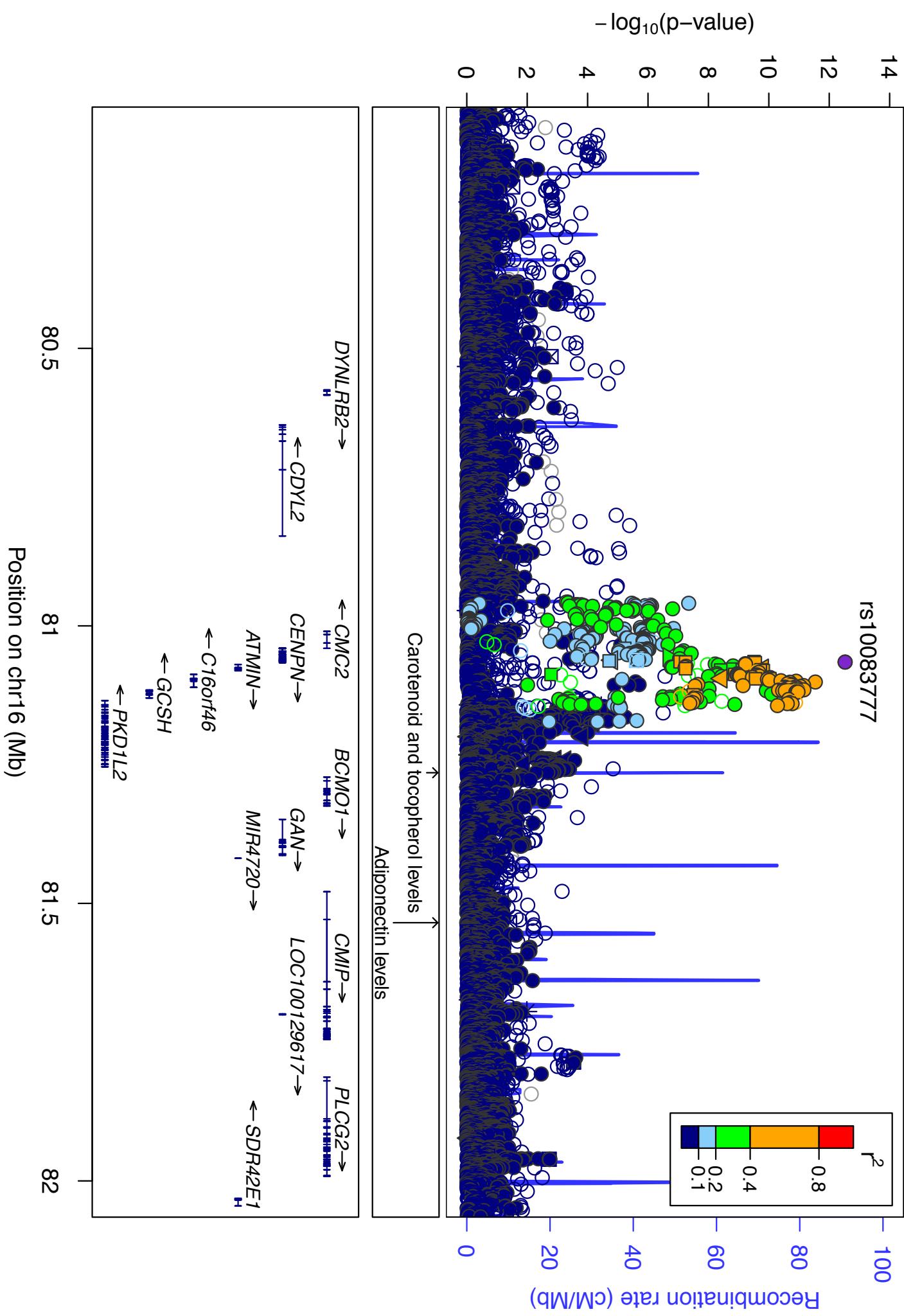


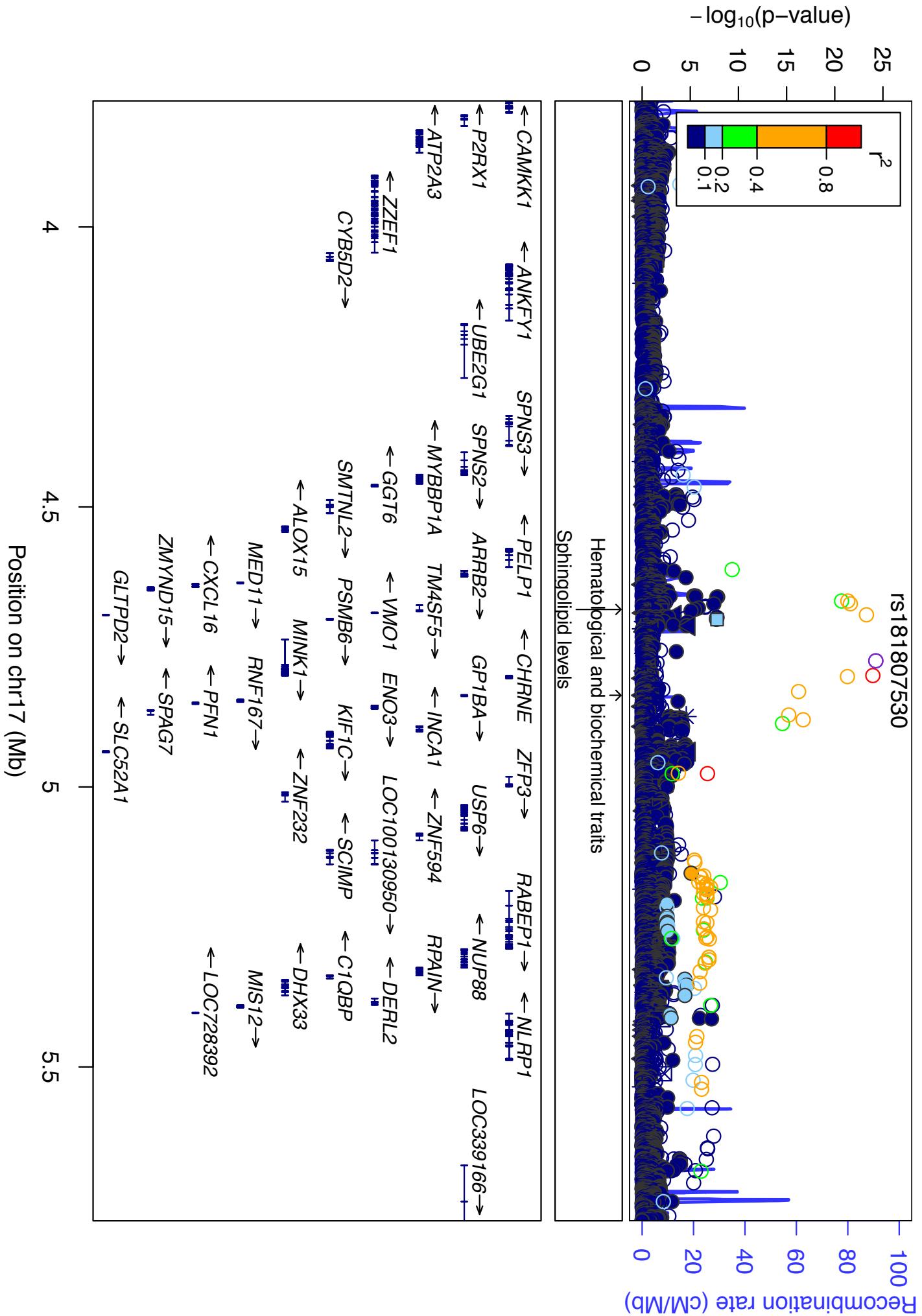


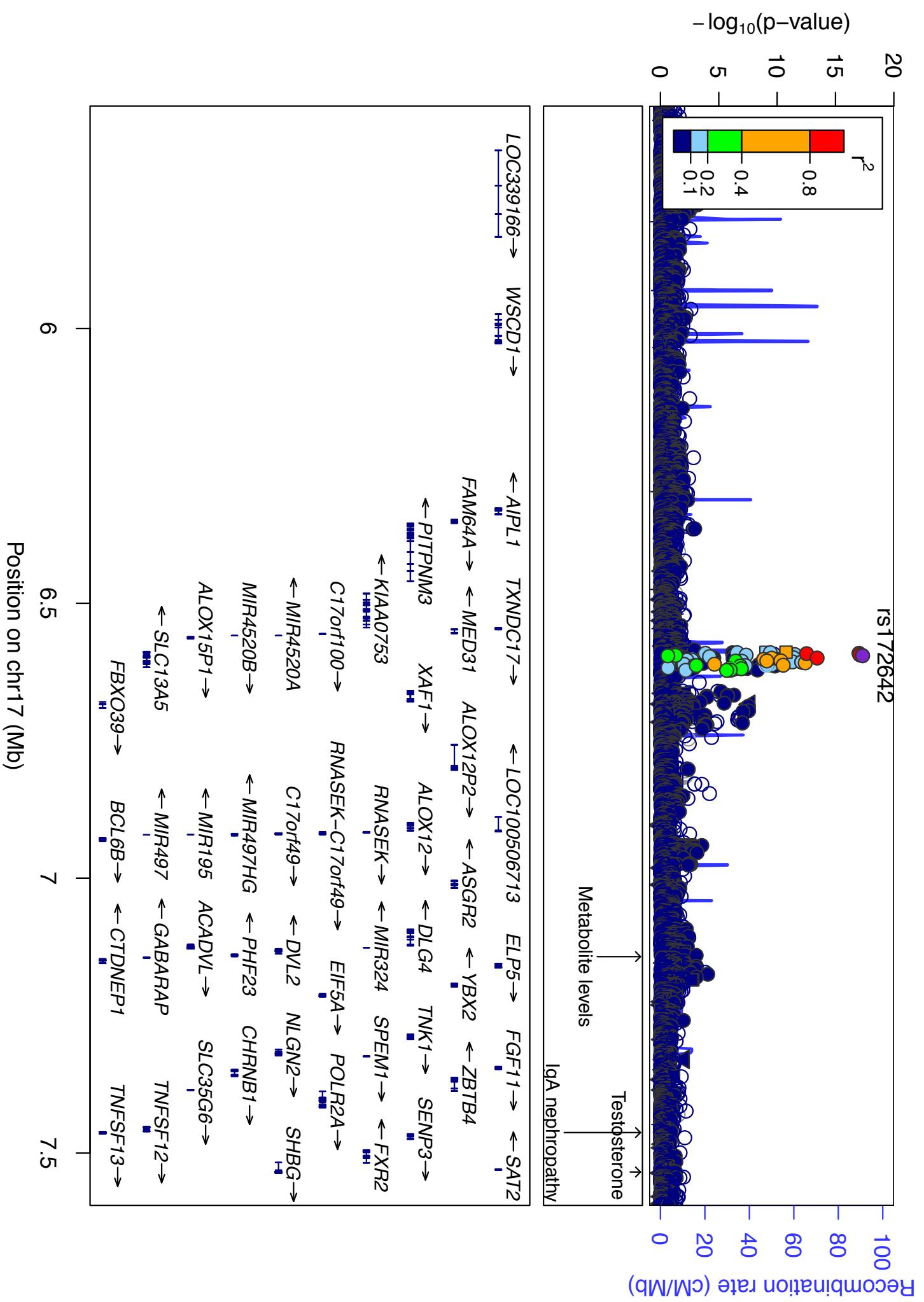


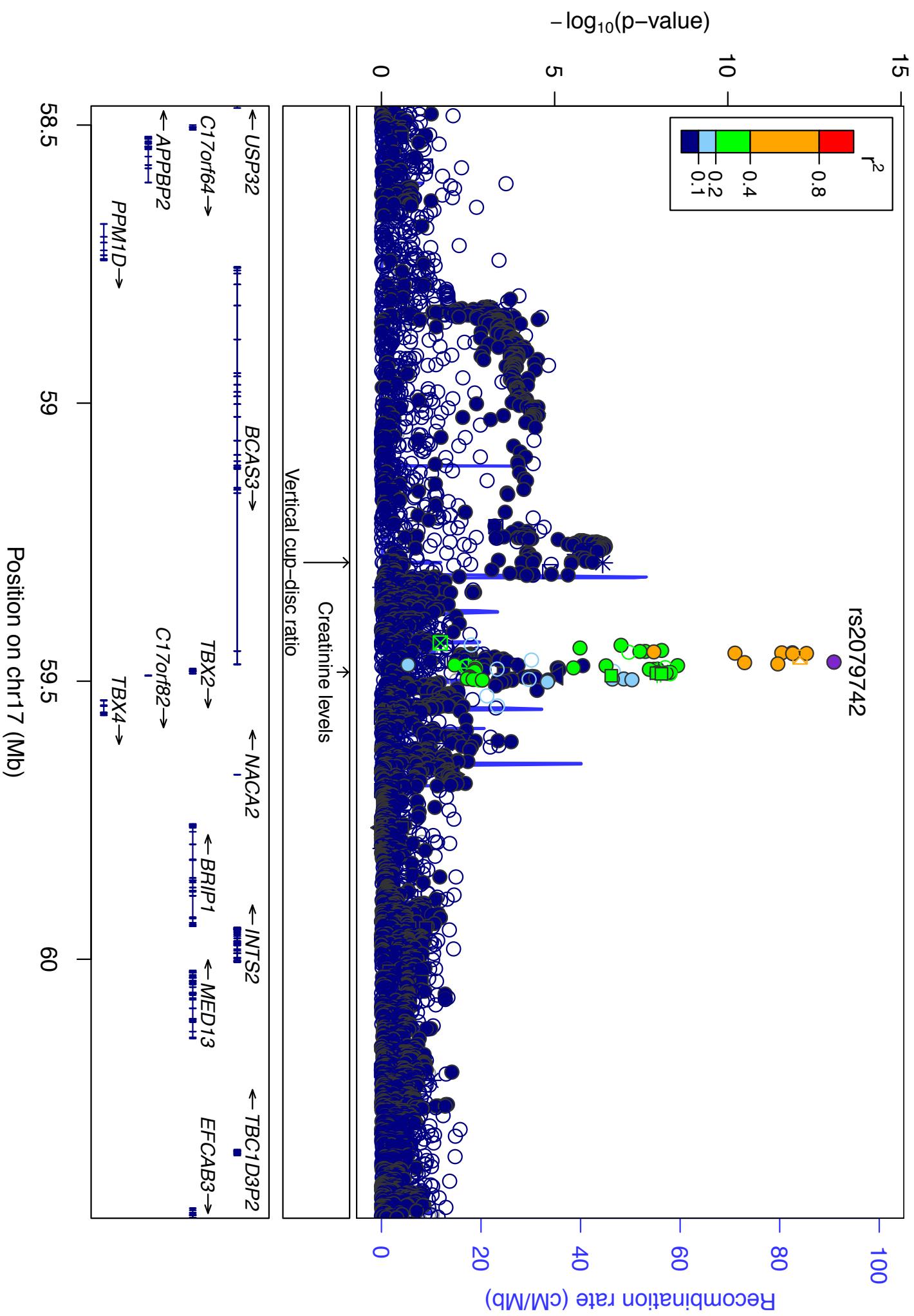


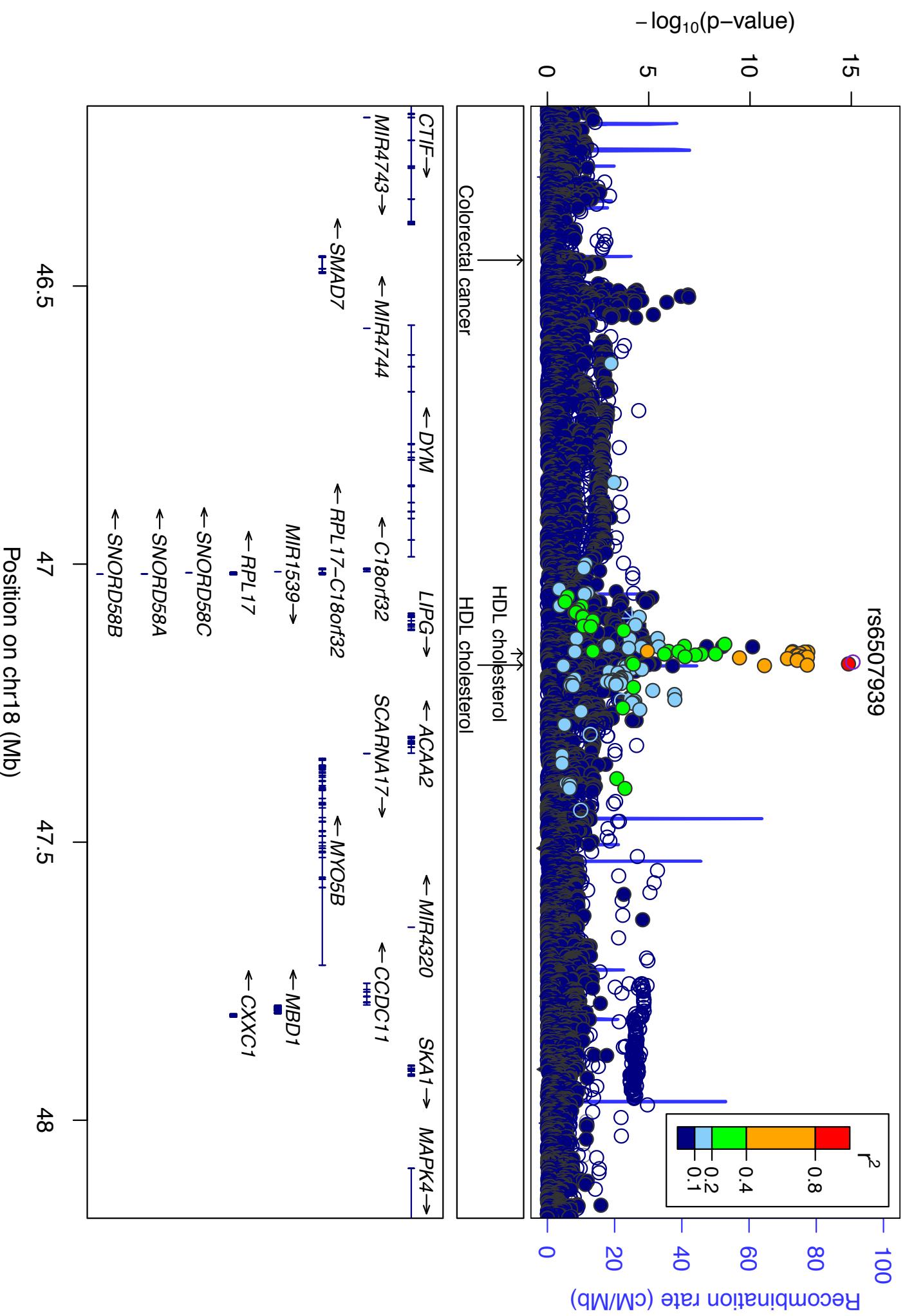


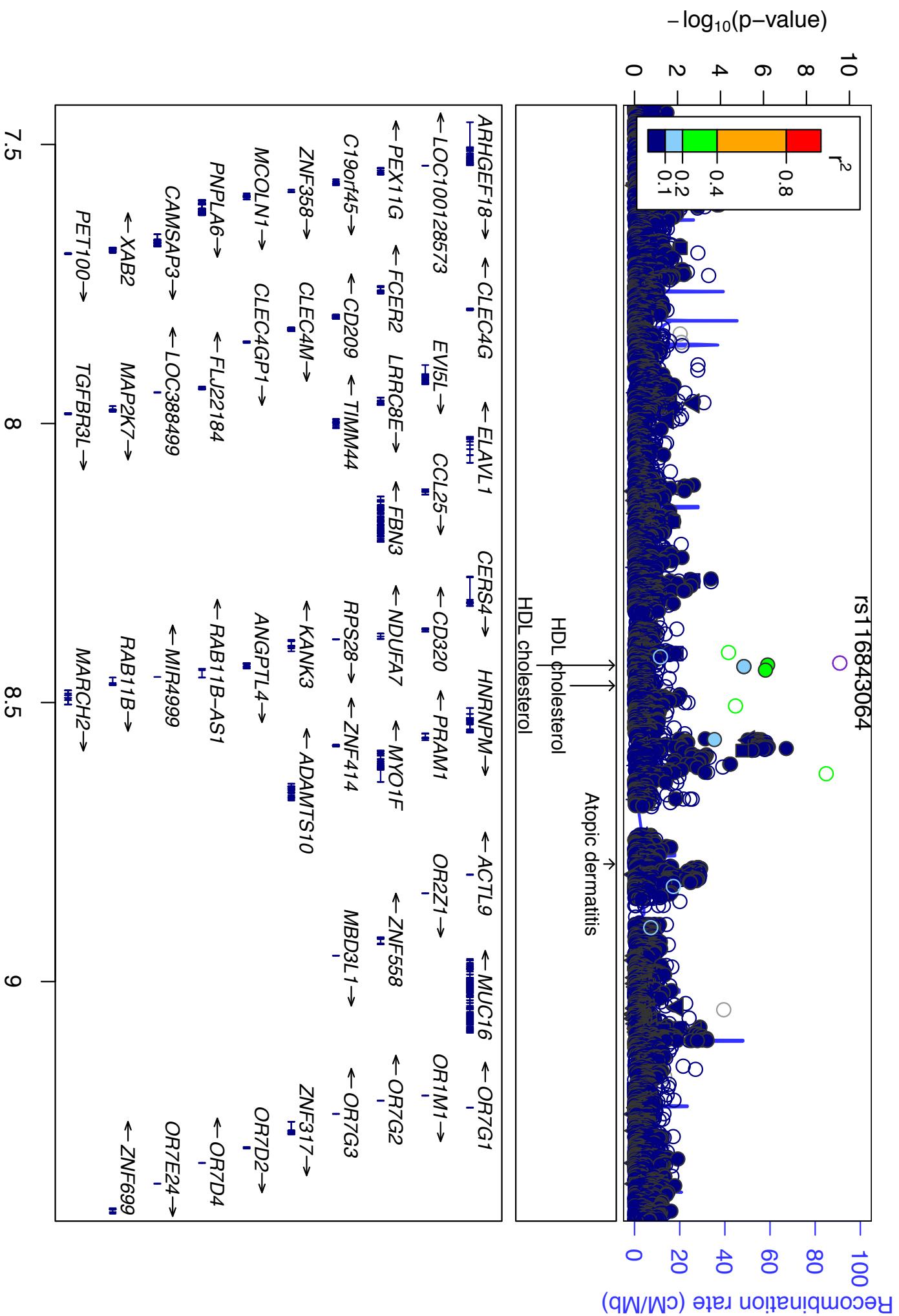


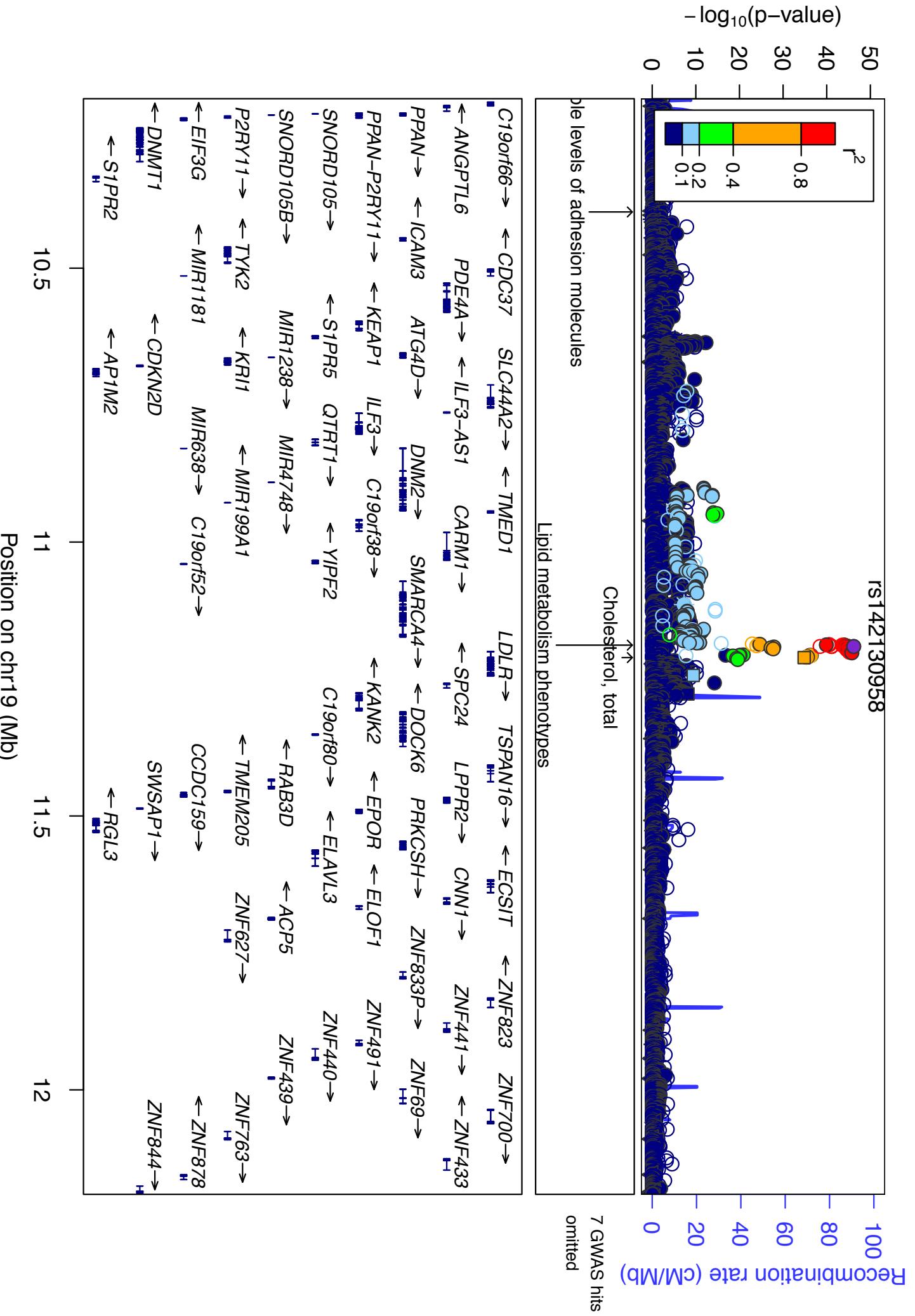


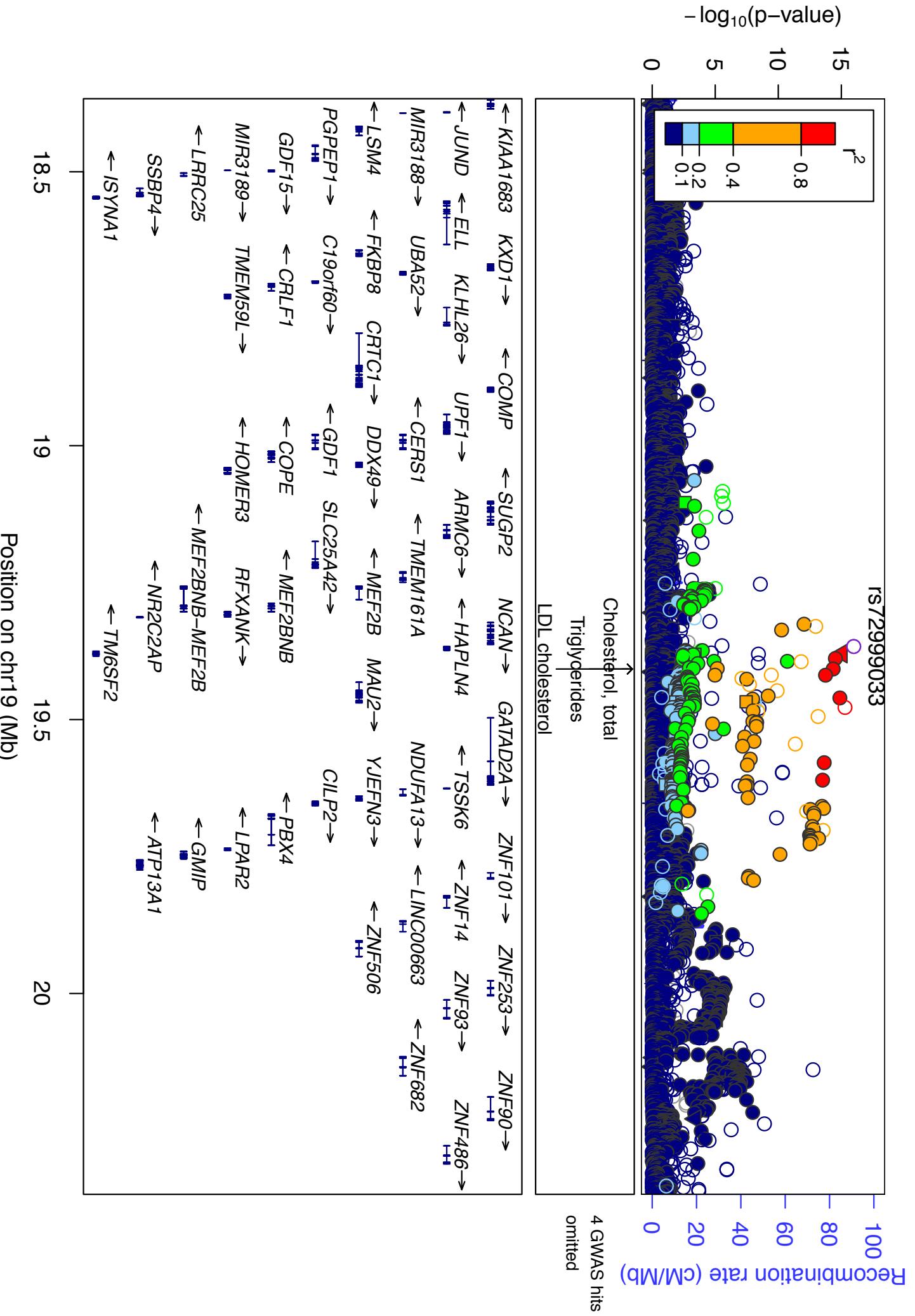


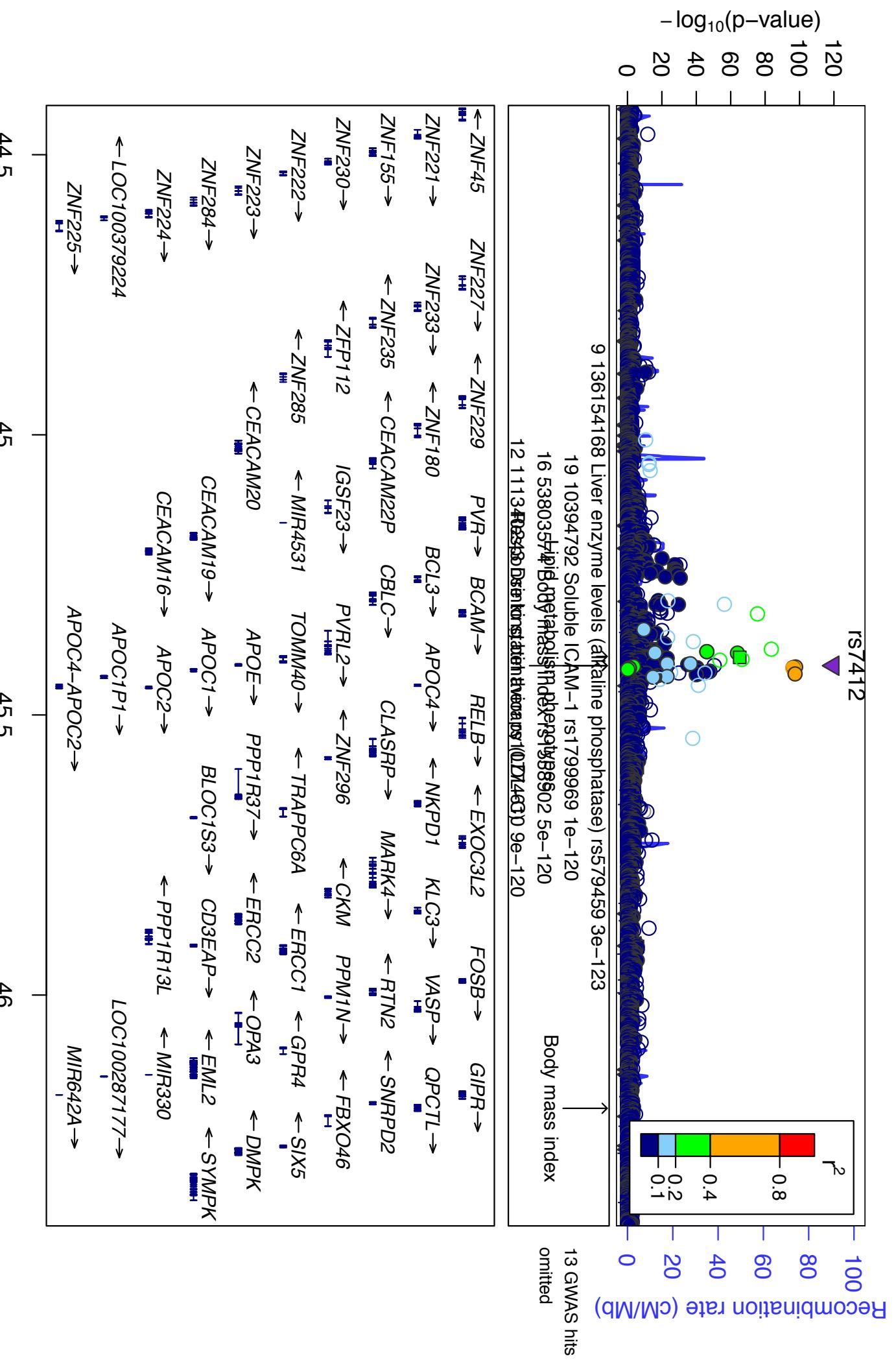


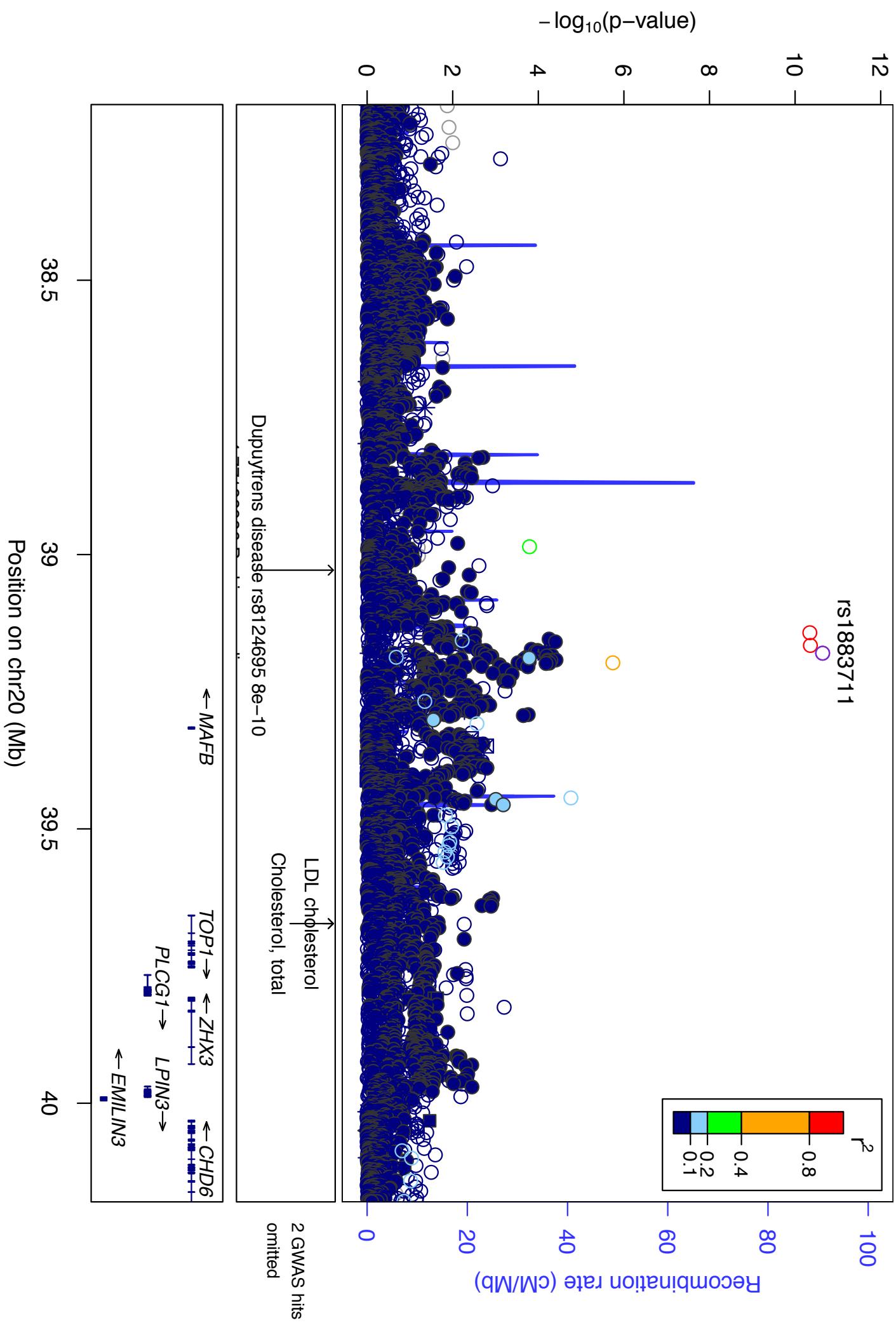


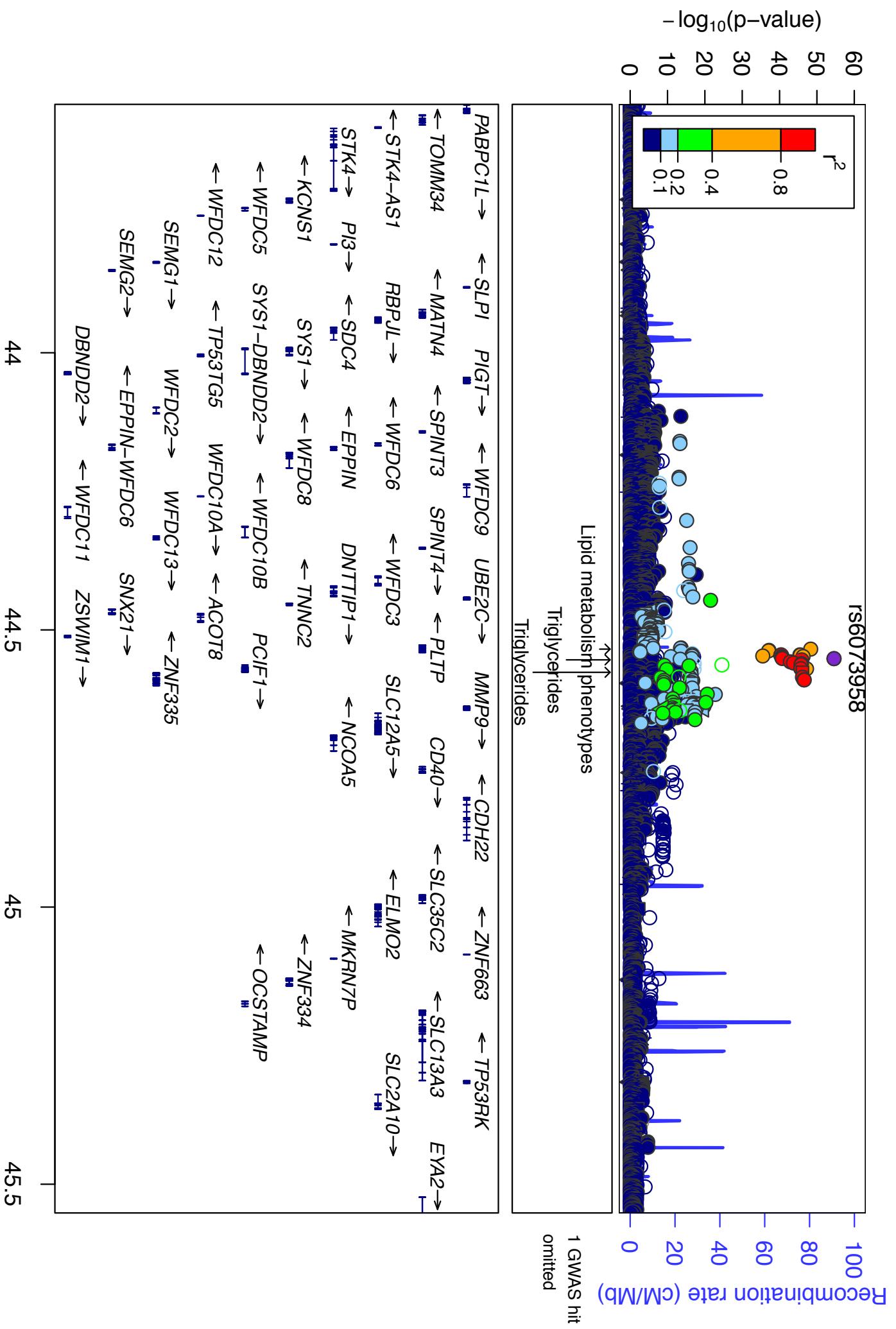












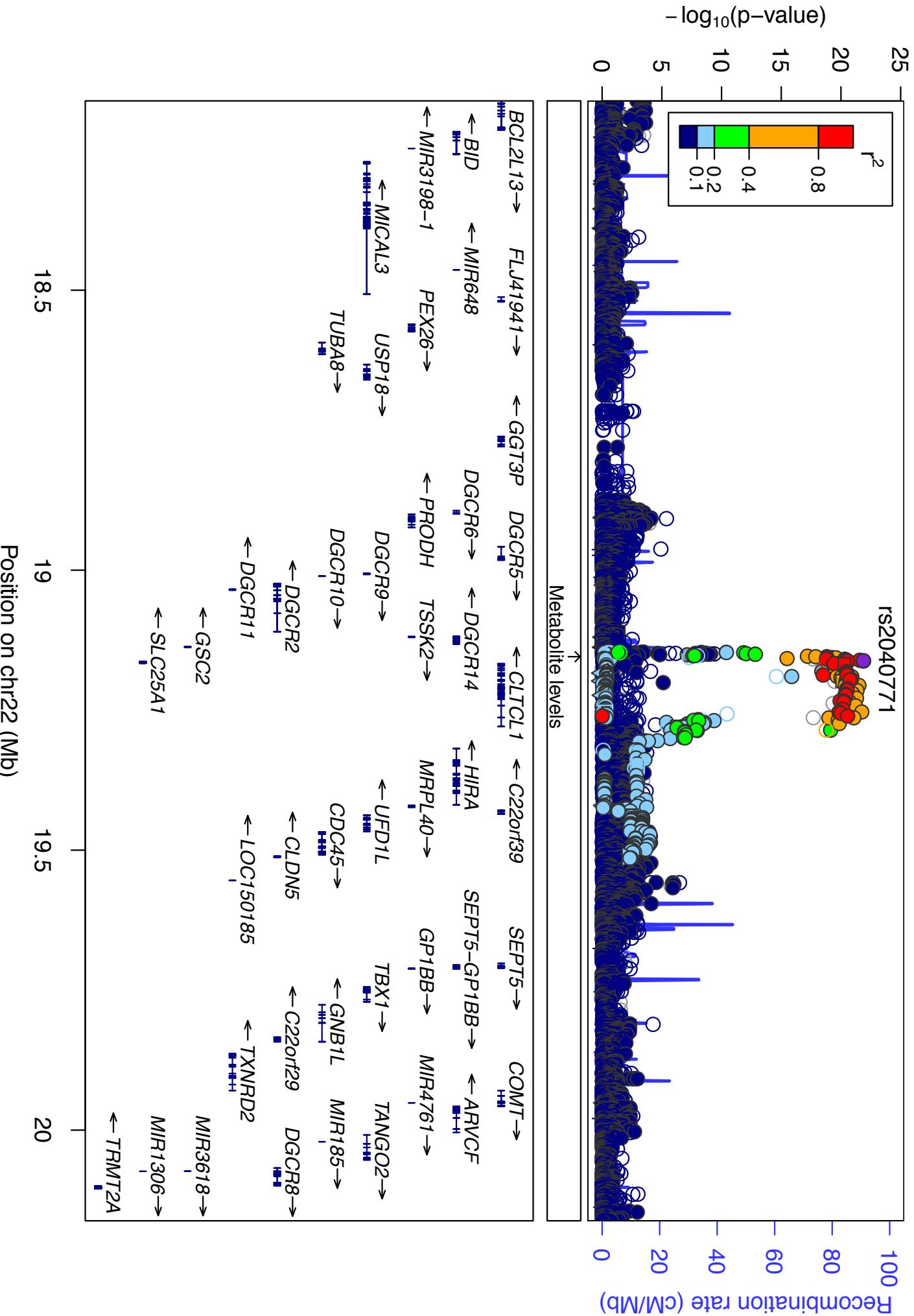
Position on chr20 (Mb)

44

44.5

45

45.5



Supplementary Figure 2. Forest plots of individual cohort and meta-analysis association results for all lead SNP associations in **Table 2** and **Supplementary Table 4**.

**SNP rs191448950, Chr 1, Pos 55584844, L.LDL.FC**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



-1 0 0.5 1 1.5

**SNP rs1168041, Chr 1, Pos 62960250, Serum.TG**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



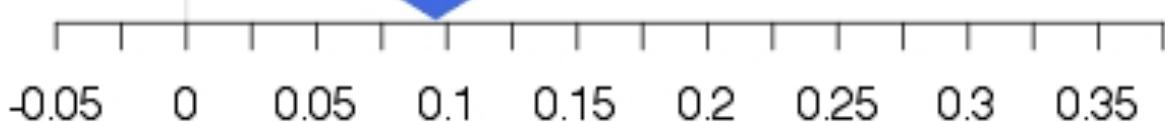
PROTE



FTC



**Meta-analysis**



**SNP chr1:109818158:I, Chr 1, Pos 109818158, L.LDL.FC**

Study

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

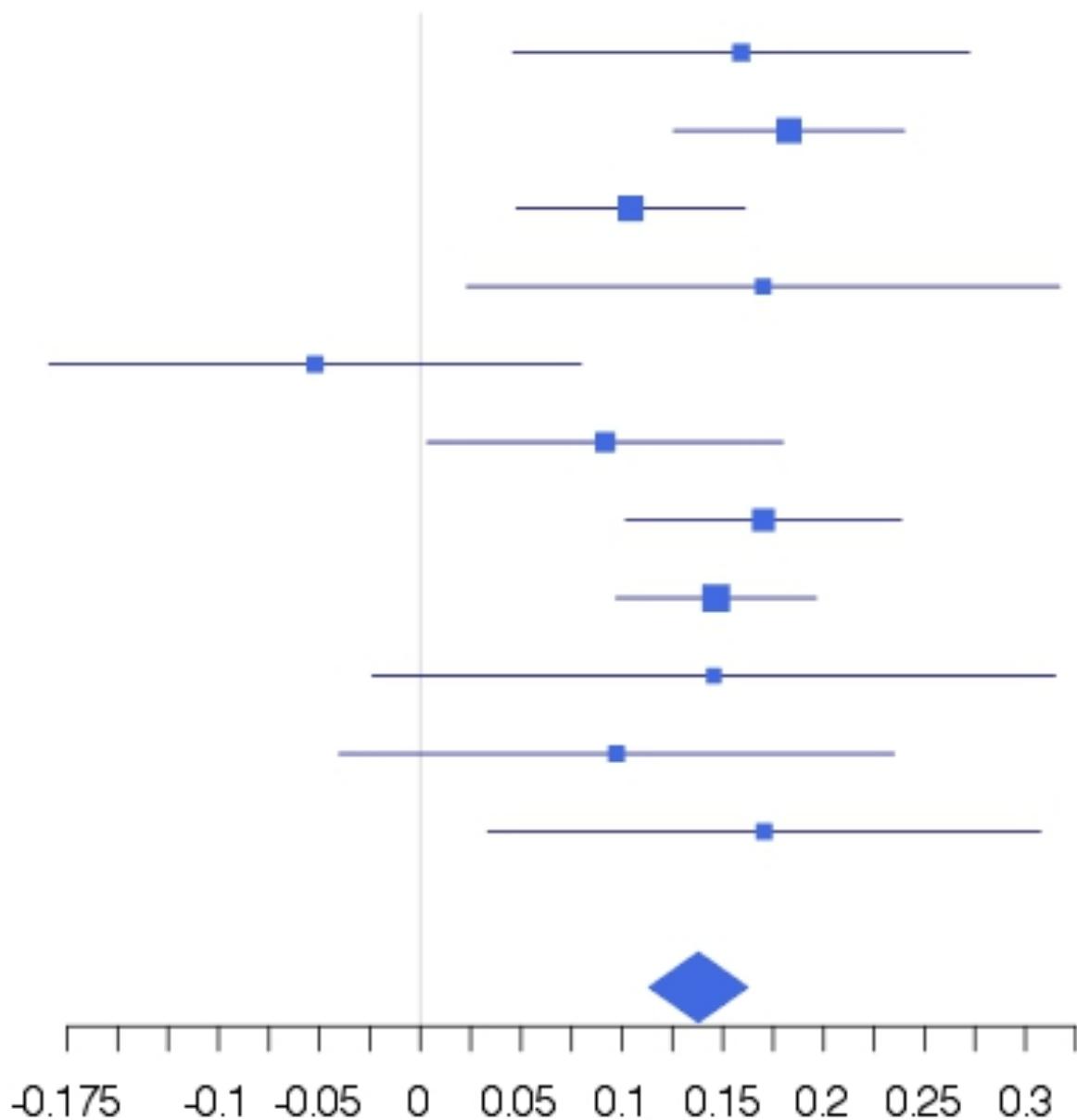
NFBC 1966

PredictCVD

PROTE

FTC

**Meta-analysis**



**SNP rs4503368, Chr 1, Pos 161190250, XL.HDL.PL**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



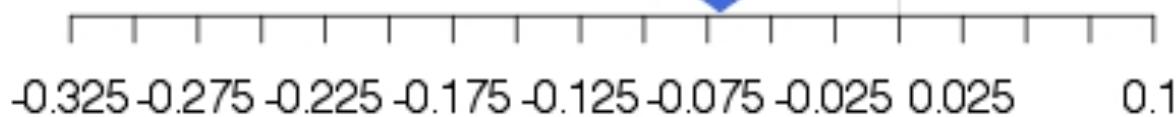
PROTE



FTC



**Meta-analysis**



**SNP rs590820, Chr 1, Pos 230309619, M.HDL.FC**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



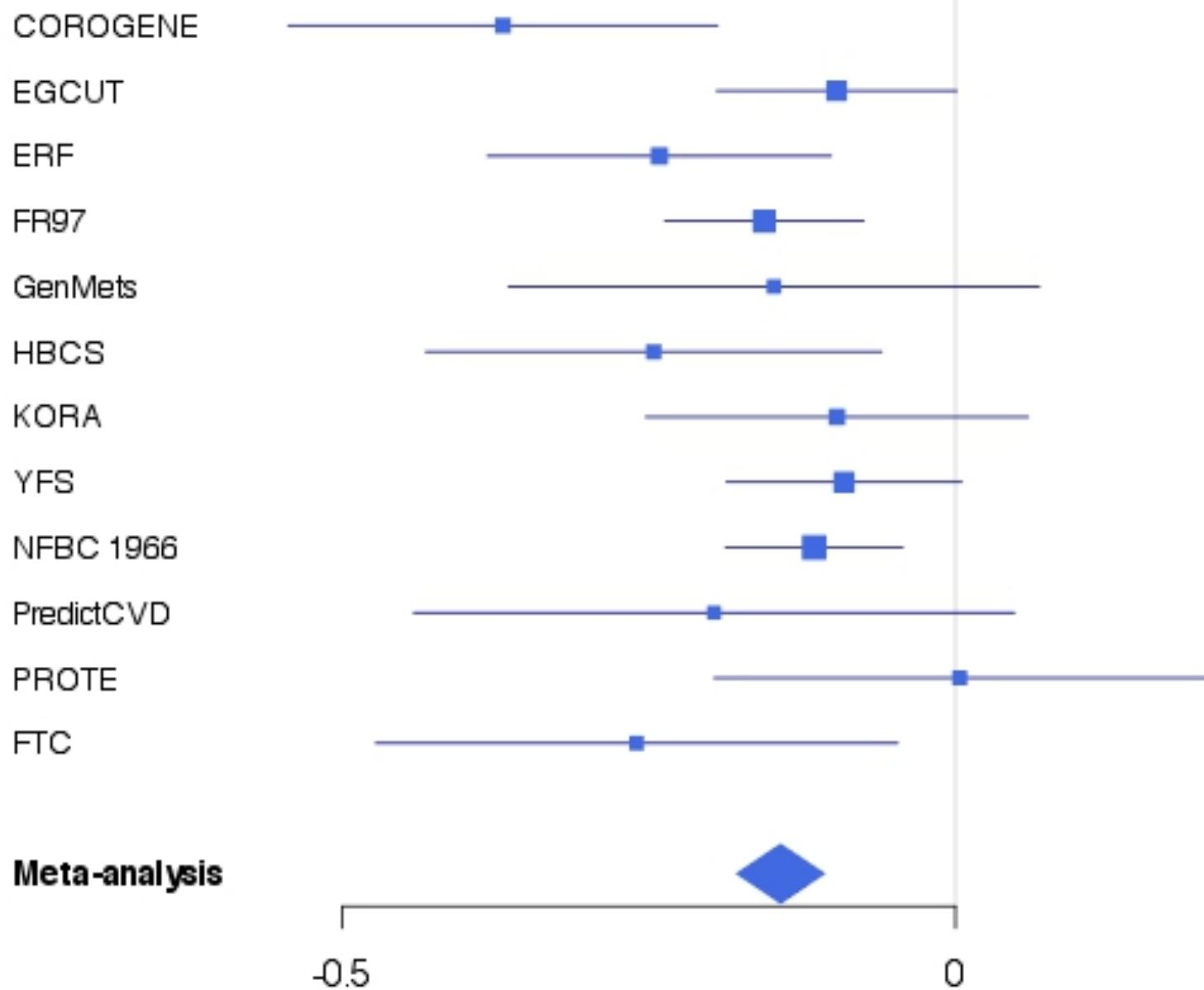
**Meta-analysis**



-0.275 -0.225 -0.175 -0.125 -0.075 -0.025 0.025 0.075 0.125 0.175

### **SNP rs6756629, Chr 2, Pos 44065090, LDL.C**

Study



## SNP rs10211524, Chr 2, Pos 65208074, Val

### Study

COROGENE

EGCUT

ERF

FR97

GenMets

HBCS

KORA

YFS

LLS

NFBC 1966

NTR

PredictCVD

PROTE

FTC

### Meta-analysis

-0.075 -0.025 0.025 0.075 0.125 0.175 0.225 0.275 0.325



### **SNP rs560887, Chr 2, Pos 169763148, Glc**

#### **Study**

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



NTR



PredictCVD



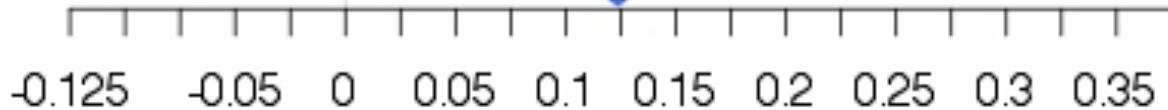
PROTE



FTC



#### **Meta-analysis**



**SNP rs1047891, Chr 2, Pos 211540507, Gly**

Study

COROGENE



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



PredictCVD



FTC



**Meta-analysis**



0 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70

### SNP rs952275, Chr 2, Pos 21221399, IDL.FC

Study

COROGENE

EGCUT

ERF

FR97

GenMets

HBCS

KORA

YFS

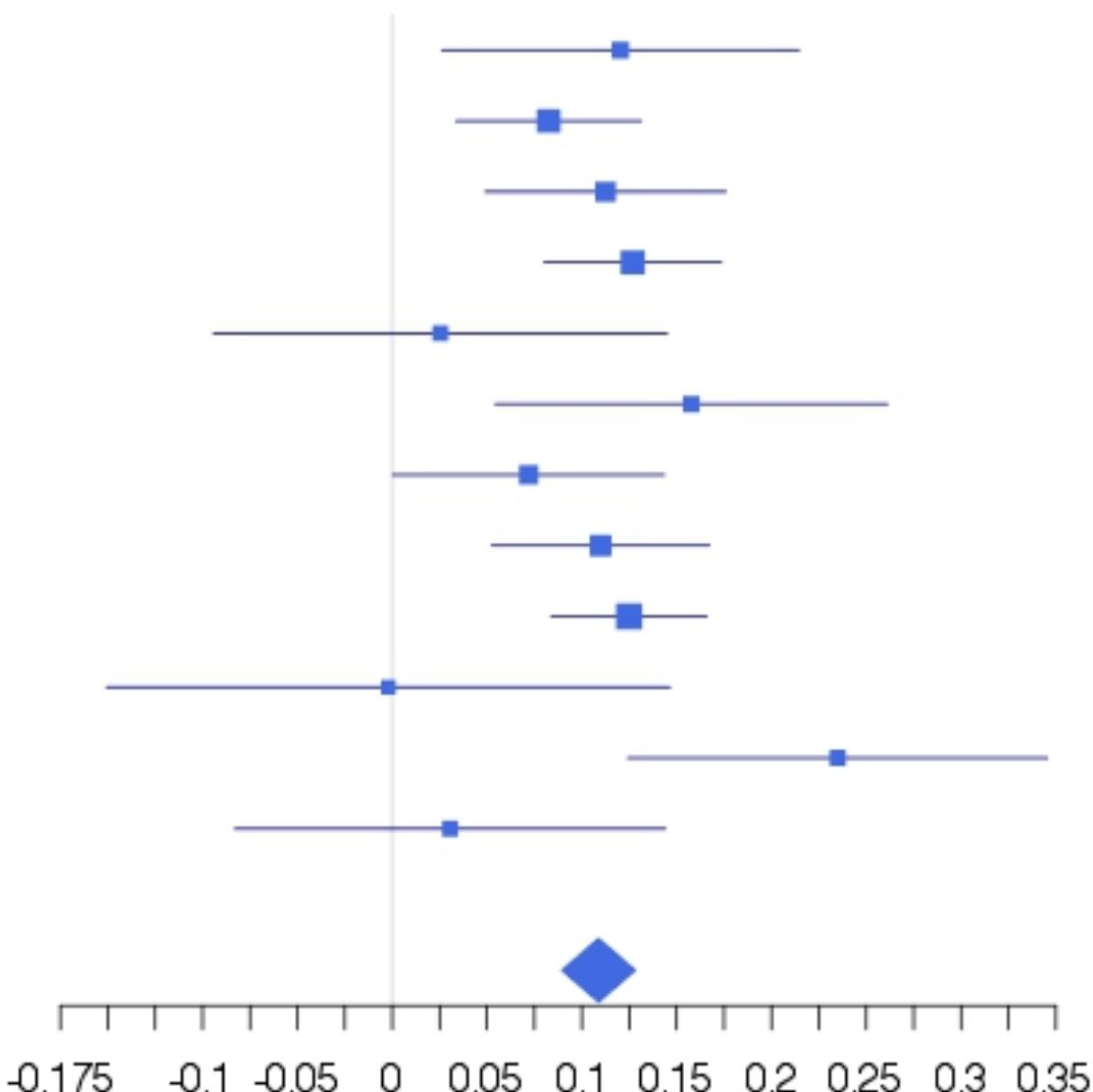
NFBC 1966

PredictCVD

PROTE

FTC

Meta-analysis



**SNP rs1260326, Chr 2, Pos 27730940, Ala**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



NTR



PredictCVD



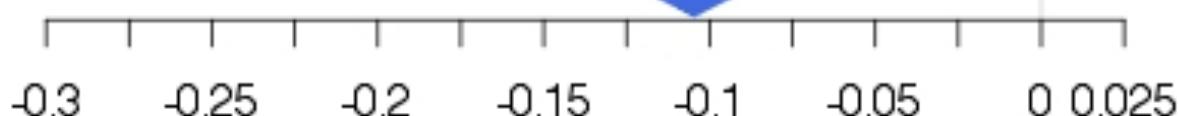
PROTE



FTC



**Meta-analysis**



### **SNP chr3:125905336:D, Chr 3, Pos 125905336, Gly**

#### **Study**

COROGENE



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



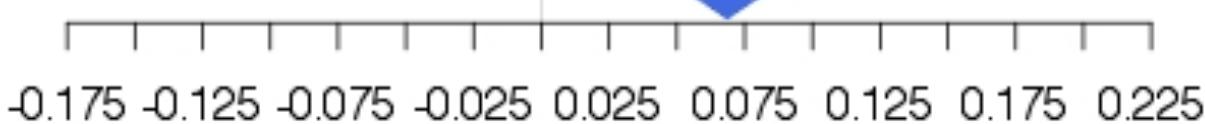
PredictCVD



FTC



#### **Meta-analysis**



**SNP rs184650103, Chr 4, Pos 74850649, Alb**

Study

COROGENE

EGCUT

FR97

GenMets

HBCS

YFS

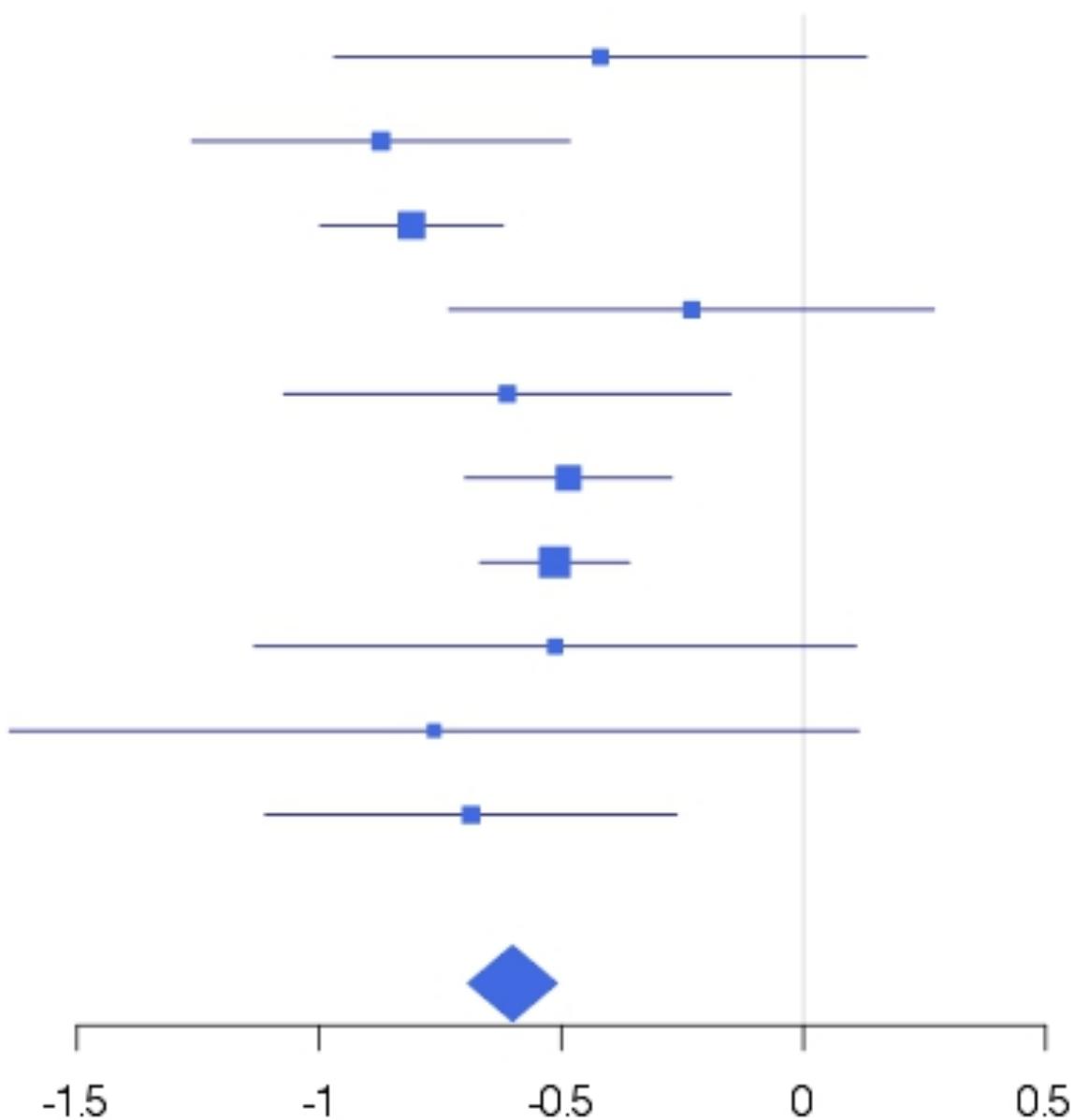
NFBC 1966

PredictCVD

PROTE

FTC

**Meta-analysis**



**SNP rs1986734, Chr 4, Pos 77420787, Crea**

Study

COROGENE



EGCUT

ERF

FR97

GenMets

HBCS

KORA

YFS

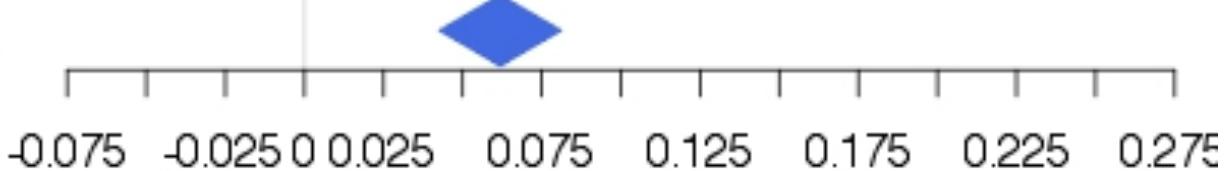
NFBC 1966

NTR

PredictCVD

PROTE

**Meta-analysis**



### **SNP rs9637599, Chr 4, Pos 89206230, Val**

Study

COROGENE

EGCUT

ERF

FR97

GenMets

HBCS

KORA

YFS

LLS

NFBC 1966

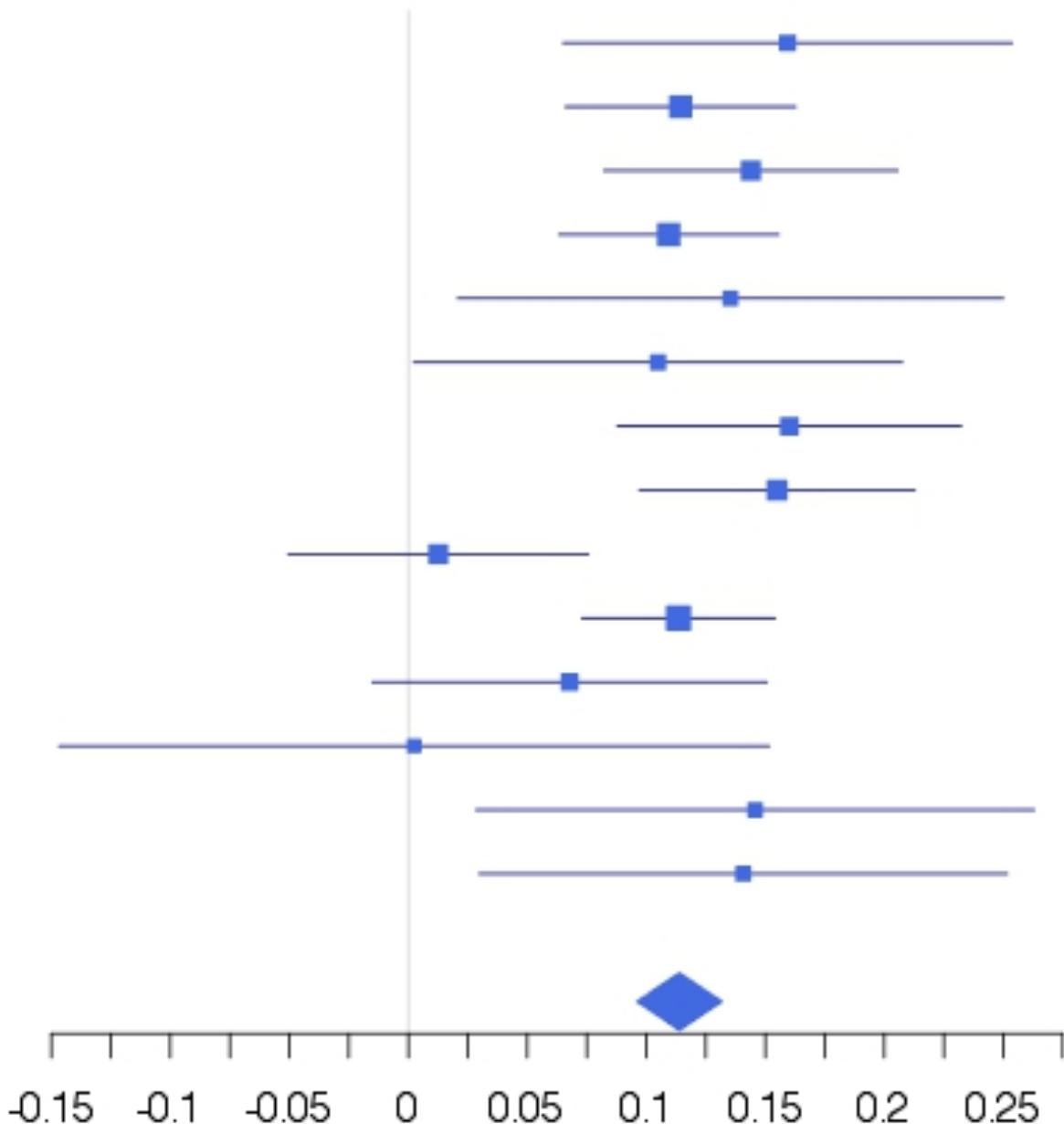
NTR

PredictCVD

PROTE

FTC

**Meta-analysis**



### **SNP rs3733402, Chr 4, Pos 187158034, His**

Study

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

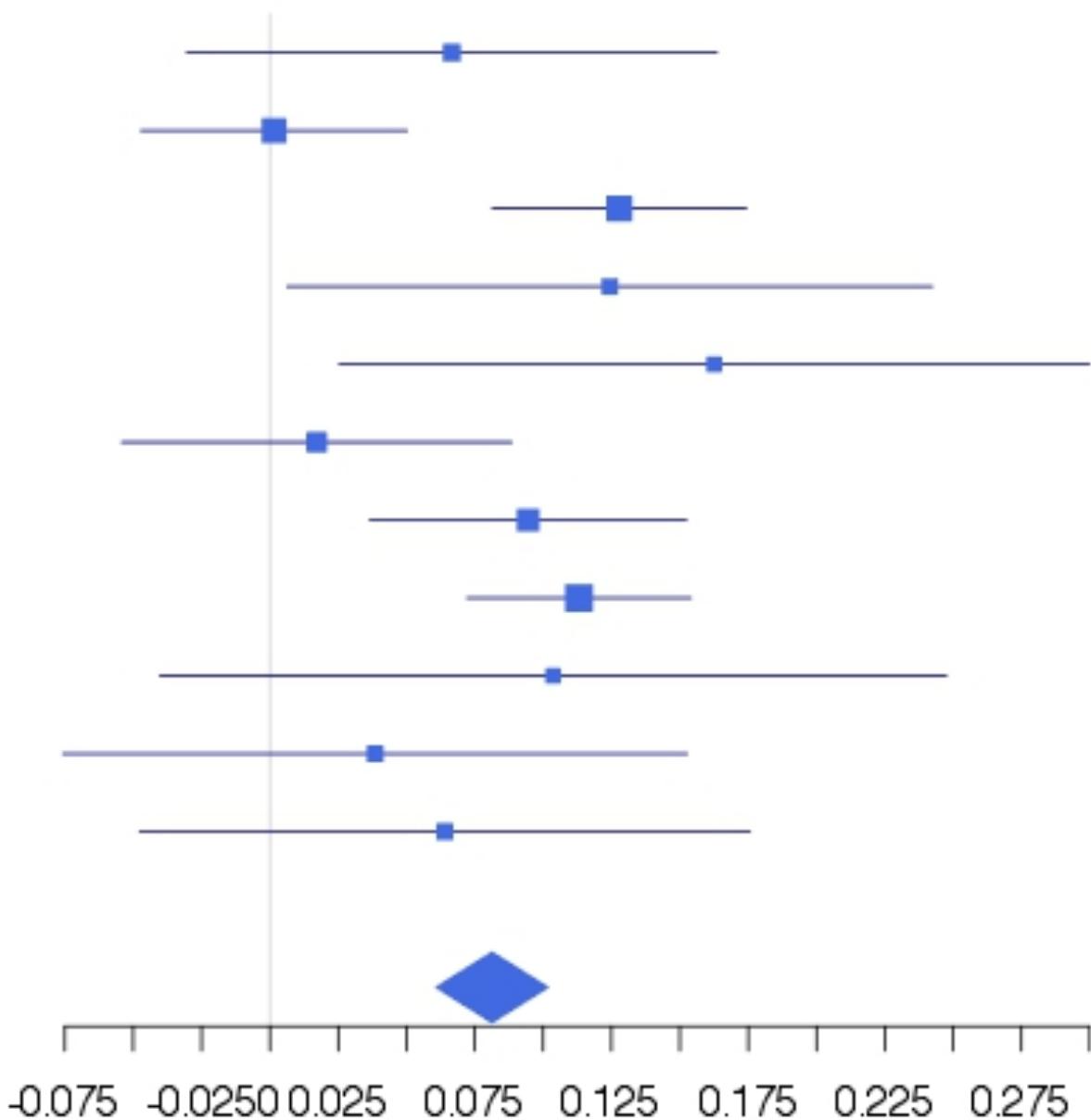
NFBC 1966

PredictCVD

PROTE

FTC

**Meta-analysis**



### SNP rs2921604, Chr 5, Pos 14867948, Cit

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



PredictCVD



PROTE



FTC



Meta-analysis



-0.025 0.025 0.075 0.125 0.175 0.225 0.275 0.325

### **SNP rs4703667, Chr 5, Pos 74613906, S.LDLC**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



-0.075 -0.025 0 0.025 0.075 0.125 0.175 0.225 0.275

## **SNP rs2731672, Chr 5, Pos 176842474, Phe**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



NTR



PredictCVD



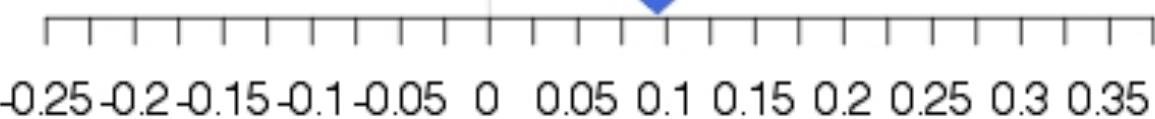
PROTE



FTC



**Meta-analysis**



**SNP rs14399, Chr 6, Pos 111543944, Tyr**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



PredictCVD



PROTE



FTC



**Meta-analysis**



**SNP rs10455872, Chr 6, Pos 161010118, VLDL.D**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



**SNP rs73066442, Chr 7, Pos 21592973, L.LDL.PL**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



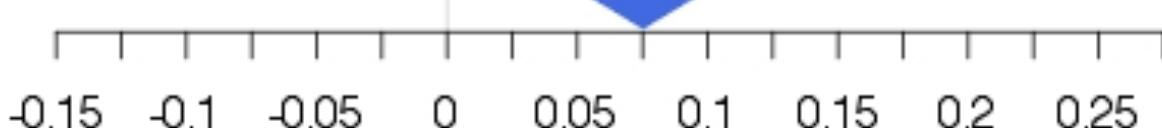
PROTE



FTC



**Meta-analysis**



### **SNP rs878521, Chr 7, Pos 44255643, Glc**

Study

COROGENE

EGCUT

ERF

GenMets

HBCS

KORA

YFS

LLS

NFBC 1966

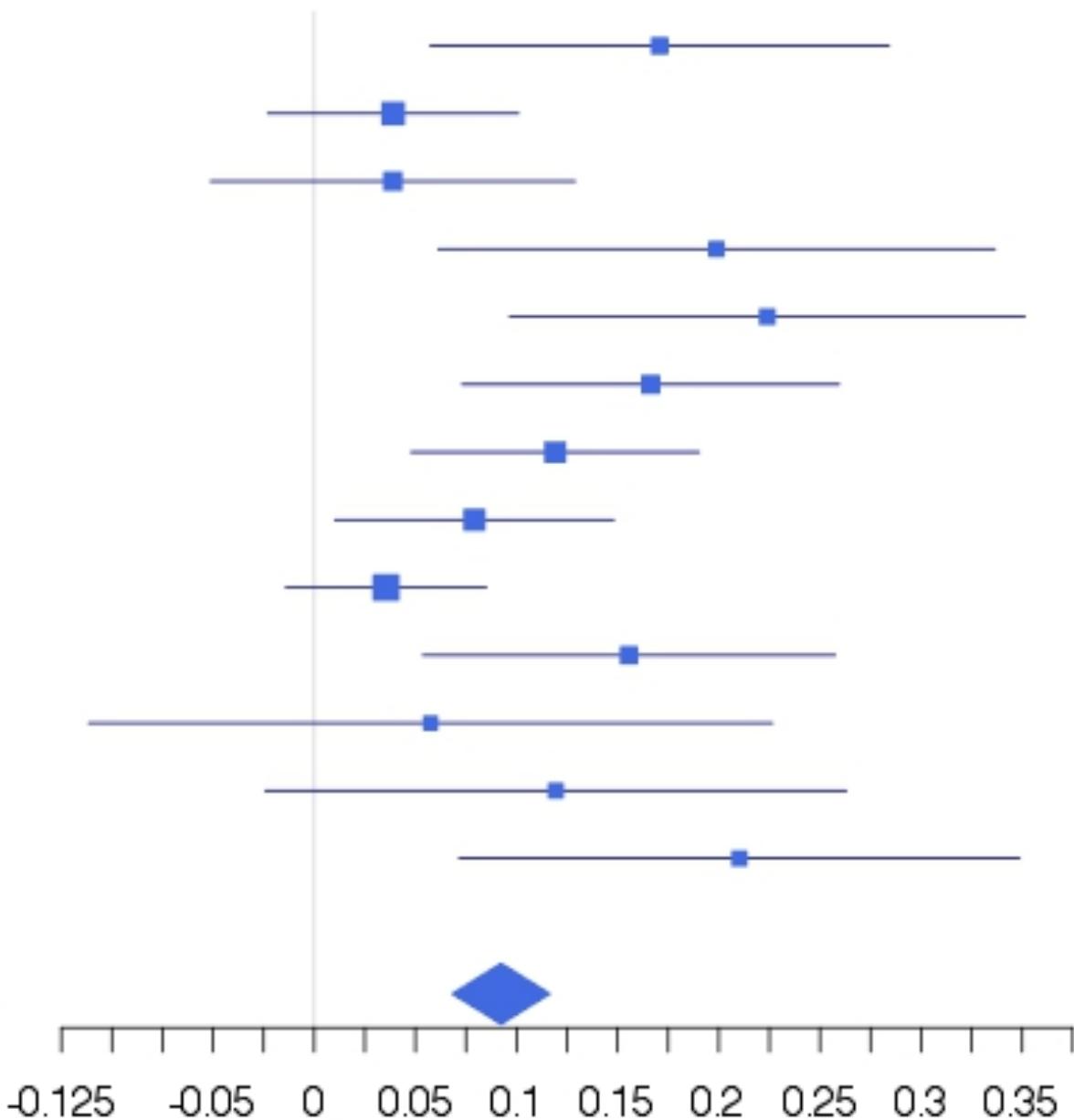
NTR

PredictCVD

PROTE

FTC

**Meta-analysis**



**SNP rs17145750, Chr 7, Pos 73026378, VLDL.D**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



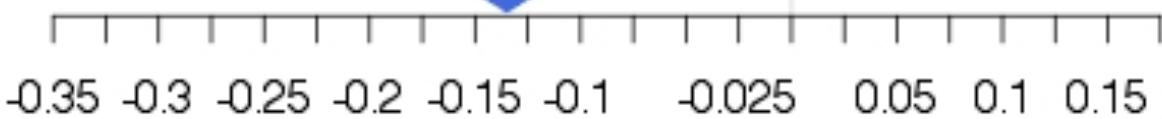
PROTE



FTC



**Meta-analysis**



### **SNP rs10265221, Chr 7, Pos 151414329, Crea**

Study

COROGENE



EGCUT

ERF

FR97

GenMets

HBCS

KORA

YFS

LLS

NFBC 1966

NTR

PredictCVD

PROTE

FTC

**Meta-analysis**

-0.125 -0.075 -0.025 0.025 0.075 0.125 0.175 0.225 0.275

### **SNP rs2169387, Chr 8, Pos 9181395, Gly**

Study

COROGENE



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



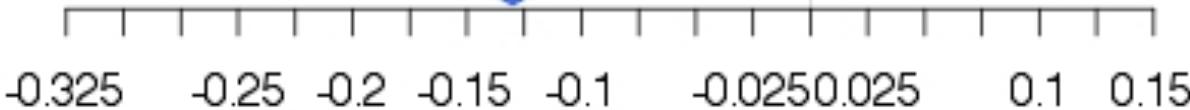
PredictCVD



FTC



**Meta-analysis**



**SNP rs115849089, Chr 8, Pos 19912370, S.VLDL.TG**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



-0.5

0

## **SNP rs2954029, Chr 8, Pos 126490972, XS.VLDL.TG**

### **Study**

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

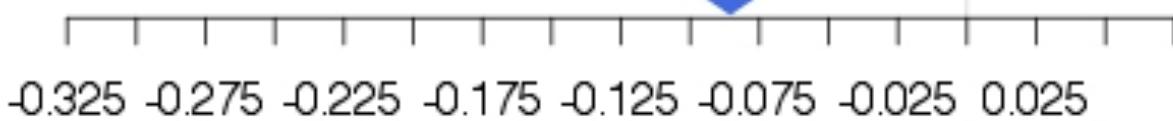
NFBC 1966

PredictCVD

PROTE

FTC

### **Meta-analysis**



### **SNP chr9:5877295:I, Chr 9, Pos 5877295, Gly**

Study

COROGENE



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



PredictCVD



FTC



**Meta-analysis**



0

0.5

1

**SNP rs686030, Chr 9, Pos 15304782, XL.HDL.FC**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



### **SNP rs2575876, Chr 9, Pos 107665739, XL.HDL.C**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



**SNP chr9:136149709:D, Chr 9, Pos 136149709, IDL.L**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



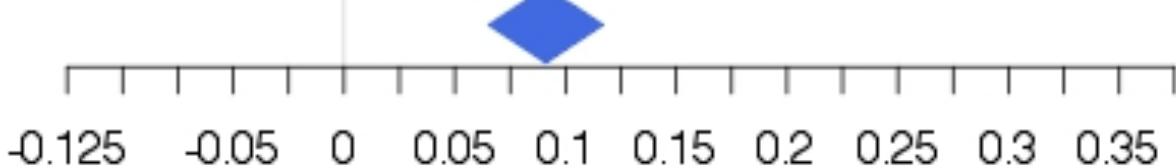
PROTE



FTC



**Meta-analysis**



### SNP rs7078003, Chr 10, Pos 99359412, Gln

Study

COROGENE

EGCUT

ERF

FR97

GenMets

HBCS

KORA

YFS

LLS

NFBC 1966

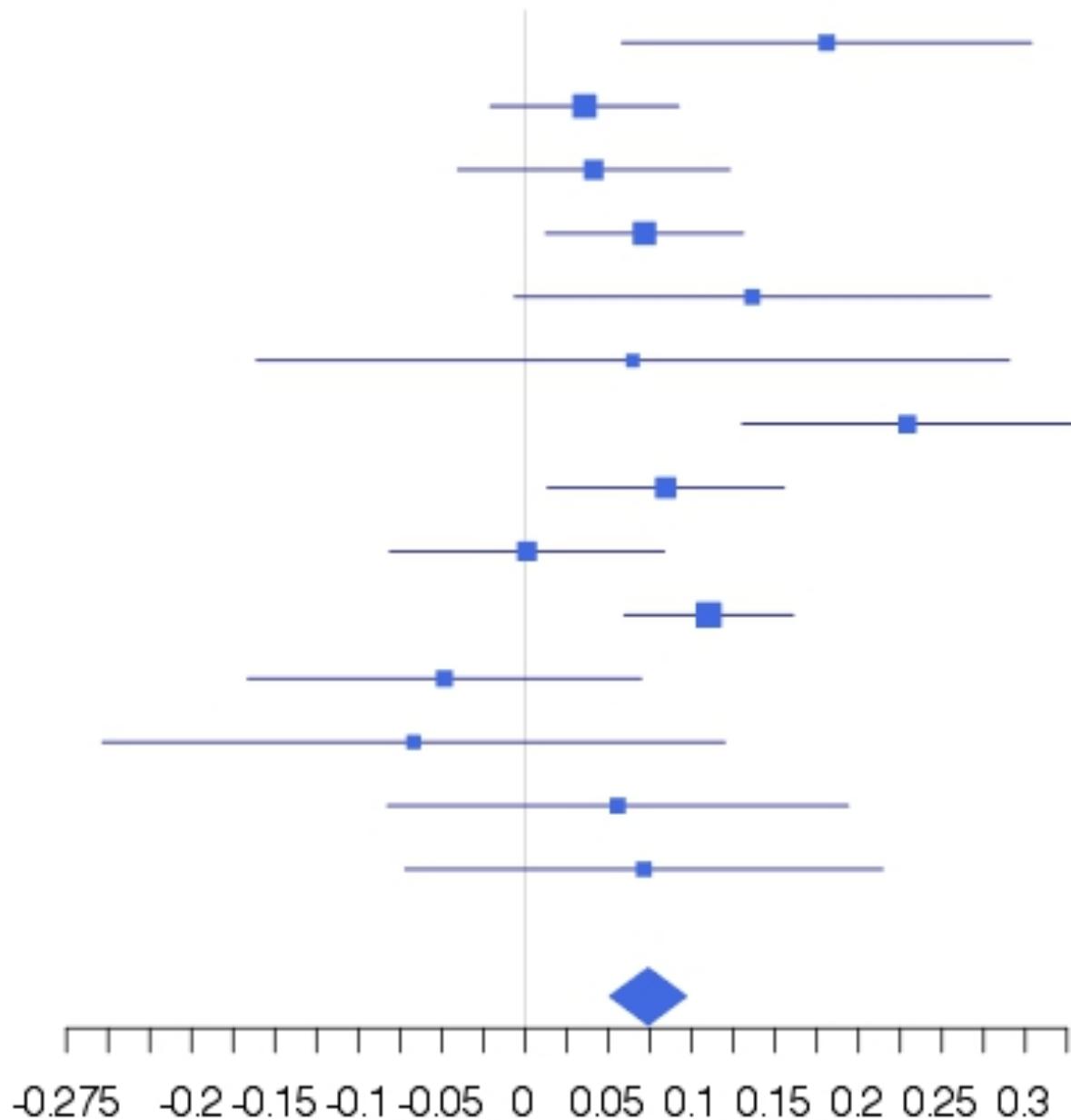
NTR

PredictCVD

PROTE

FTC

Meta-analysis



### **SNP rs174547, Chr 11, Pos 61570783, otPUFA**

#### **Study**

COROGENE



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



#### **Meta-analysis**



-0.675 -0.6 -0.525 -0.45 -0.375 -0.3 -0.225 -0.15 -0.075 0

**SNP rs186183604, Chr 11, Pos 67128733, otPUFA**

Study

COROGENE



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



### SNP rs10466351, Chr 11, Pos 92697981, Glc

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



PredictCVD



PROTE



FTC



Meta-analysis



-0.15 -0.1 -0.05 0 0.05 0.1 0.15 0.2 0.25 0.3 0.35

**SNP rs964184, Chr 11, Pos 116648917, S.VLDL.TG**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



-0.5

0

### **SNP rs4554975, Chr 12, Pos 47201814, Ala**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



PredictCVD



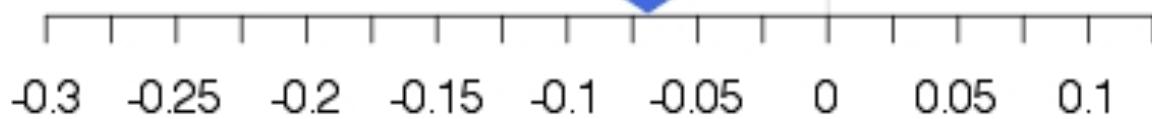
PROTE



FTC



**Meta-analysis**



### **SNP rs2657879, Chr 12, Pos 56865338, Gln**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



PredictCVD



PROTE



FTC



**Meta-analysis**



-0.475 -0.4 -0.35 -0.3 -0.25 -0.2 -0.15 -0.1 -0.05 0 0.05 0.1

### **SNP rs7954638, Chr 12, Pos 96314795, His**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



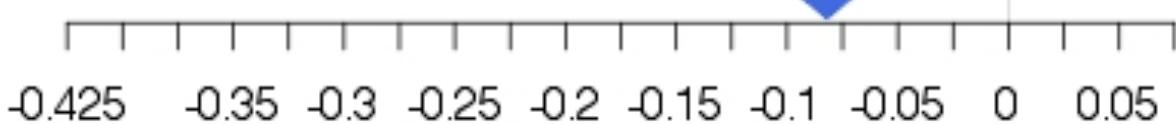
PROTE



FTC



**Meta-analysis**



### **SNP rs1718309, Chr 12, Pos 103242396, Phe**

Study

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

LLS

NFBC 1966

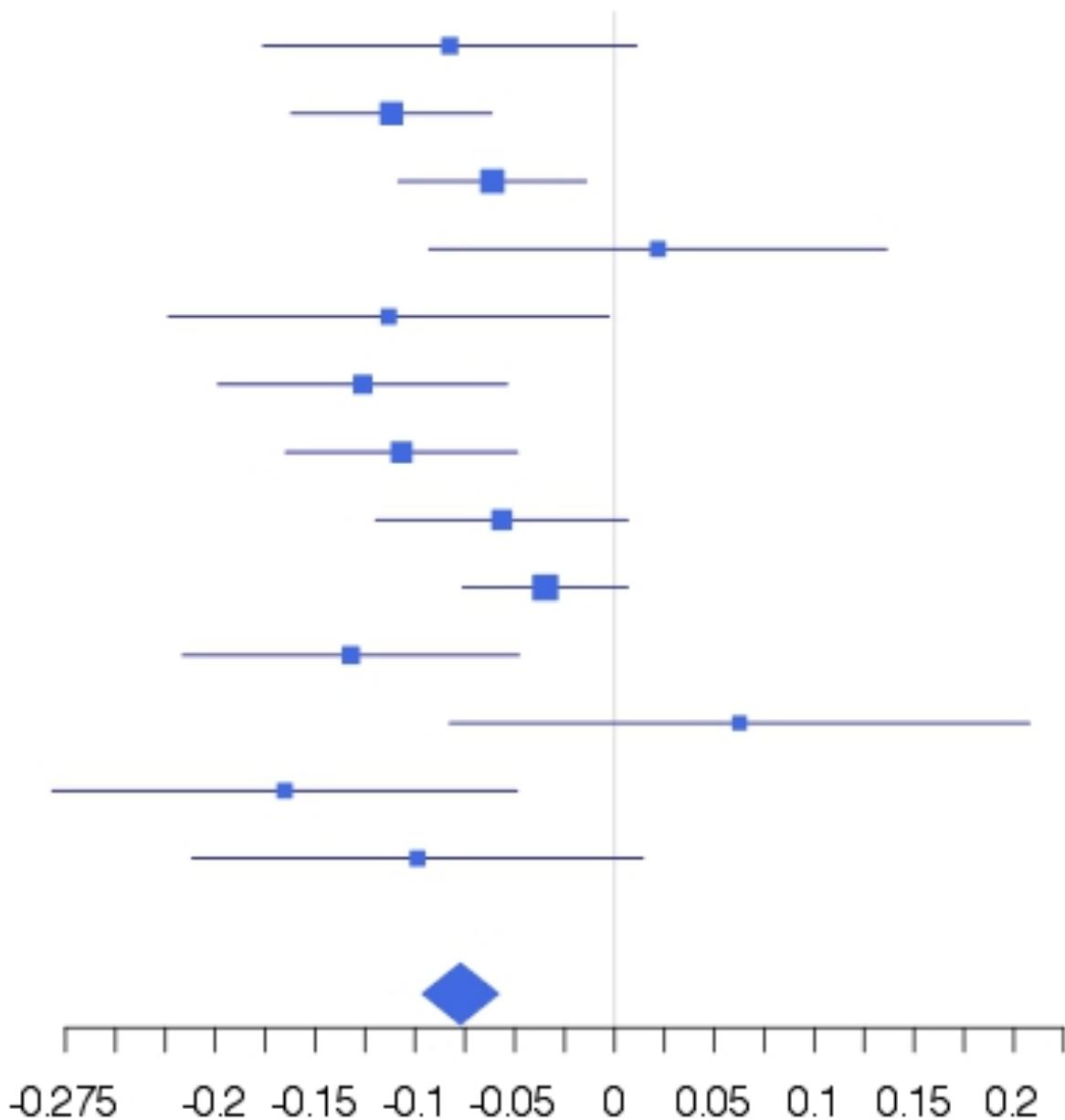
NTR

PredictCVD

PROTE

FTC

**Meta-analysis**



**SNP rs67053123, Chr 12, Pos 125353810, HDL.D**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



-0.125 -0.05 0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5

**SNP rs1998848, Chr 14, Pos 21492229, His**

Study

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

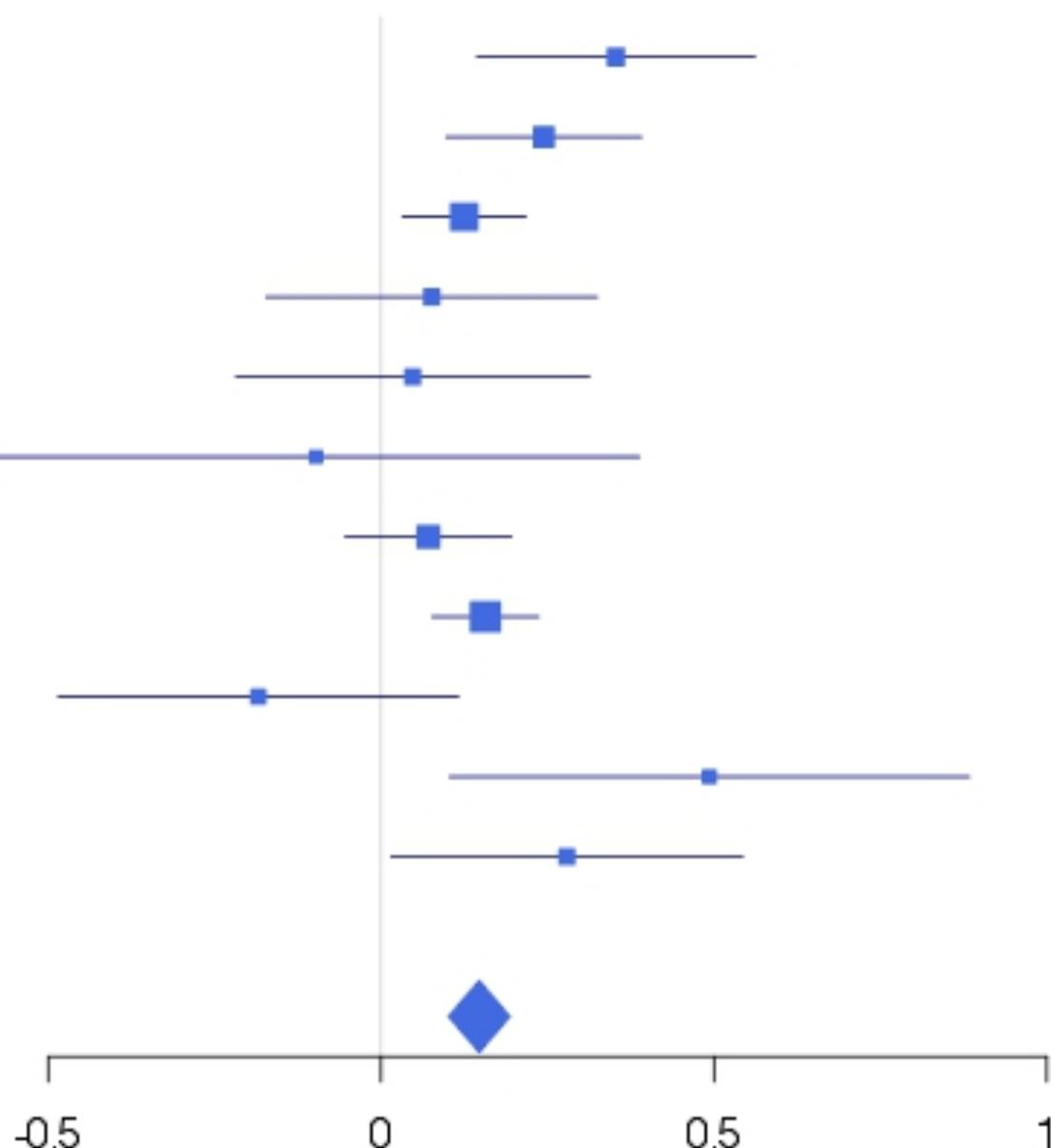
NFBC 1966

PredictCVD

PROTE

FTC

**Meta-analysis**



**SNP rs28929474, Chr 14, Pos 94844947, Gp**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



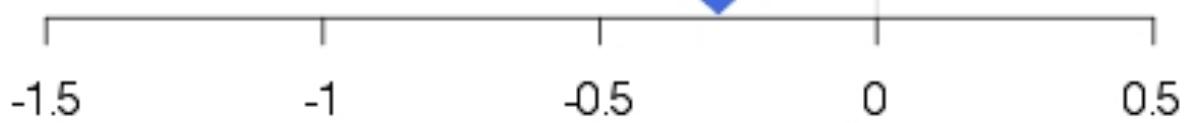
PROTE



FTC



**Meta-analysis**



### **SNP rs61524473, Chr 15, Pos 45646283, Crea**

Study

COROGENE

EGCUT

ERF

FR97

GenMets

HBCS

KORA

YFS

LLS

NFBC 1966

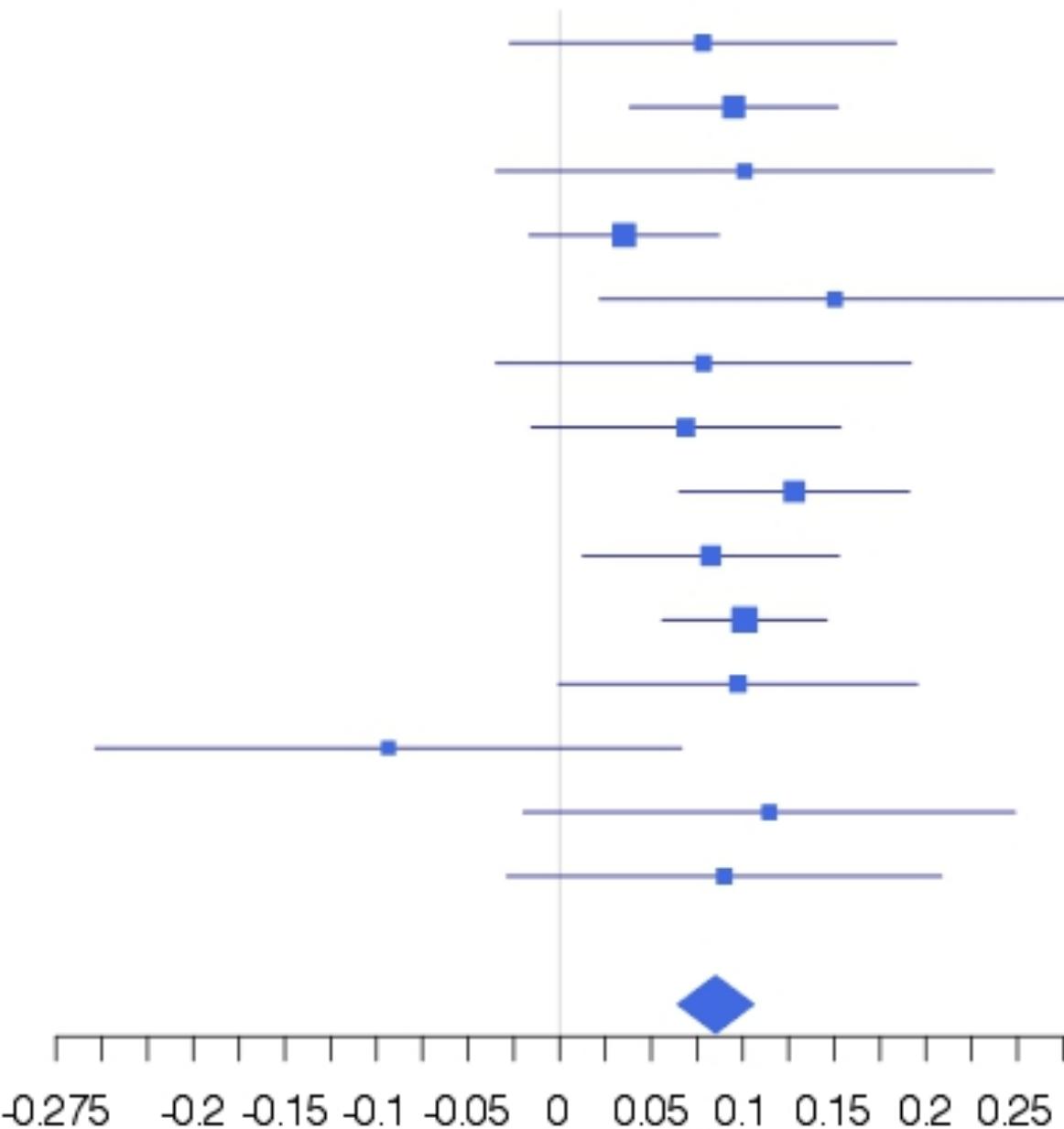
NTR

PredictCVD

PROTE

FTC

**Meta-analysis**



### **SNP rs1532085, Chr 15, Pos 58683366, XL.HDL.TG**

#### **Study**

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



#### **Meta-analysis**



-0.425 -0.375 -0.325 -0.275 -0.225 -0.175 -0.125 -0.075 -0.025

### **SNP rs11644601, Chr 16, Pos 15172118, Bis.DB.ratio**

Study

COROGENE



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



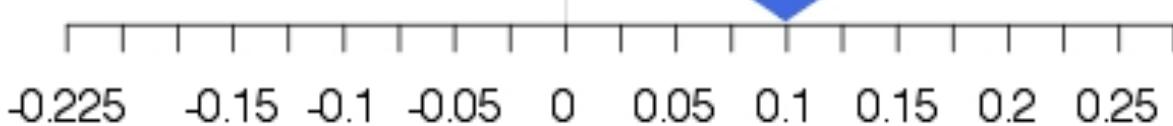
PROTE



FTC



**Meta-analysis**



### **SNP rs247617, Chr 16, Pos 56990716, HDL.C**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



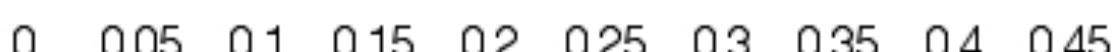
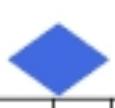
PROTE



FTC



**Meta-analysis**



## SNP rs74249229, Chr 16, Pos 69979271, Pyr

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



PredictCVD



PROTE



FTC



Meta-analysis



-0.5

0

0.5

**SNP rs77303550, Chr 16, Pos 72079657, Gp**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



-0.075 -0.025 0.025 0.075 0.125 0.175 0.225 0.275 0.325 0.375

**SNP rs10083777, Chr 16, Pos 81065282, Gly**

Study

COROGENE



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



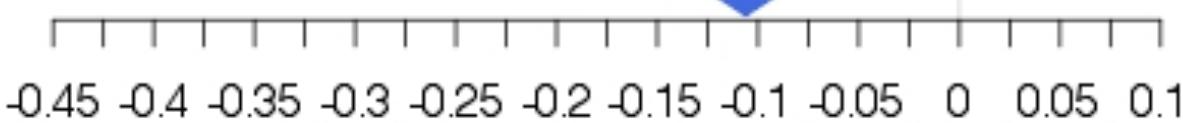
PredictCVD



FTC



**Meta-analysis**



**SNP rs181807530, Chr 17, Pos 4774814, SM**

Study

COROGENE



FR97

GenMets

HBCS

YFS

NFBC 1966

PredictCVD

FTC

**Meta-analysis**



### SNP rs172642, Chr 17, Pos 6595398, Cit

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



LLS



NFBC 1966



NTR



PredictCVD



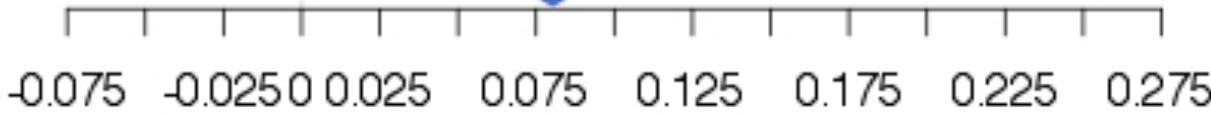
PROTE



FTC



Meta-analysis



### SNP rs2079742, Chr 17, Pos 59465697, Crea

#### Study

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

LLS

NFBC 1966

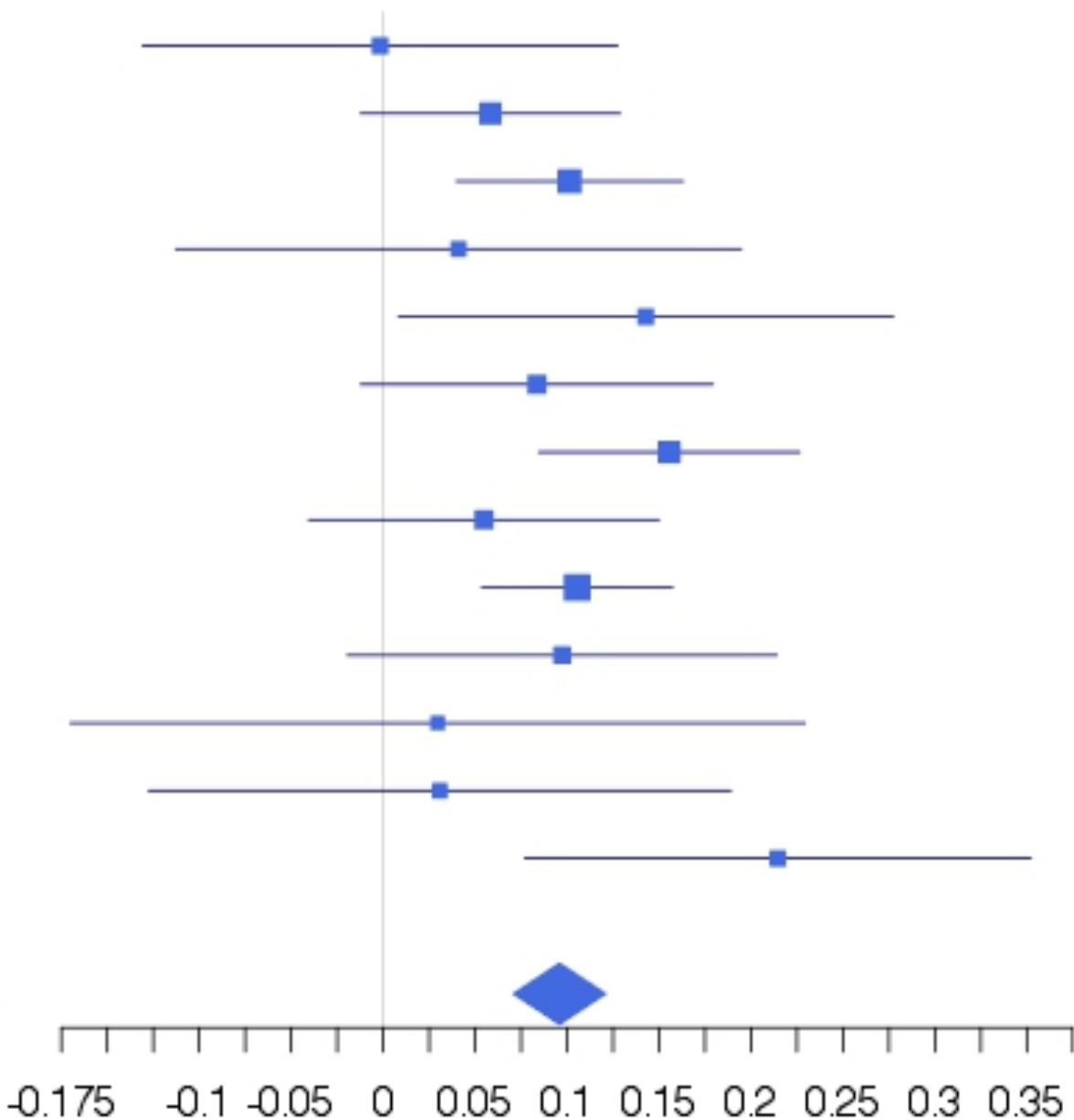
NTR

PredictCVD

PROTE

FTC

#### Meta-analysis



**SNP rs6507939, Chr 18, Pos 47176261, HDL.D**

Study

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

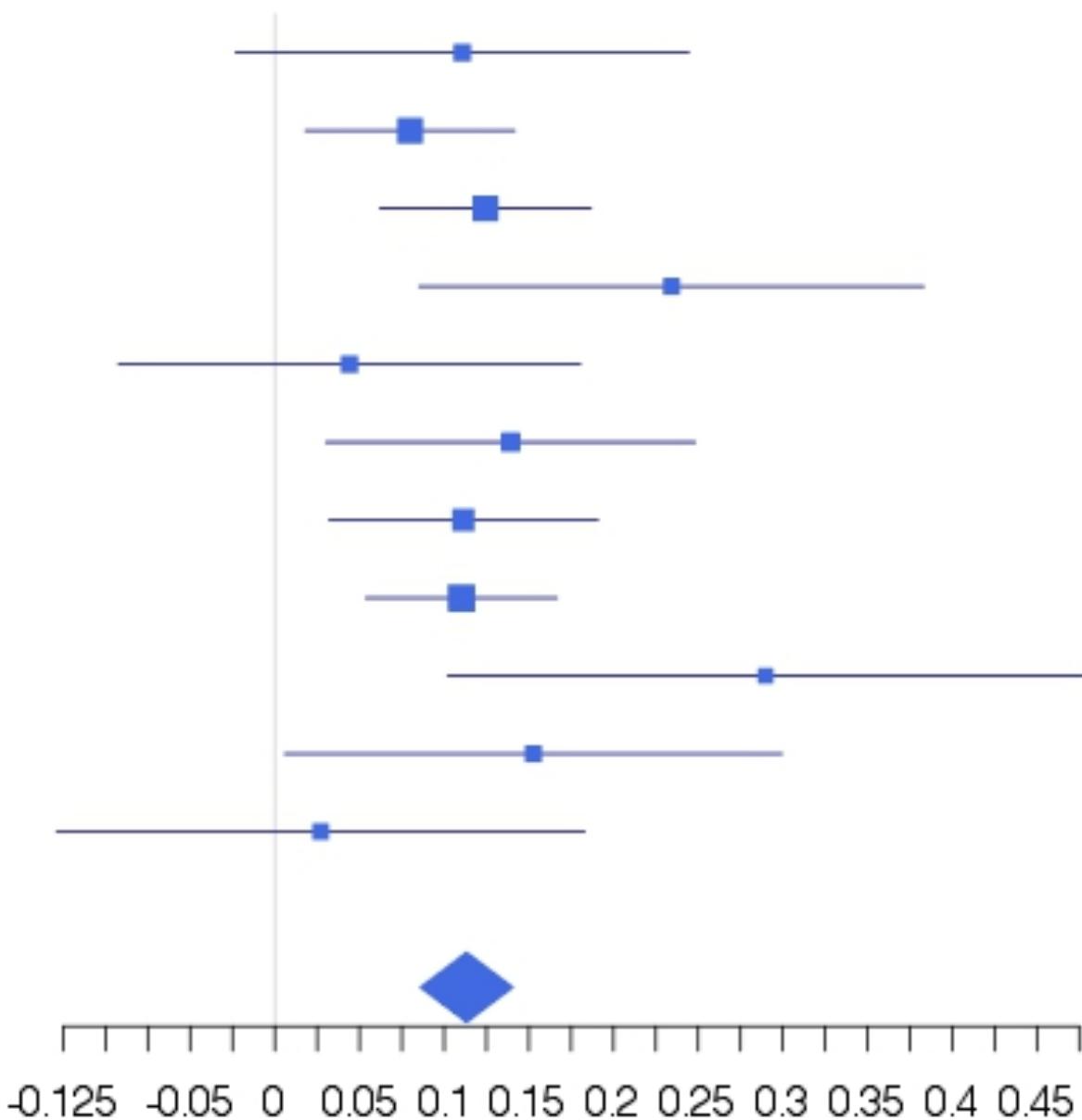
NFBC 1966

PredictCVD

PROTE

FTC

**Meta-analysis**



**SNP rs116843064, Chr 19, Pos 8429323, VLDL.D**

Study

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

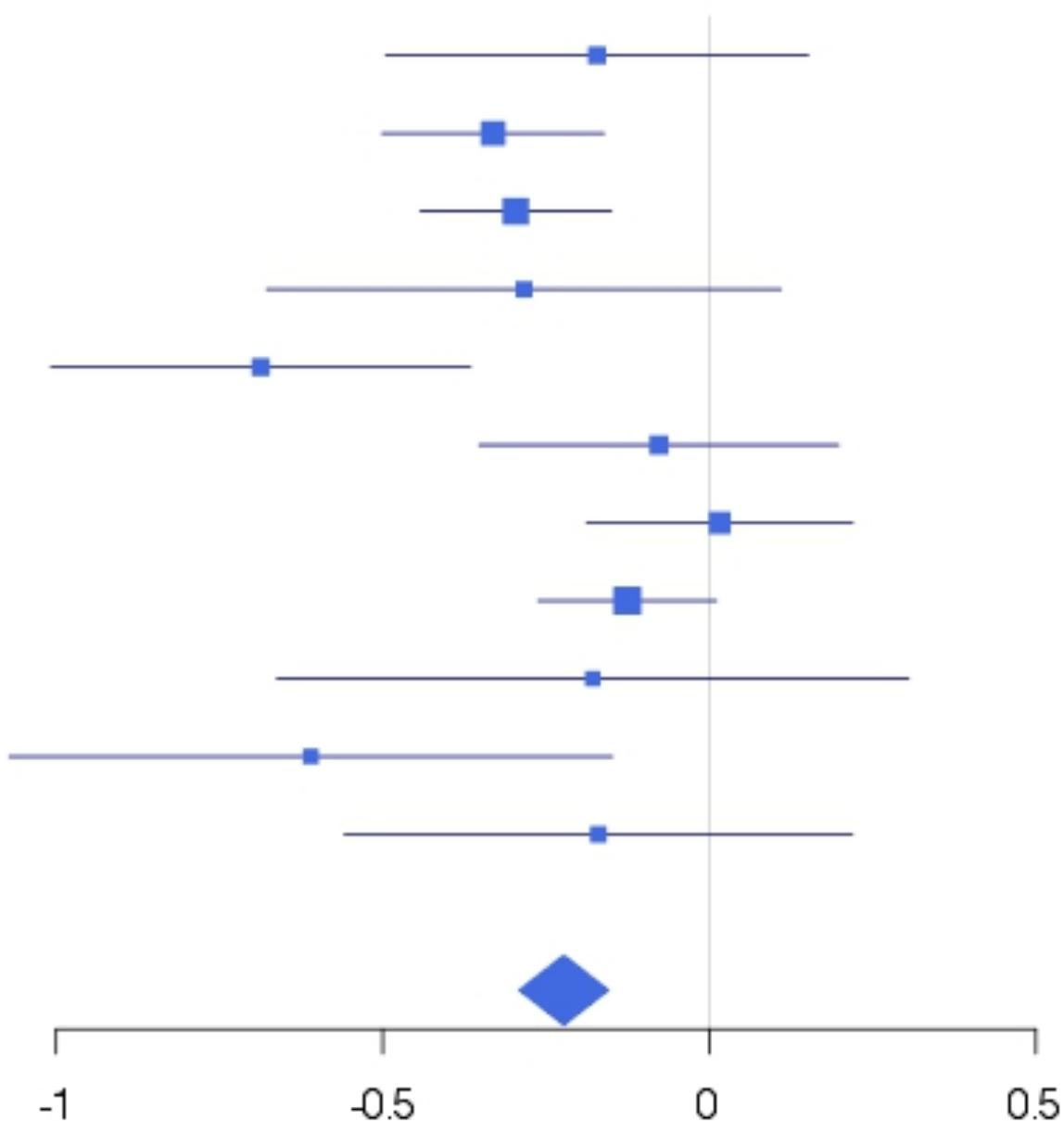
NFBC 1966

PredictCVD

PROTE

FTC

**Meta-analysis**



**SNP rs142130958, Chr 19, Pos 11190652, L.LDL.FC**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



-0.5

0

**SNP rs72999033, Chr 19, Pos 19366632, M.VLDL.C**

Study

COROGENE



EGCUT



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



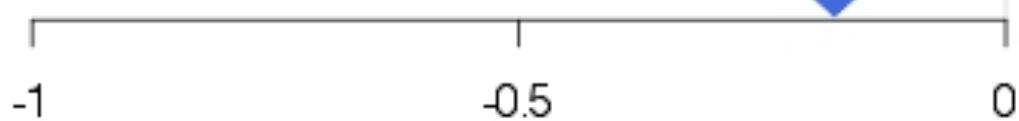
PROTE



FTC



**Meta-analysis**



### **SNP rs7412, Chr 19, Pos 45412079, L.LDL.FC**

Study

COROGENE



ERF



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



**Meta-analysis**



-0.975 -0.85 -0.75 -0.65 -0.55 -0.45 -0.35 -0.25 -0.15 0

**SNP rs1883711, Chr 20, Pos 39179822, XS.VLDL.PL**

Study

COROGENE

EGCUT

FR97

GenMets

HBCS

KORA

YFS

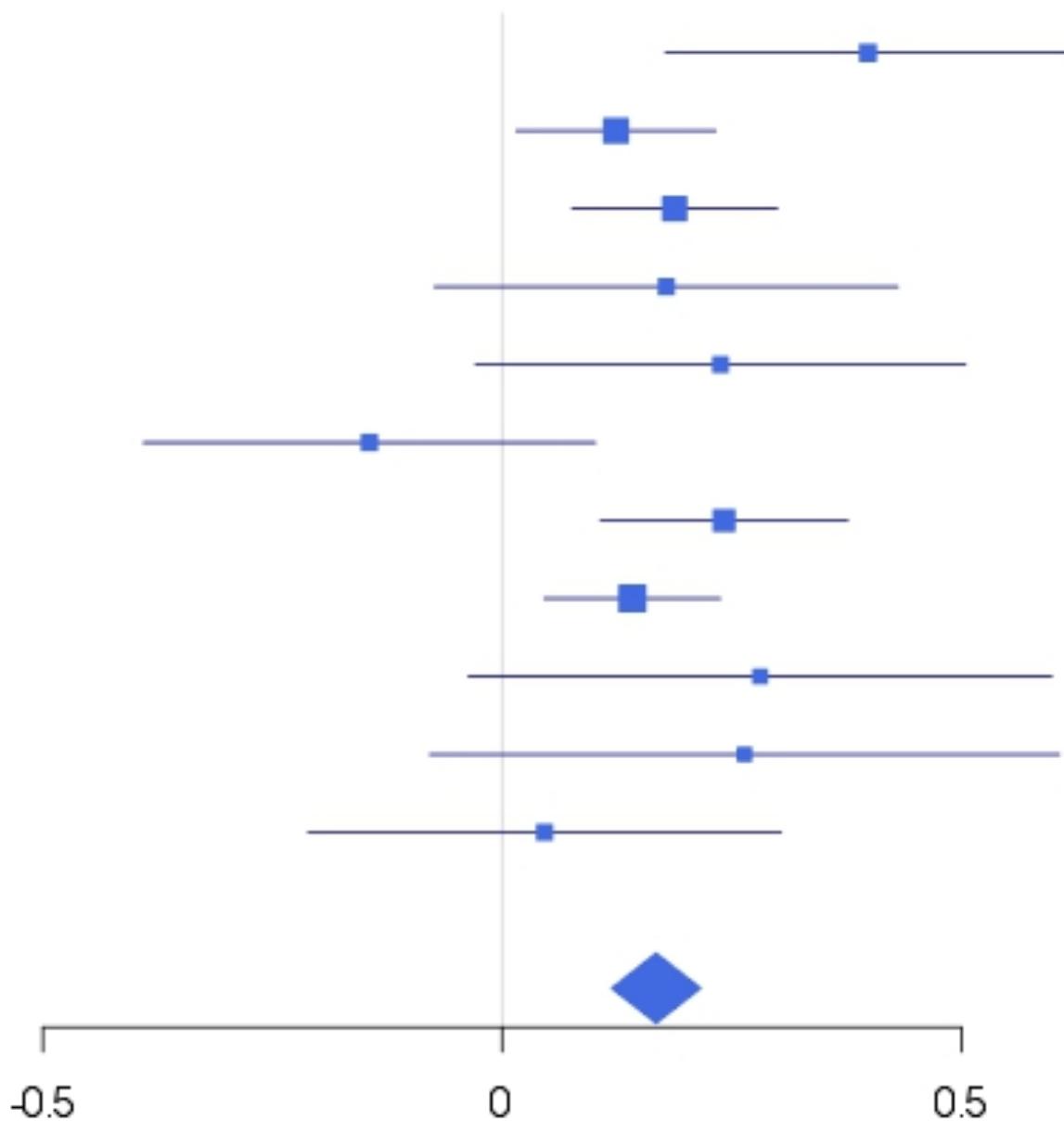
NFBC 1966

PredictCVD

PROTE

FTC

**Meta-analysis**



**SNP rs6073958, Chr 20, Pos 44551855, S.HDL.P**

Study

COROGENE



EGCUT



FR97



GenMets



HBCS



KORA



YFS



NFBC 1966



PredictCVD



PROTE



FTC



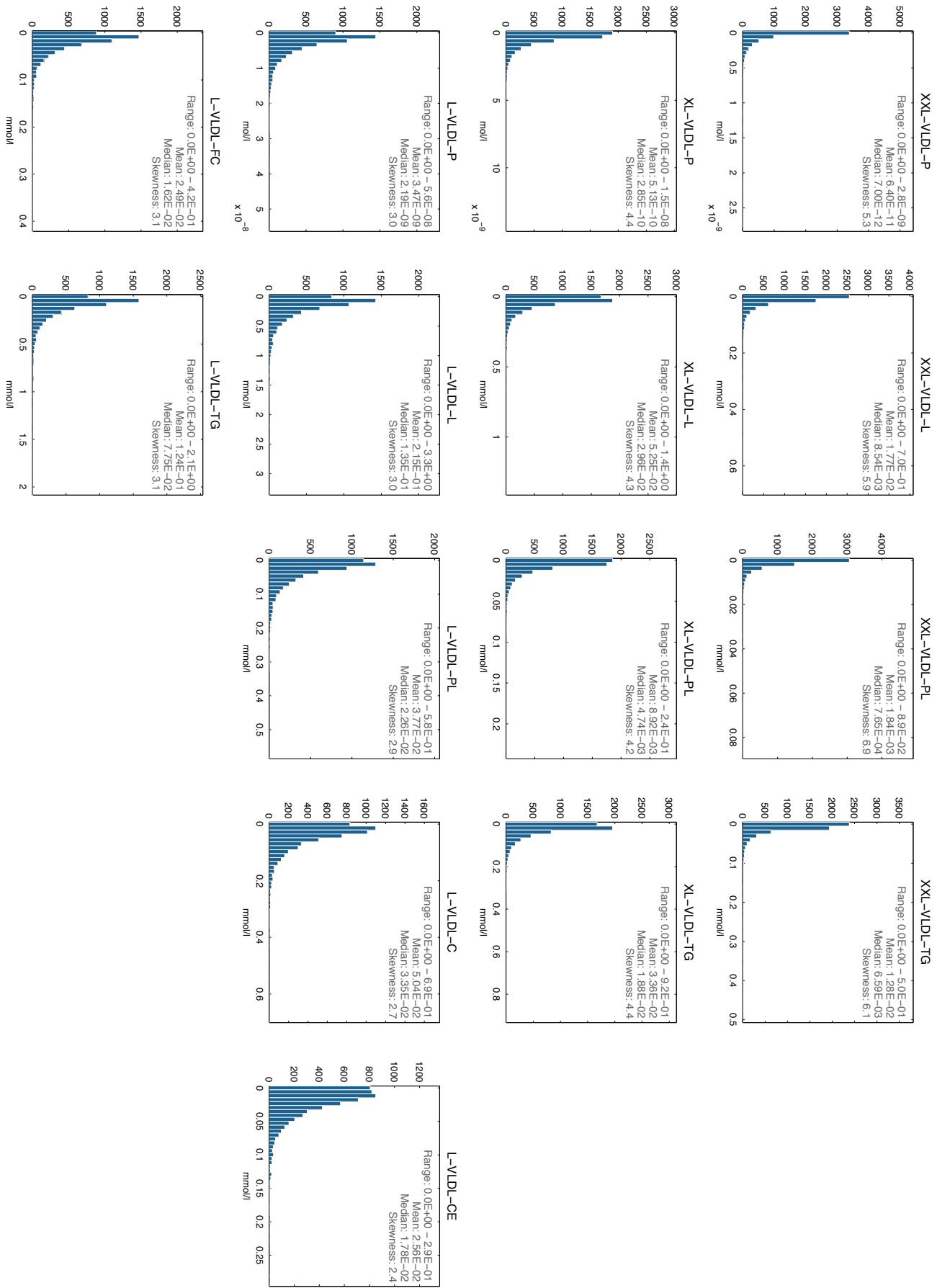
**Meta-analysis**



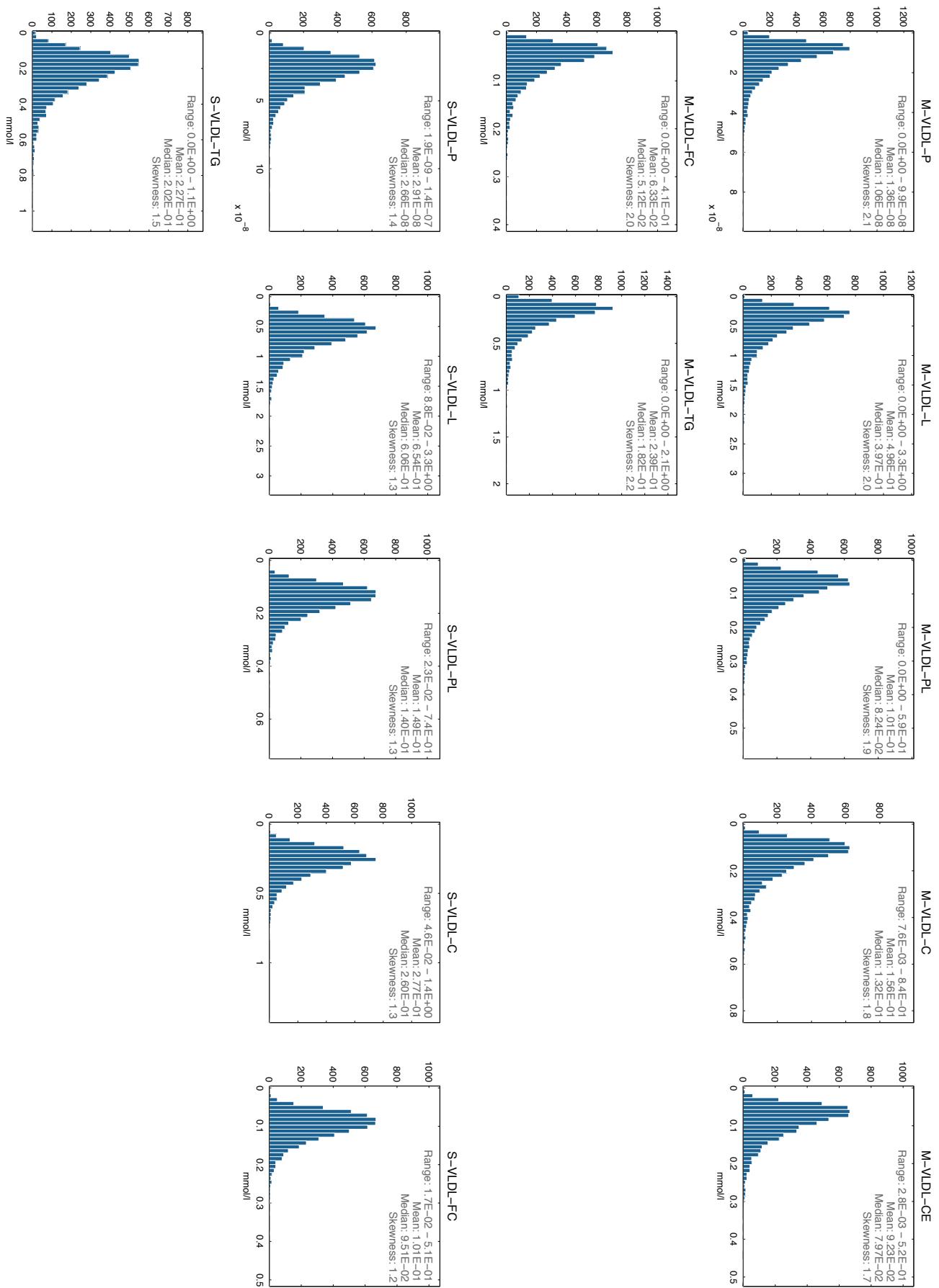
-0.025 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45

**Supplementary Figure 3.** Representative metabolite-specific original (untransformed) distributions and descriptive statistics (range, mean, median and skewness). The distributions drawn are from the largest cohort NFBC1966 with n = 5923.

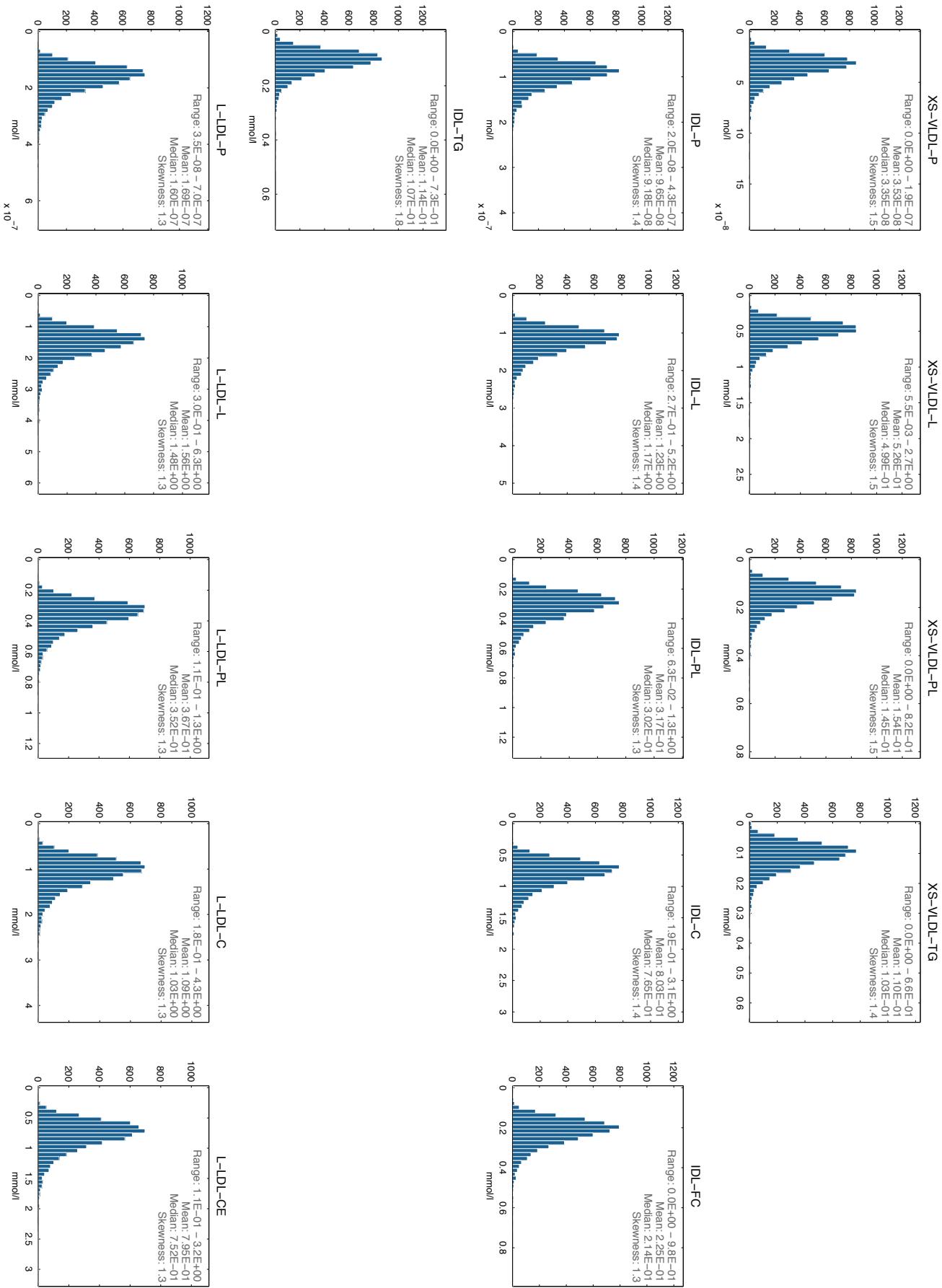
## NFC66 measures summary



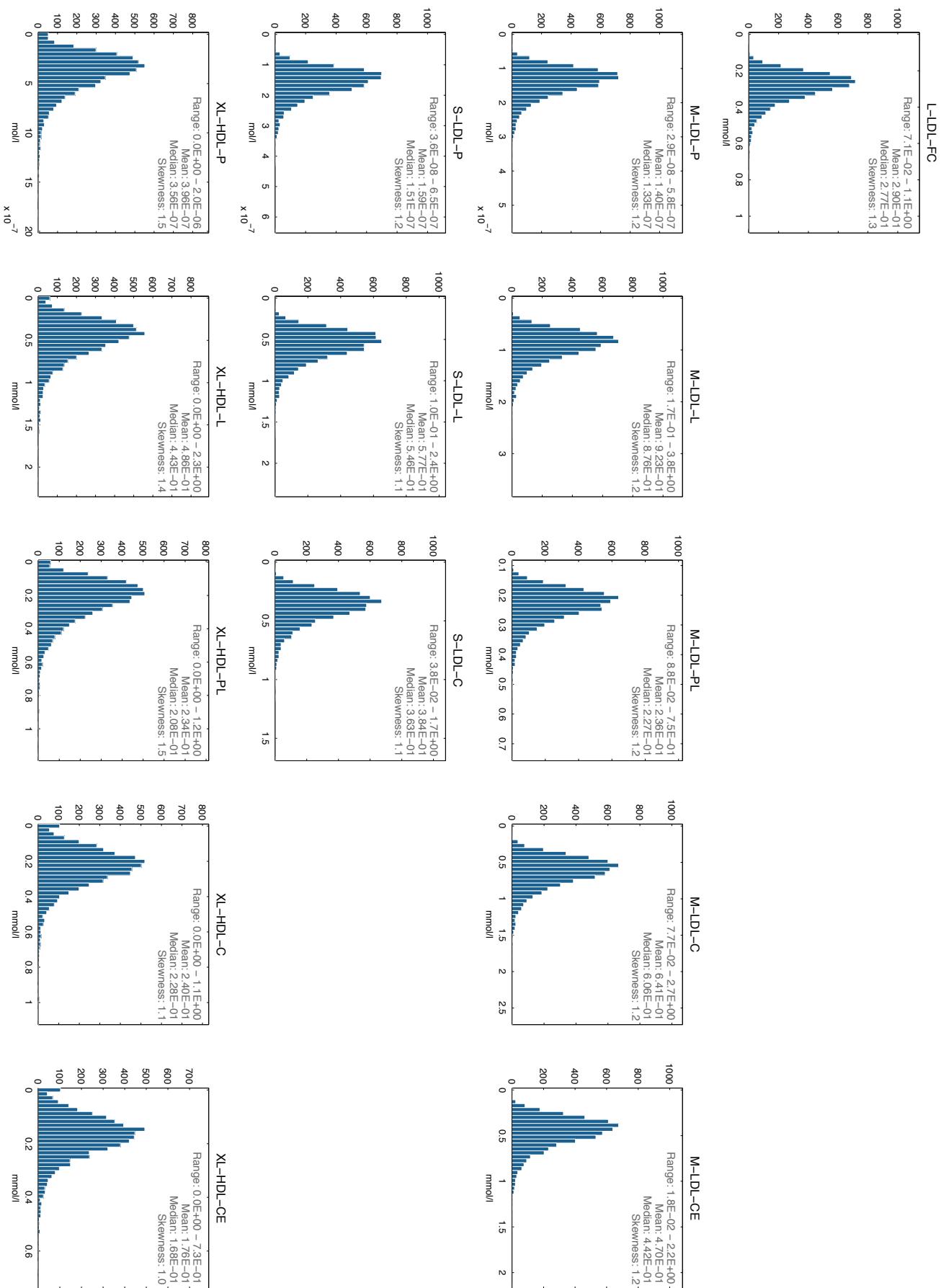
## NFBG66 measures summary



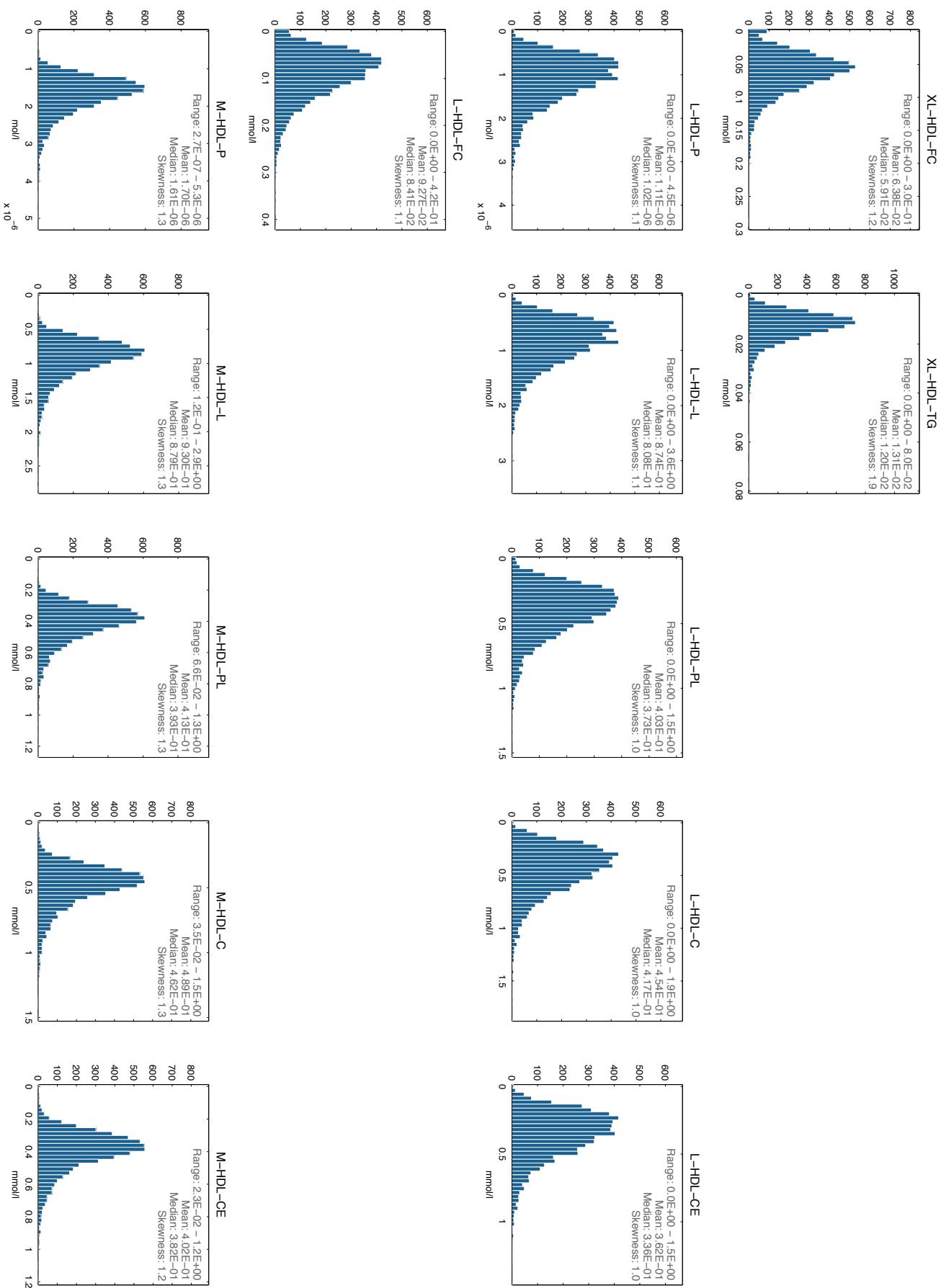
## NFC66 measures summary



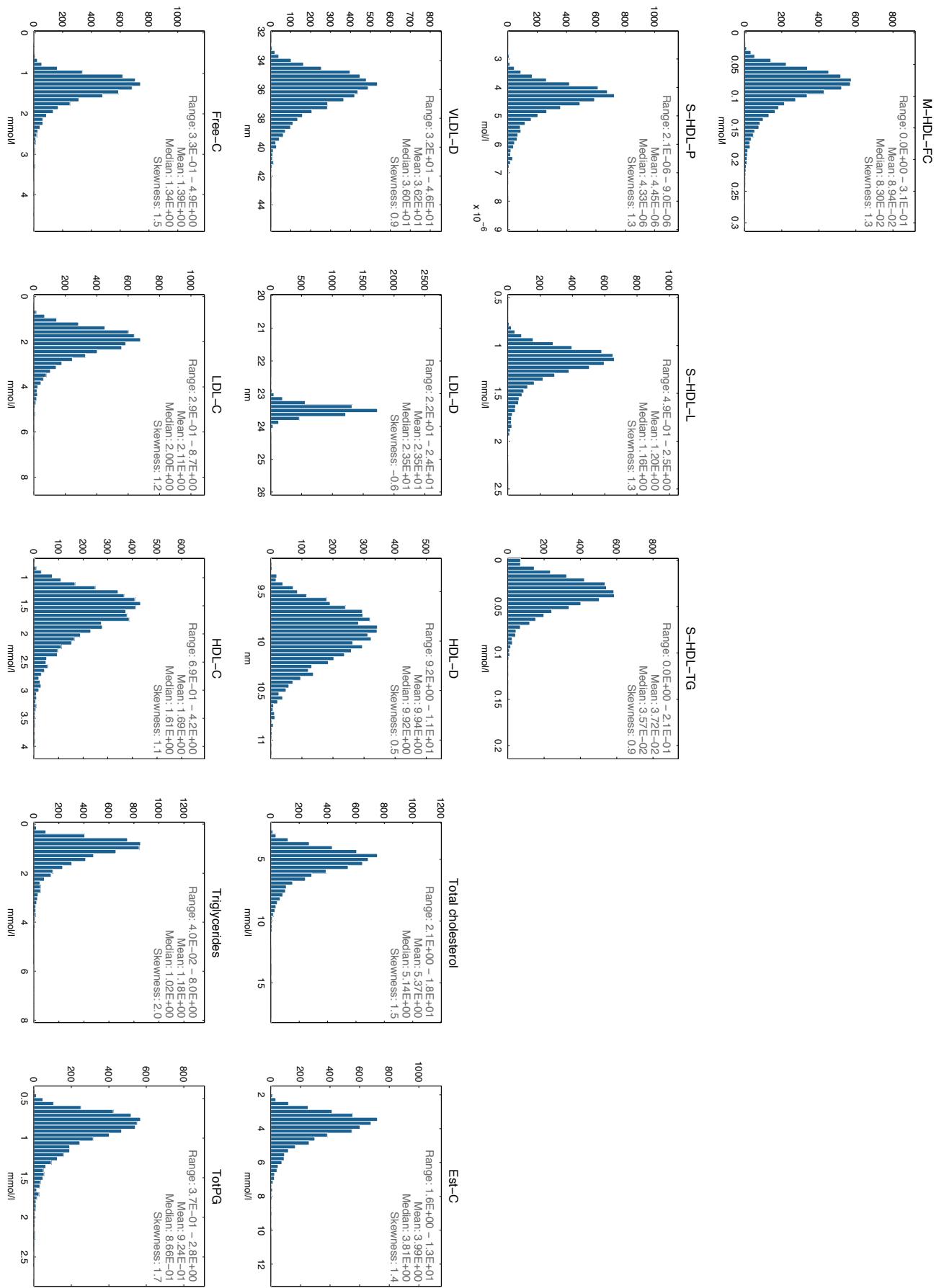
## NFBG66 measures summary



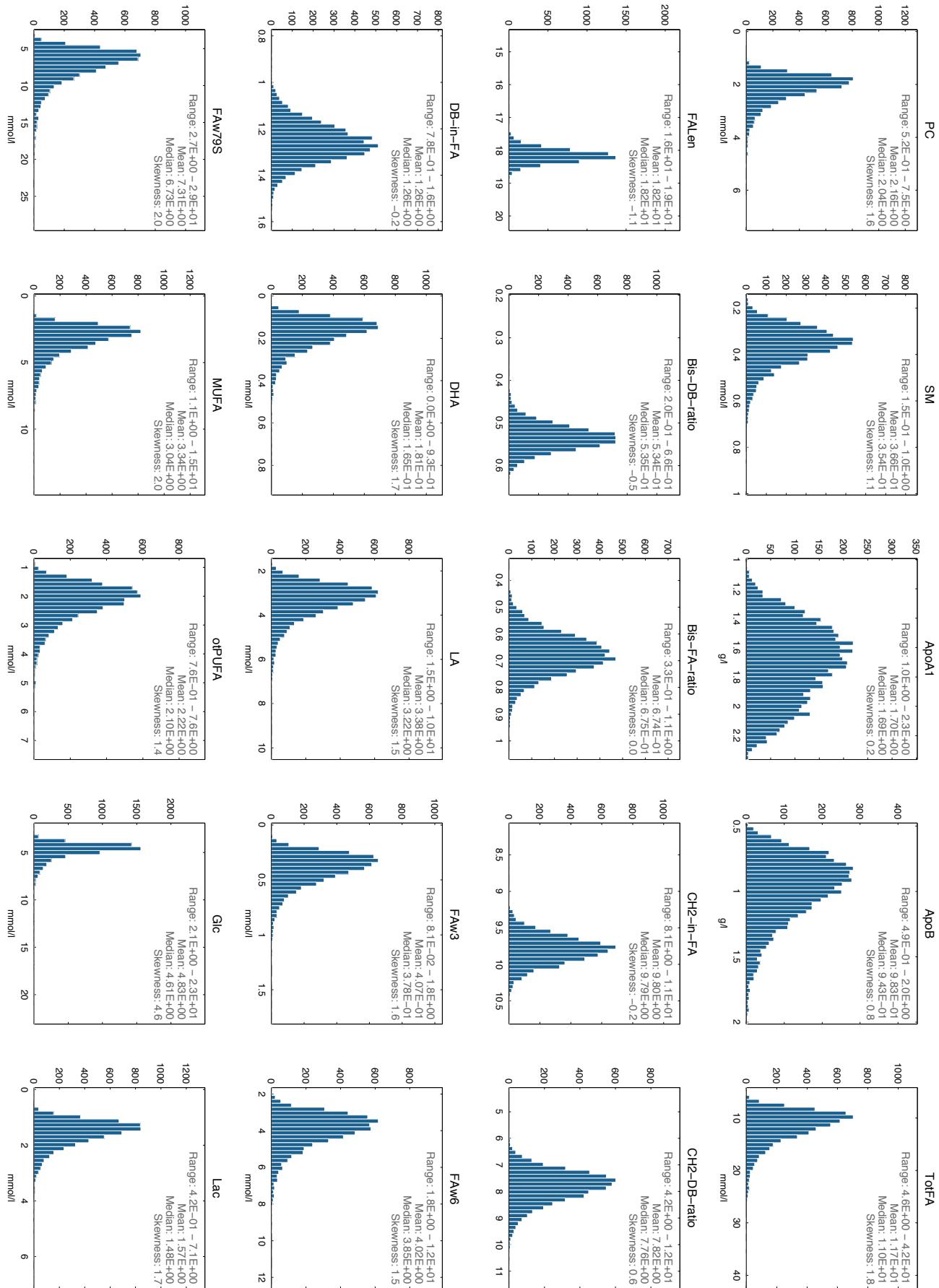
## NFBG66 measures summary



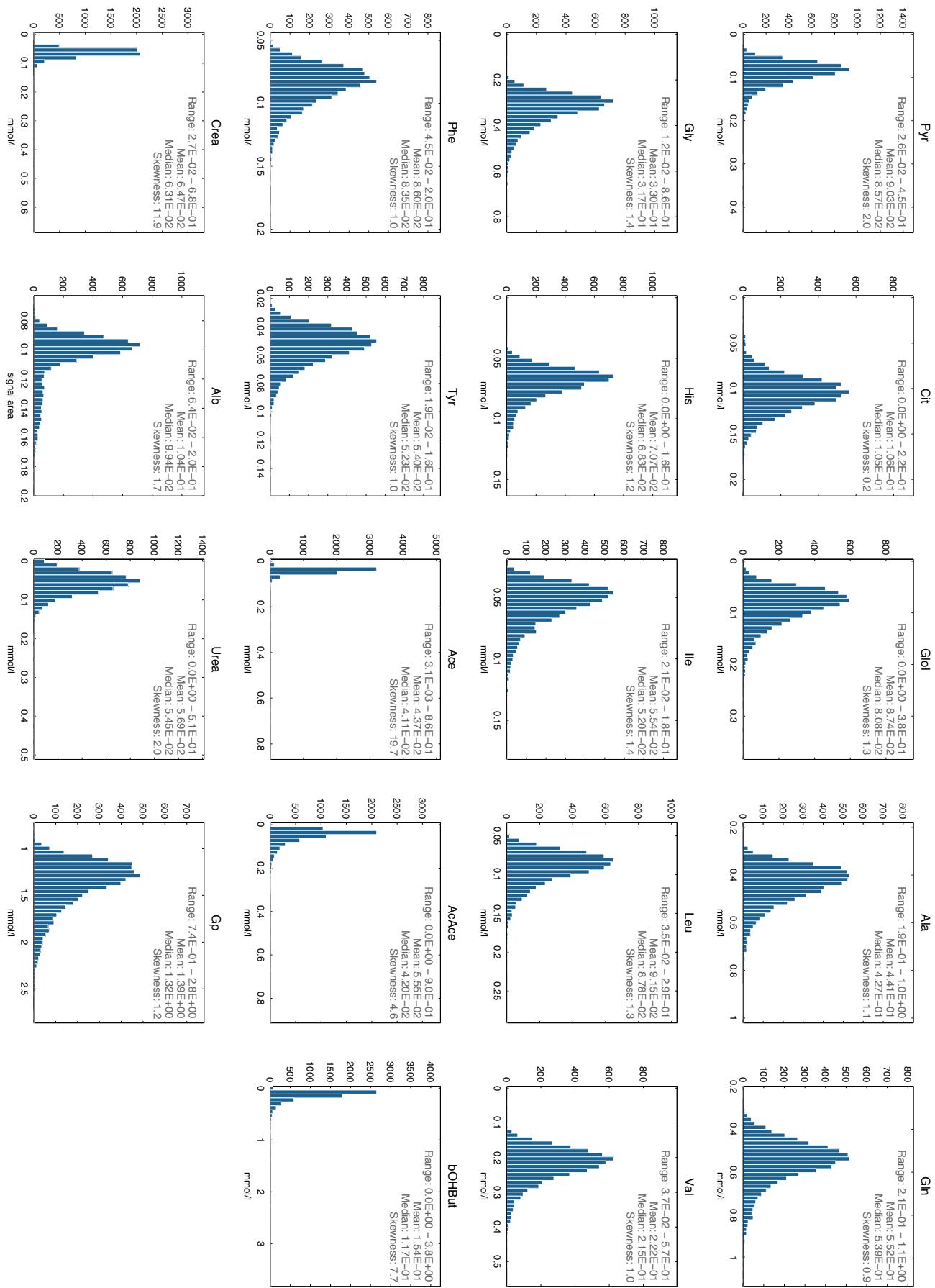
## NFBC66 measures summary



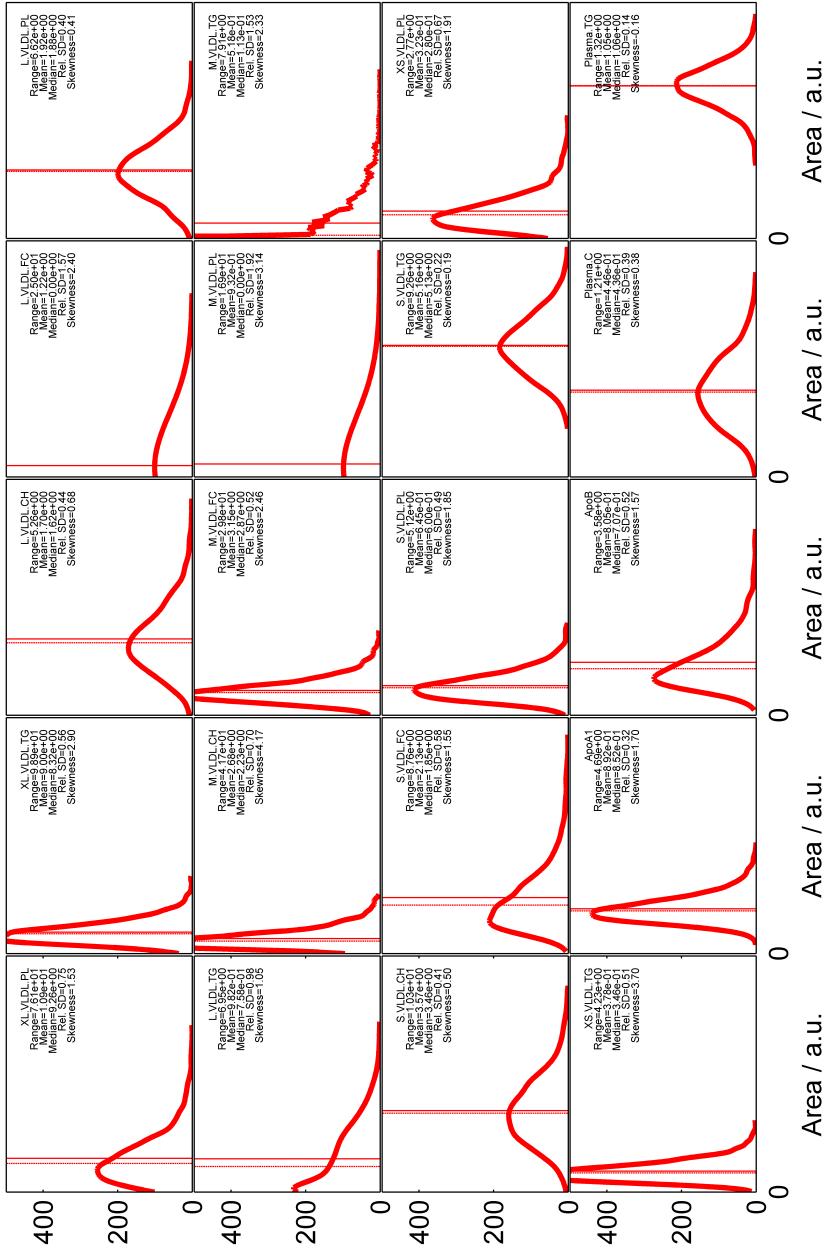
## NFC66 measures summary

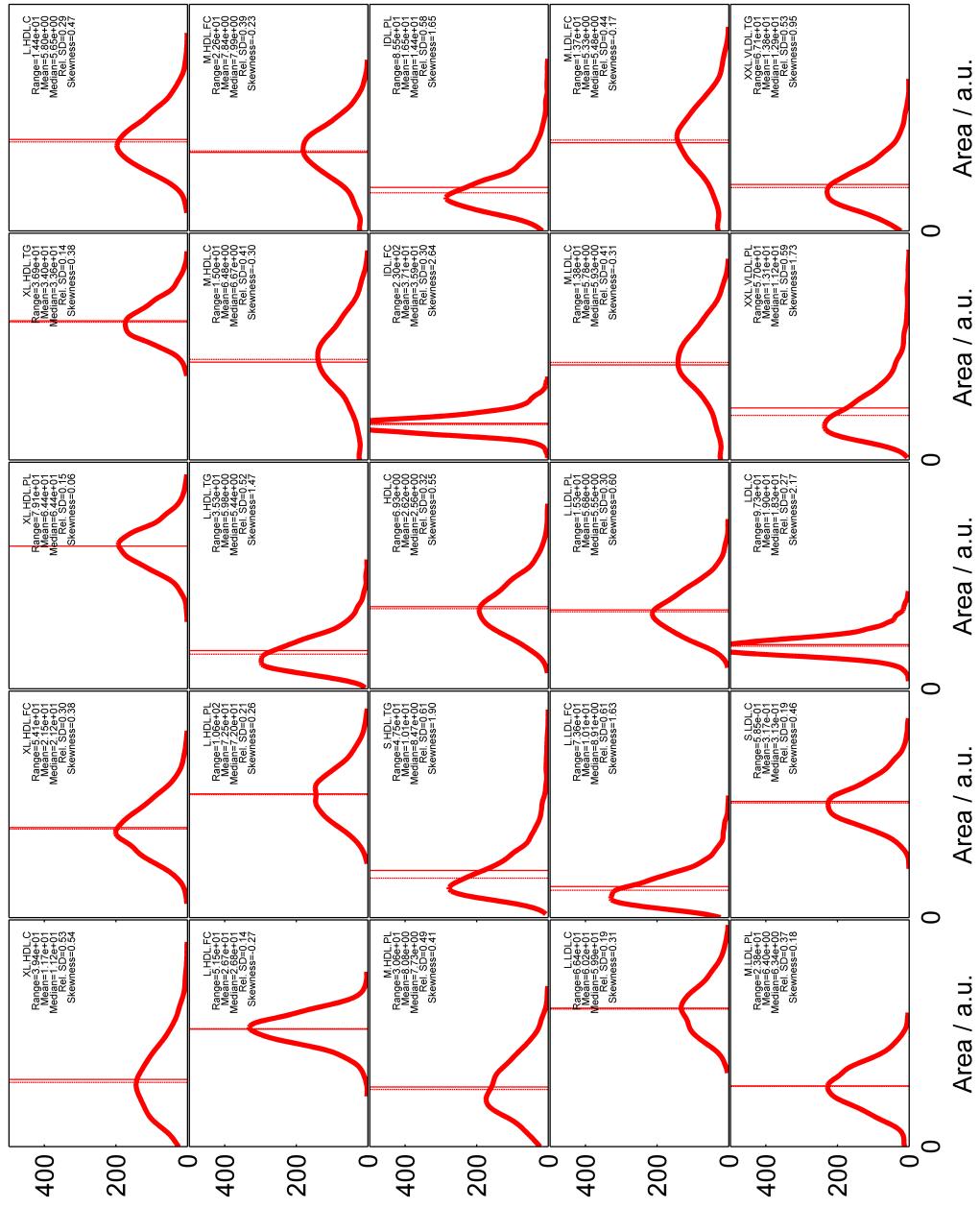


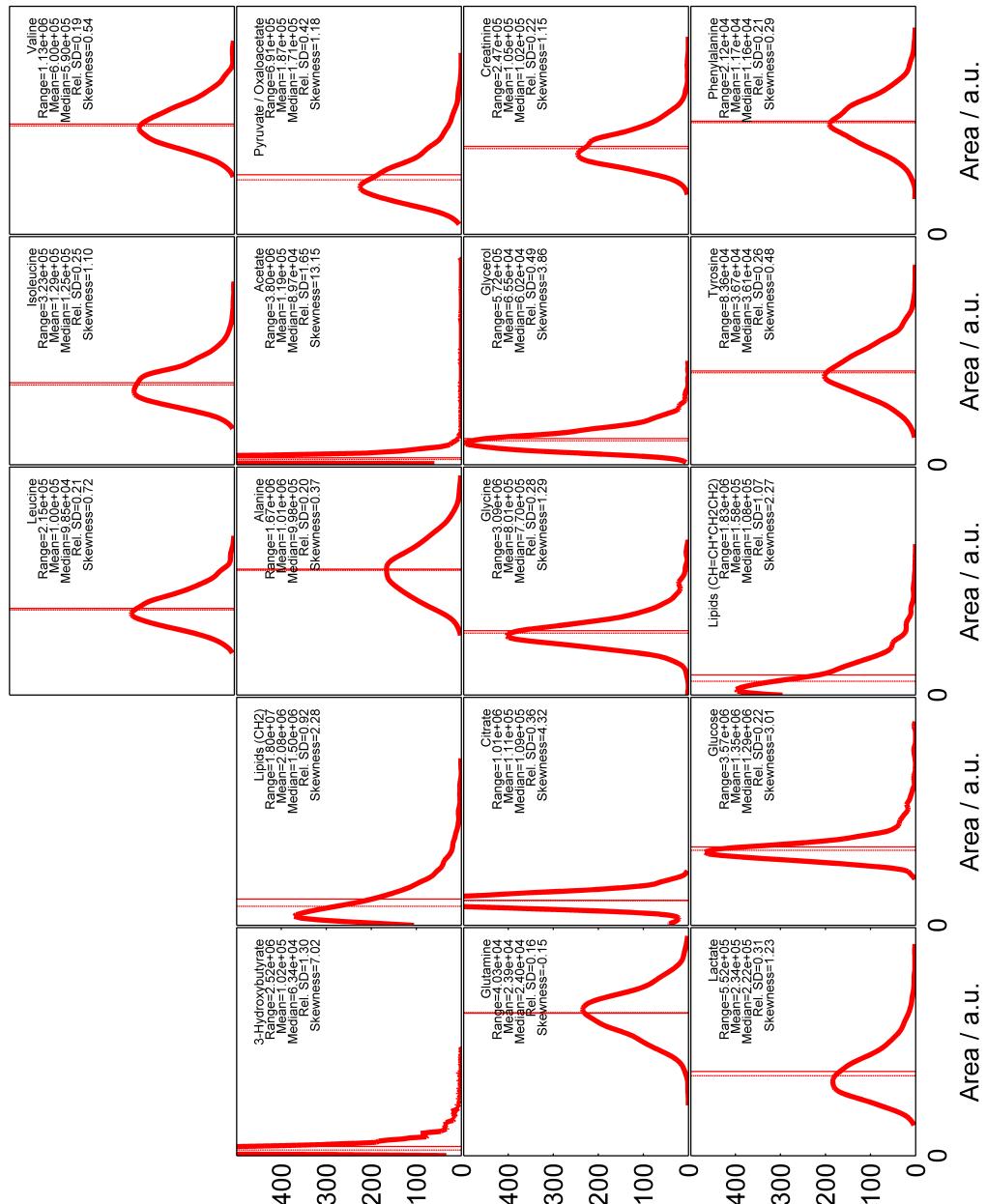
## NFBG66 measures summary



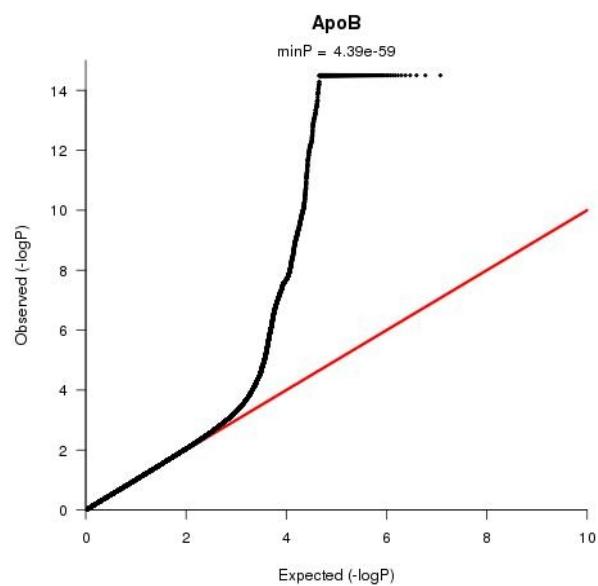
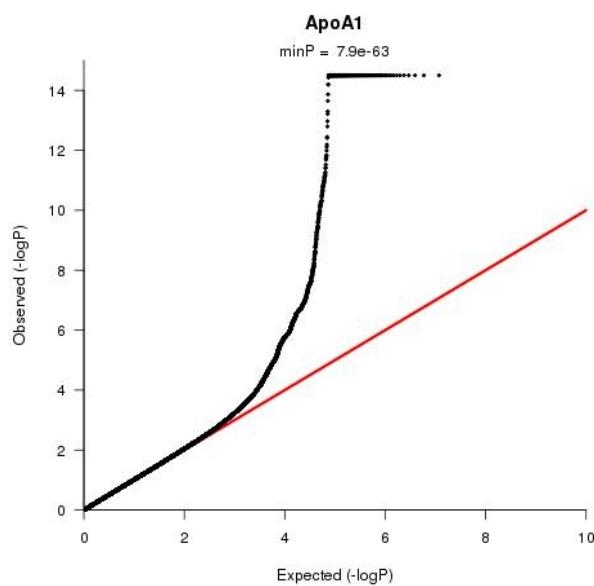
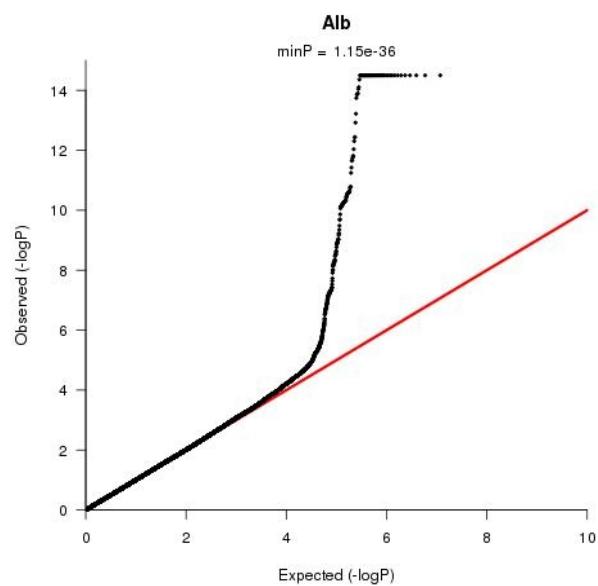
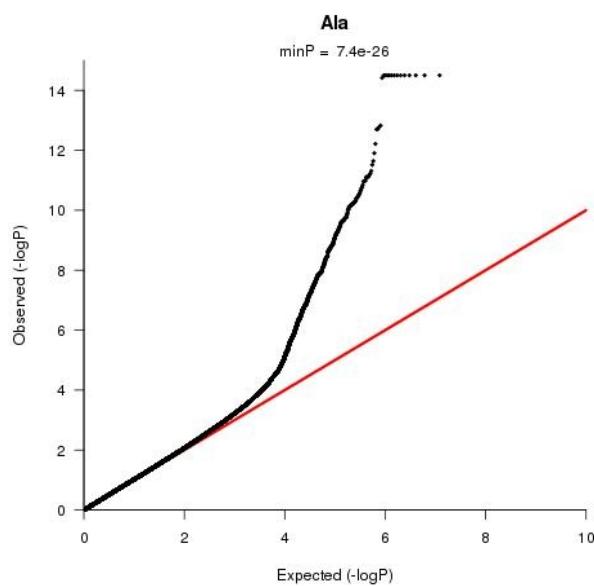
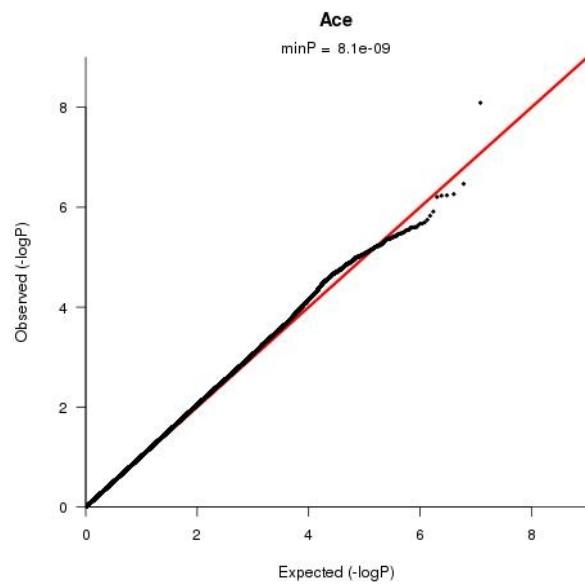
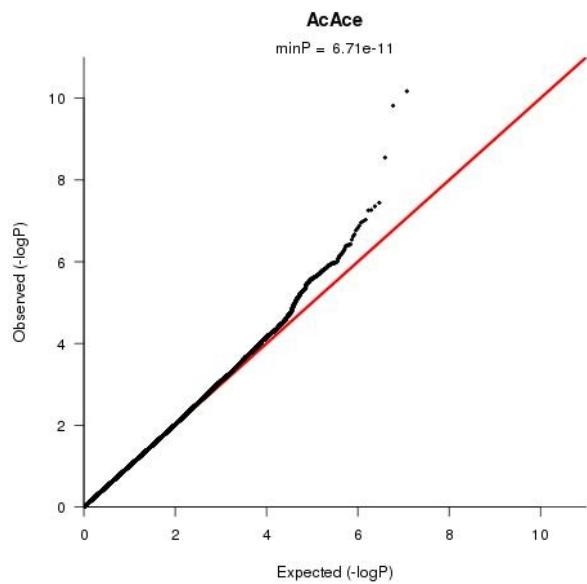
**Supplementary Figure 4.** Metabolite-specific original (untransformed) distributions and descriptive statistics (for each metabolite; range, mean, median, relative standard deviation and skewness) from ERF cohort ( $N = 2118$ ) for available metabolic measures.

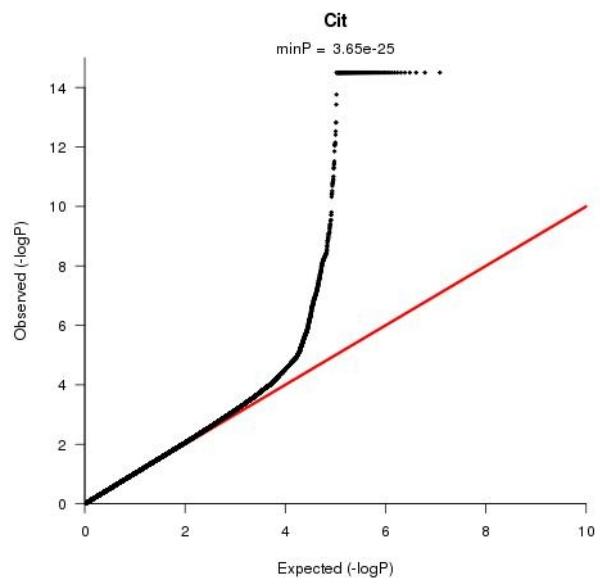
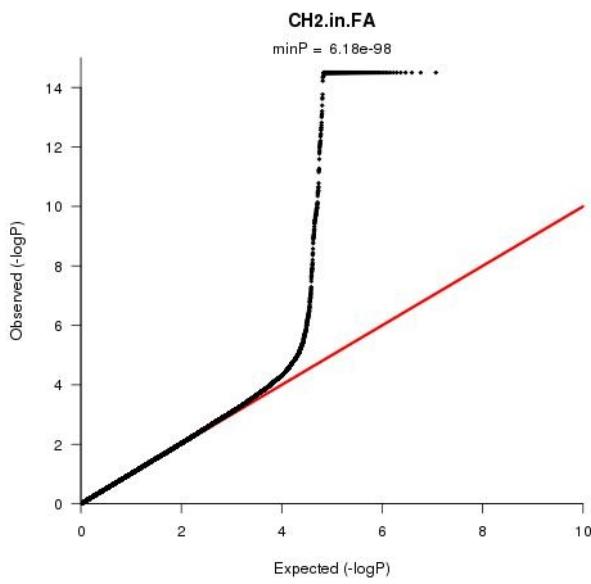
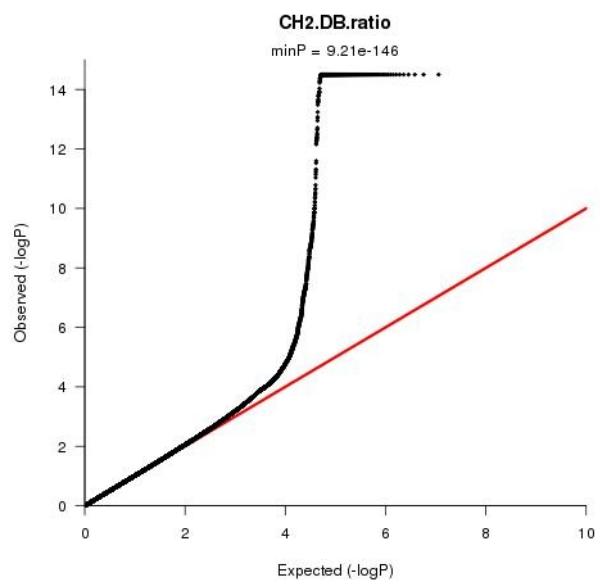
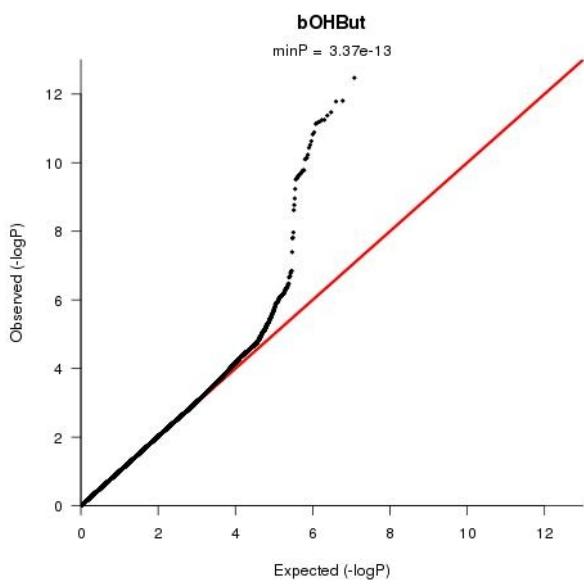
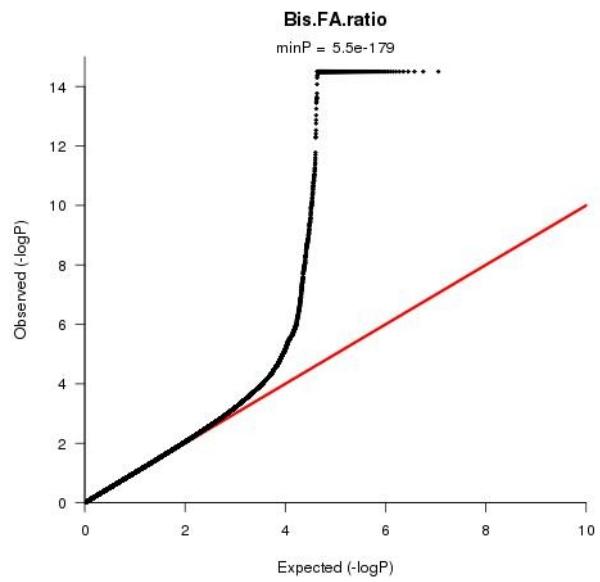
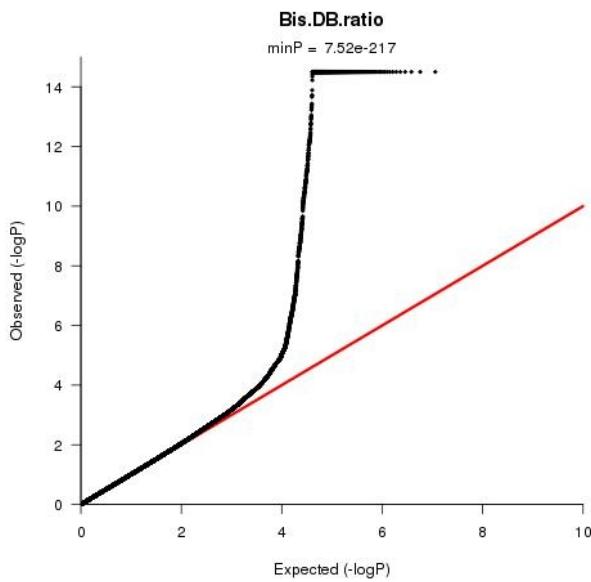


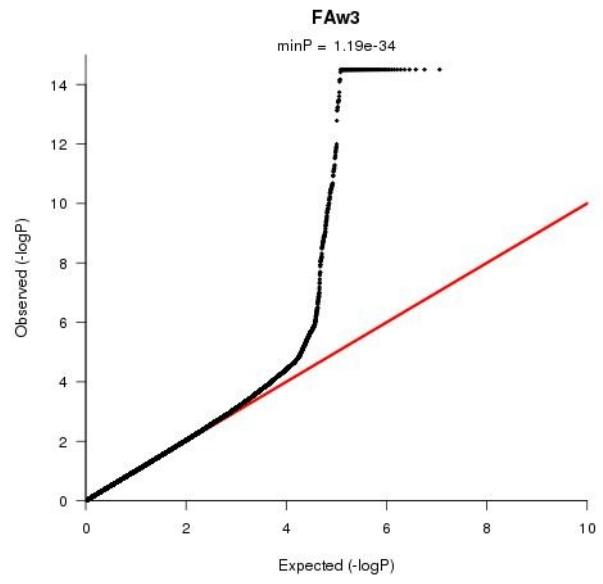
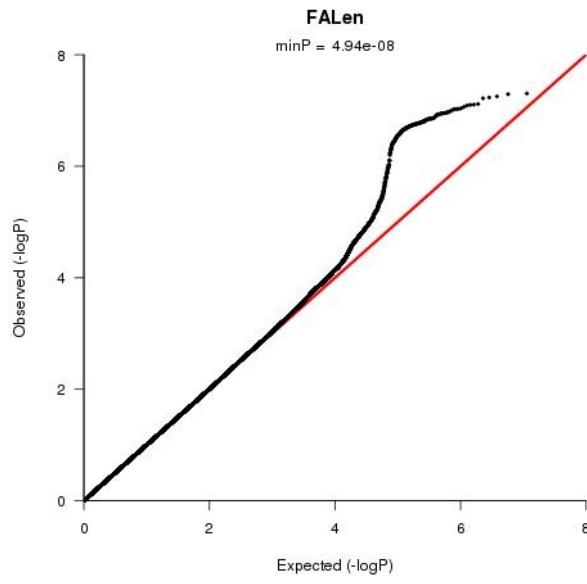
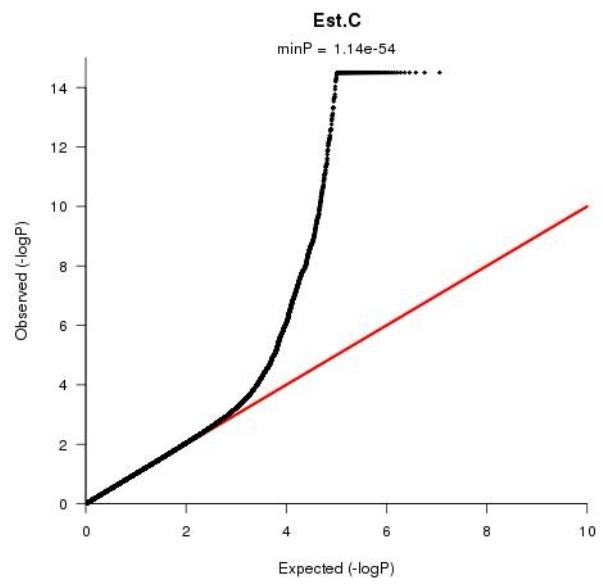
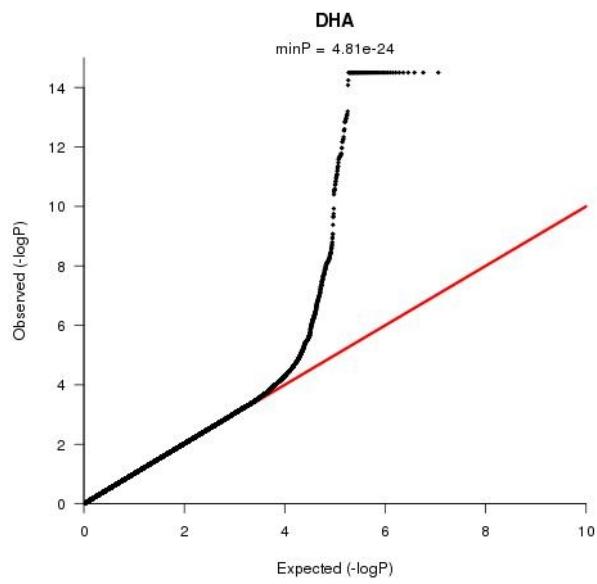
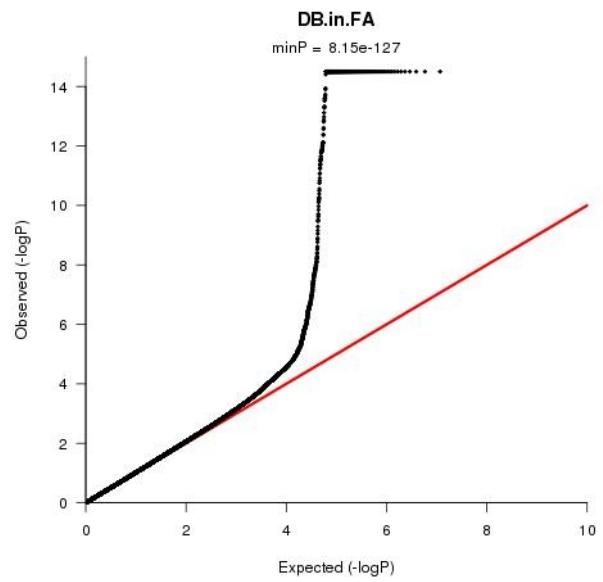
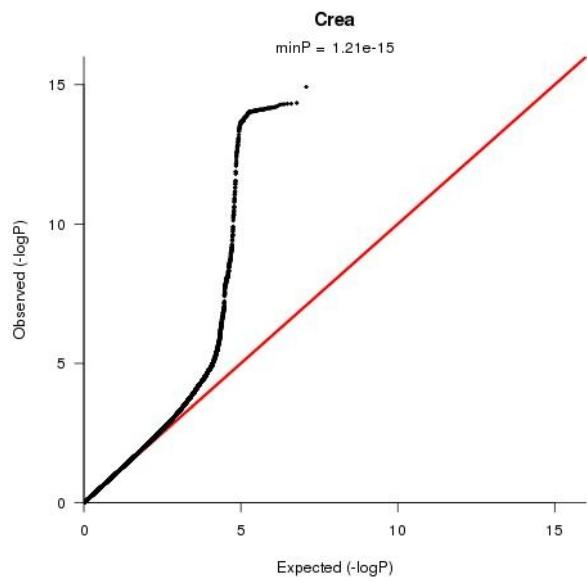


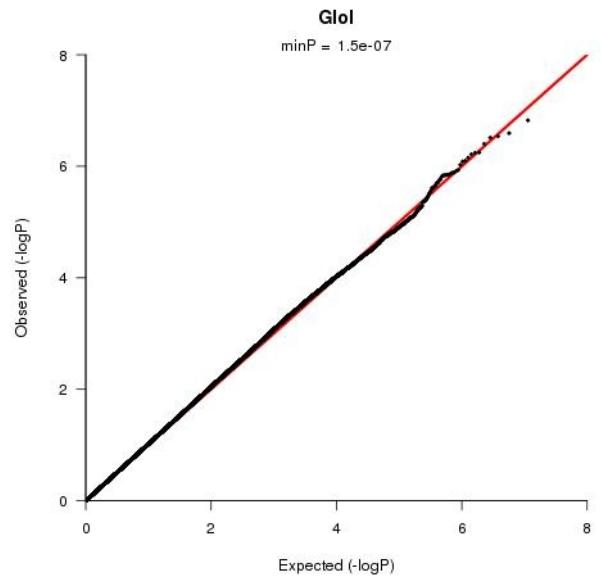
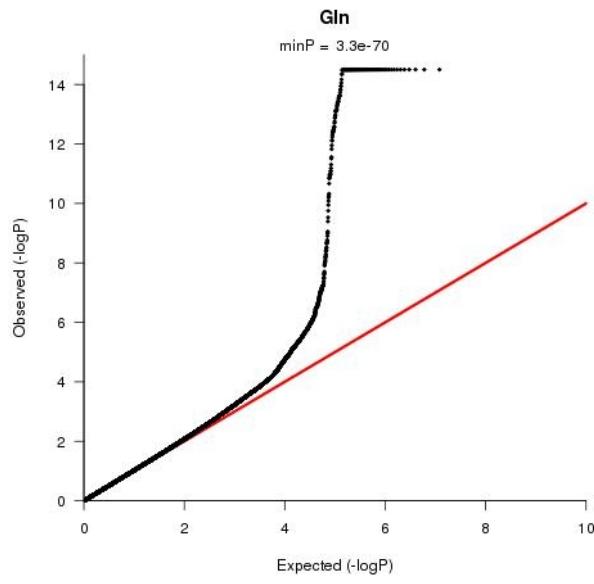
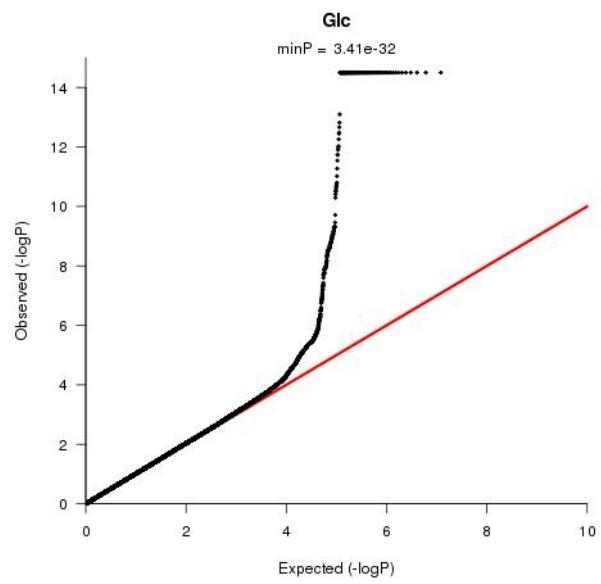
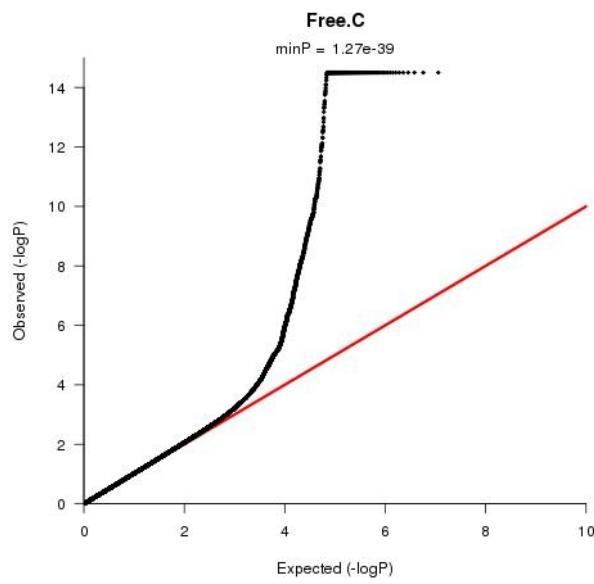
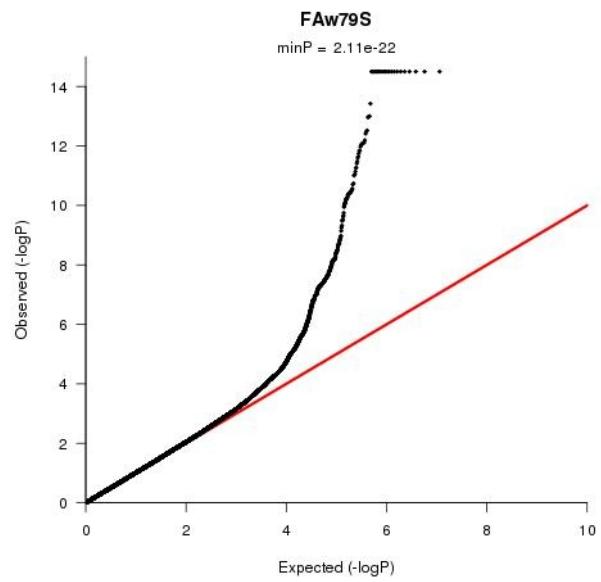
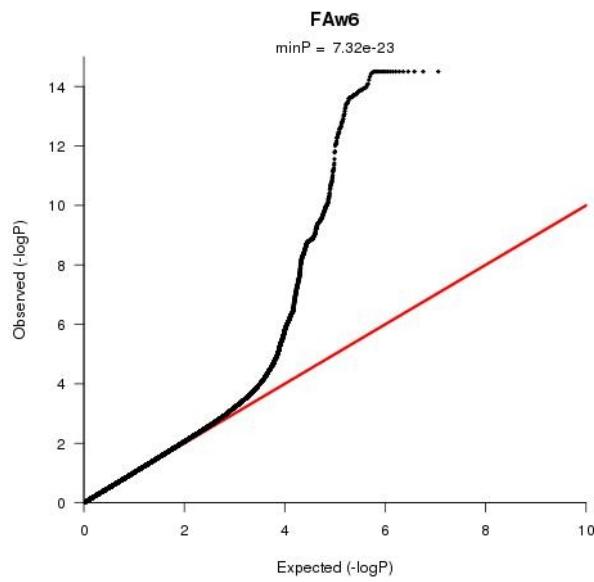


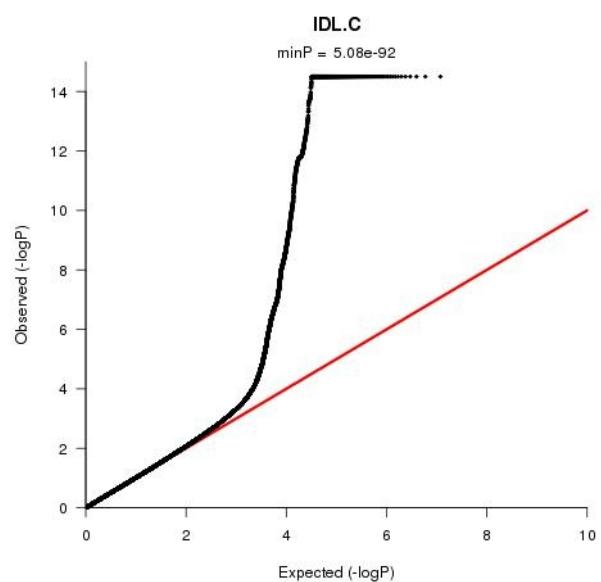
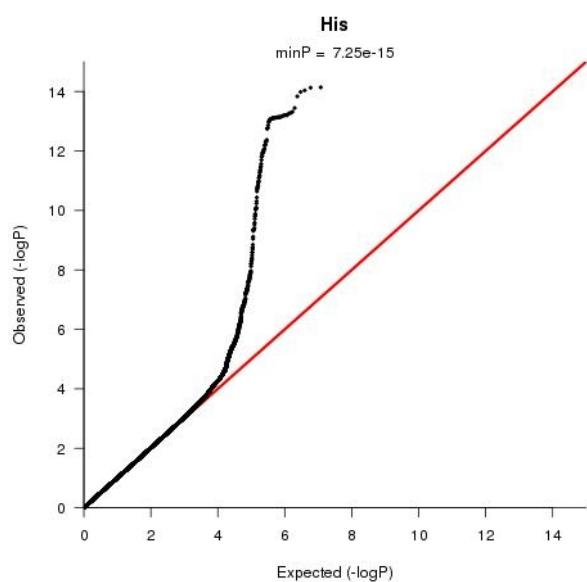
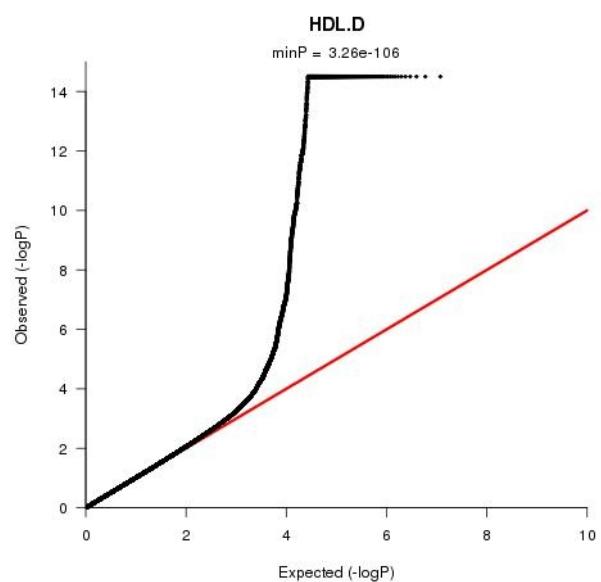
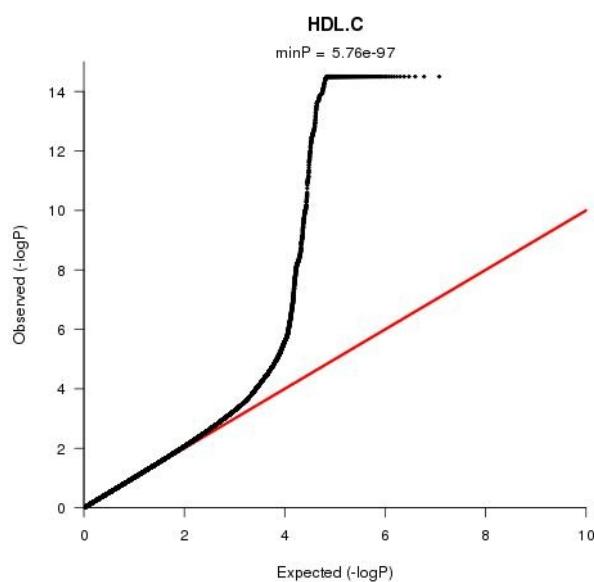
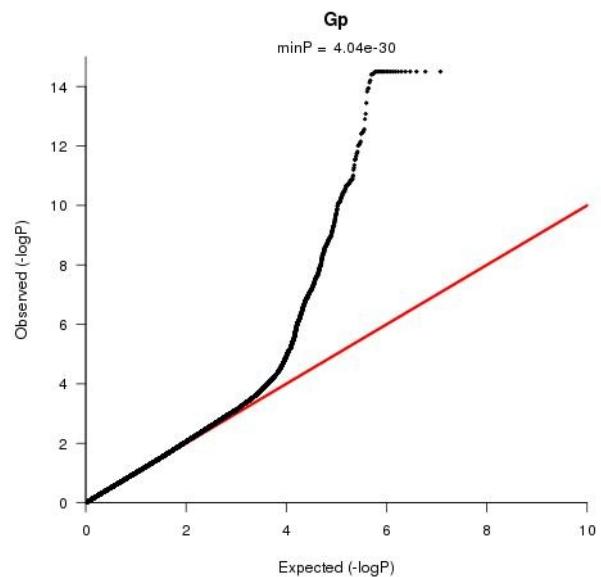
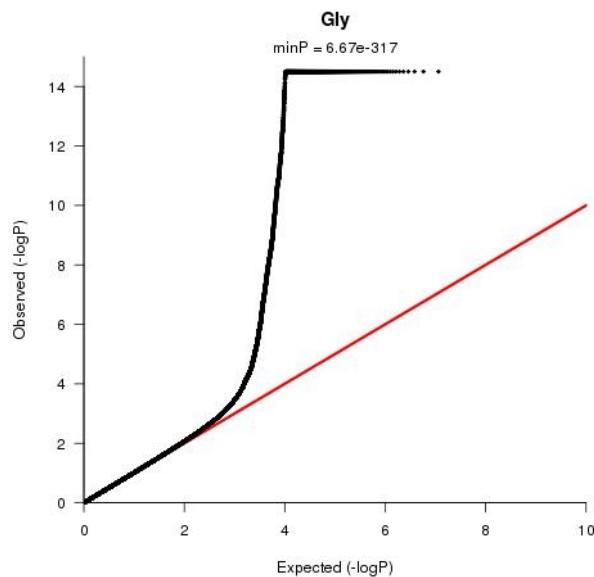
**Supplementary Figure 5.** Quantile plots for all traits. The p-value has been truncated at  $1 \times 10^{-14}$  to allow the visualization of over all inflation, i.e. observed distribution departing from expected already at large p-value distribution. The smallest p-value of the distribution is given in the upper part of each figure.

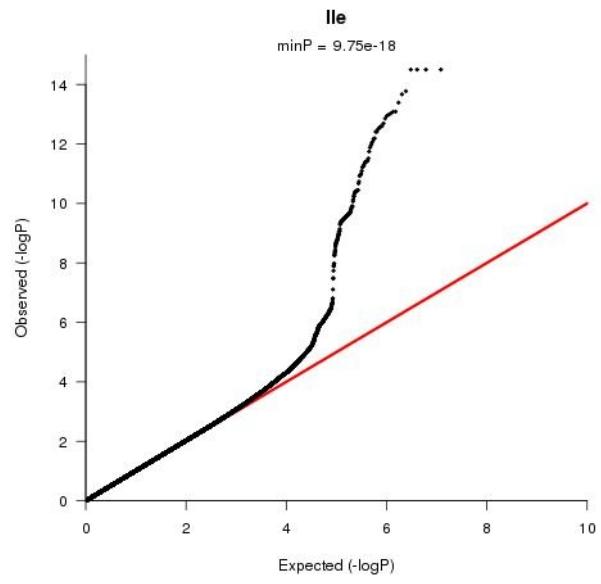
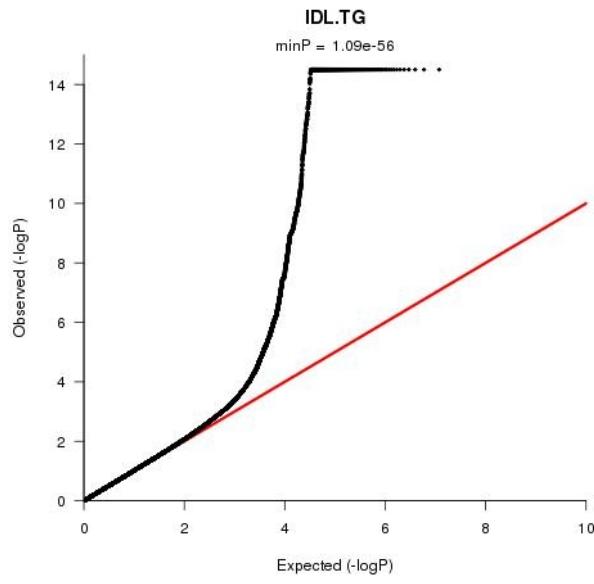
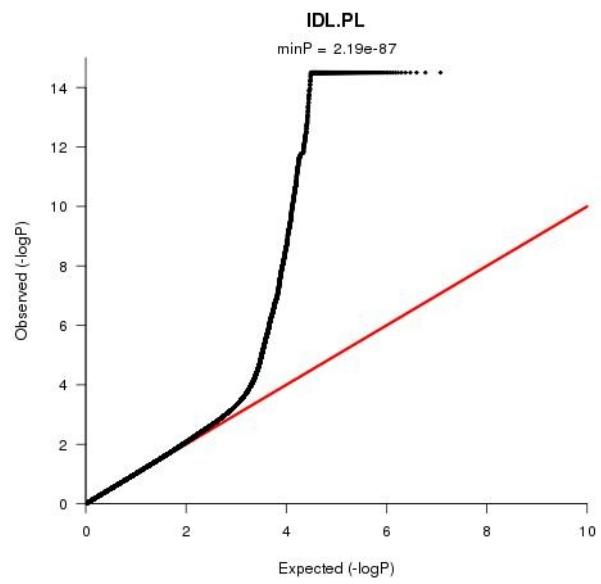
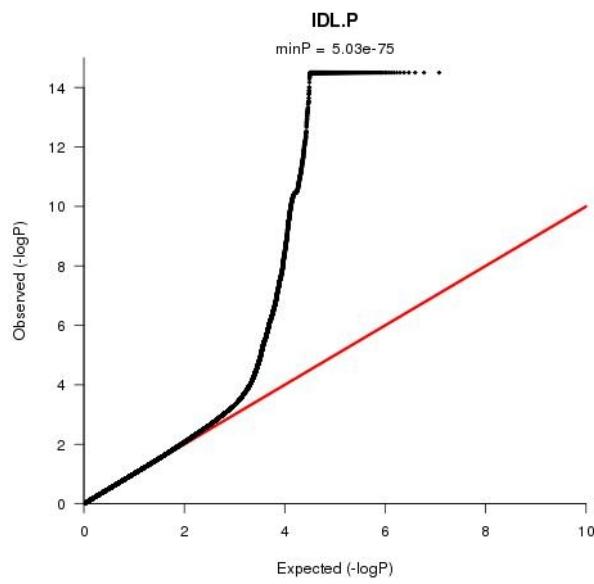
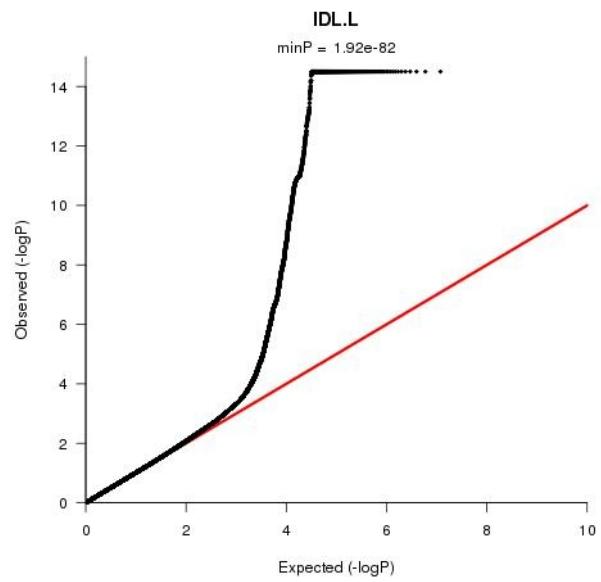
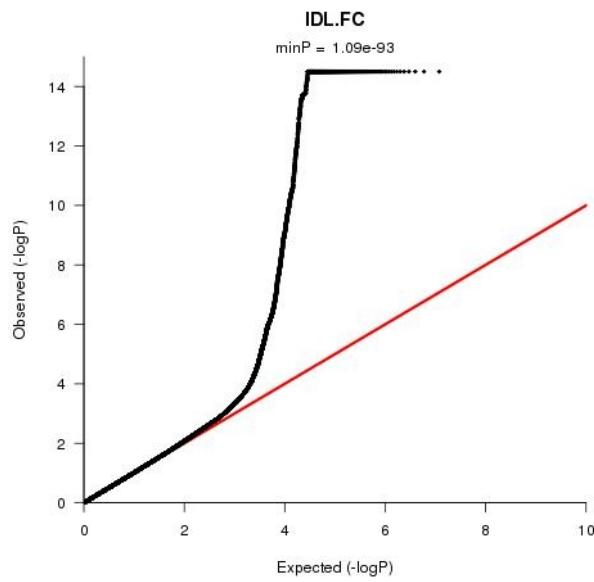


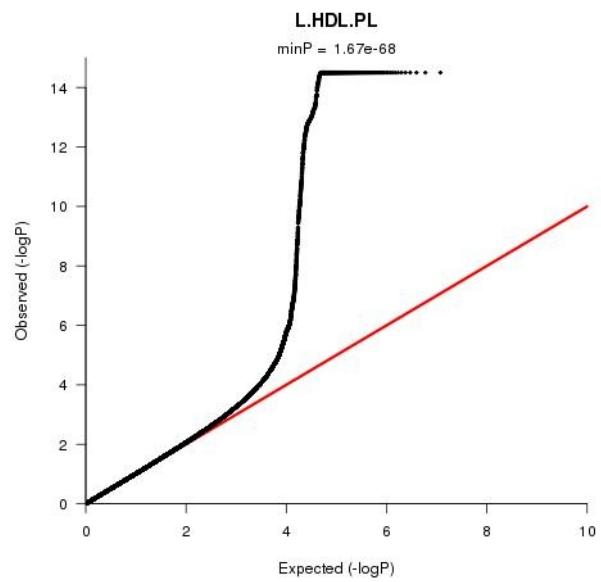
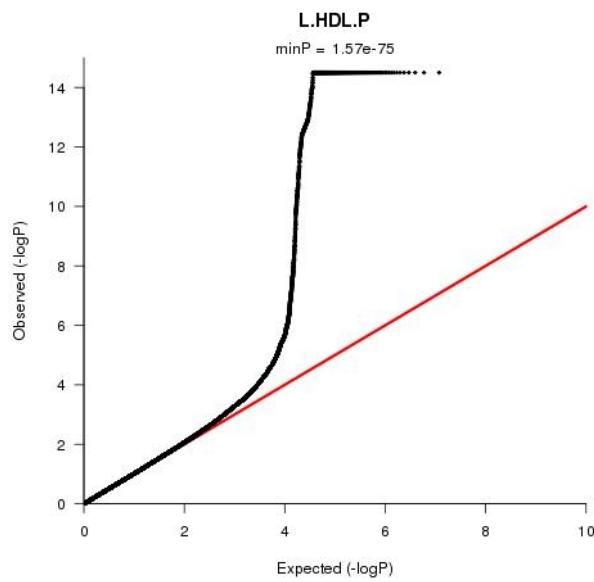
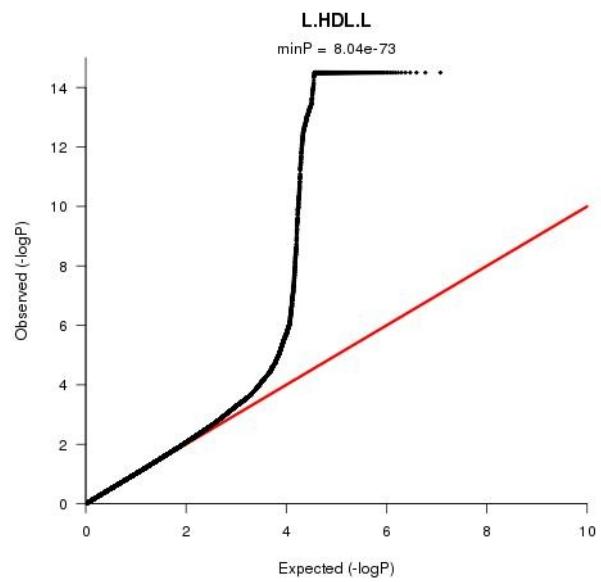
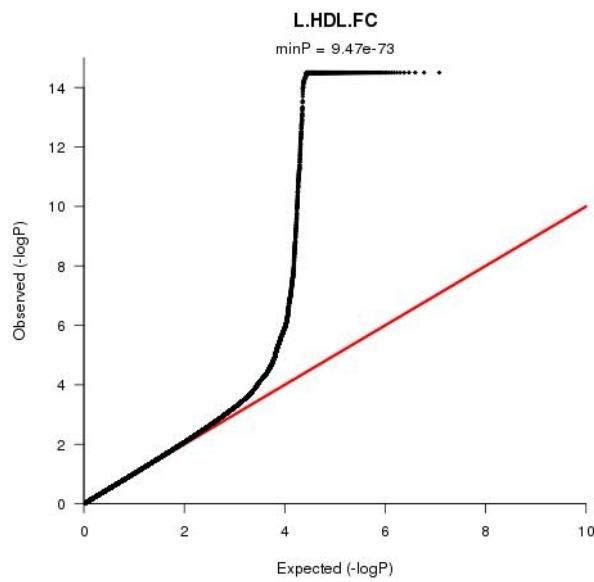
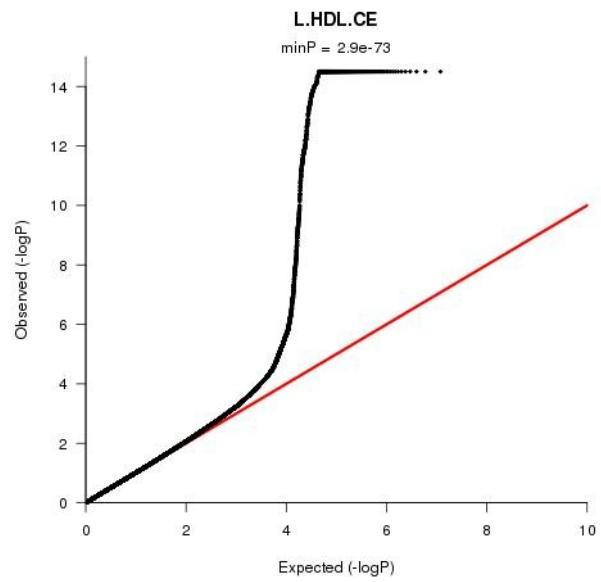
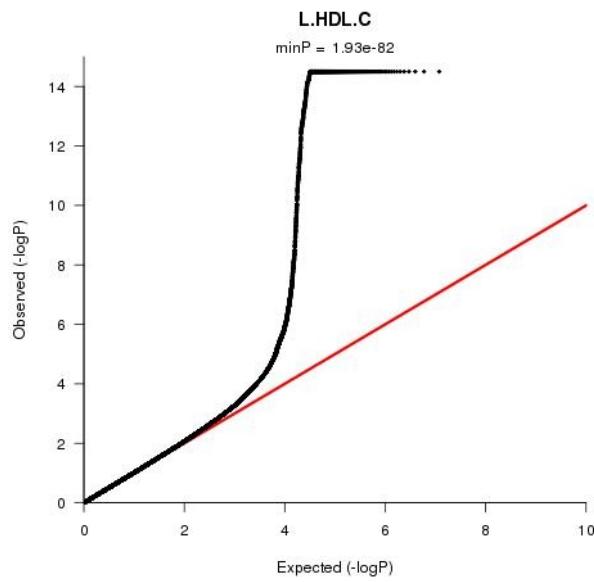


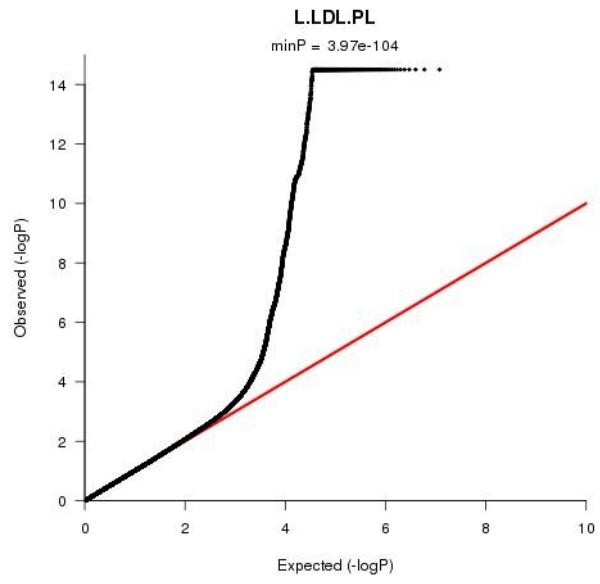
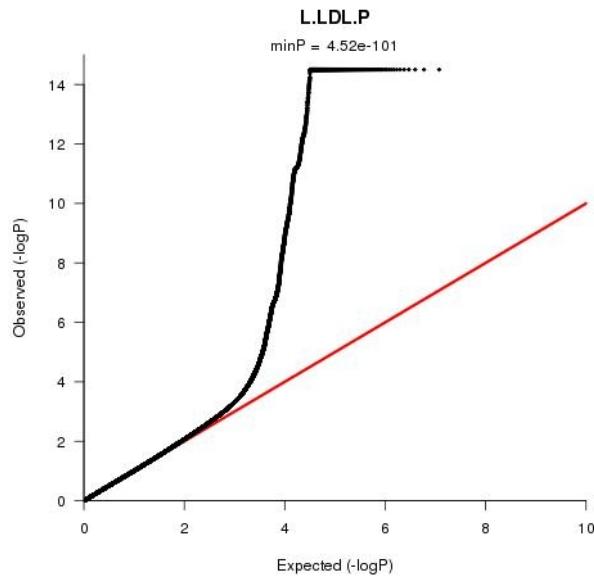
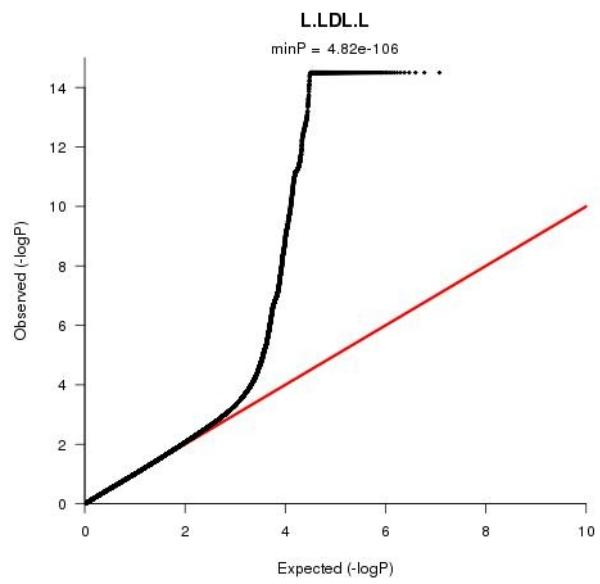
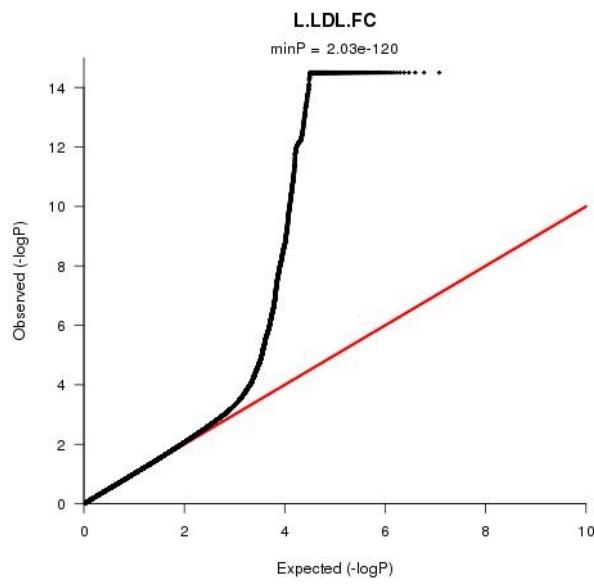
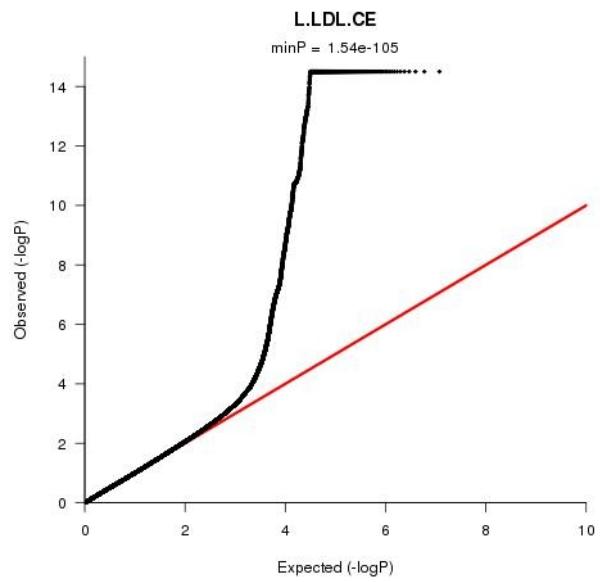
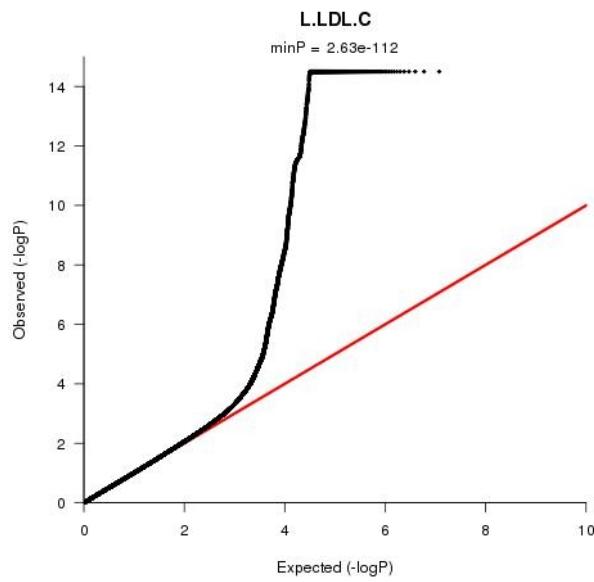


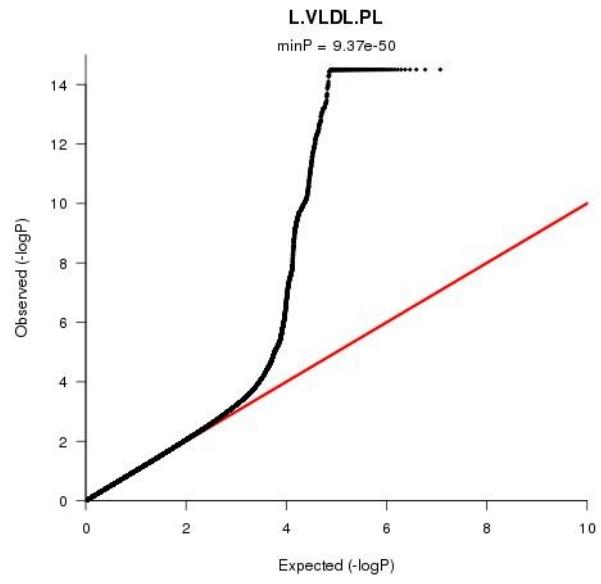
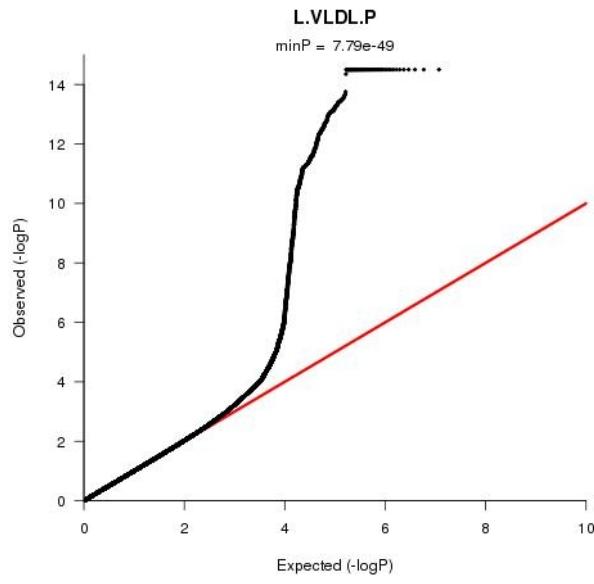
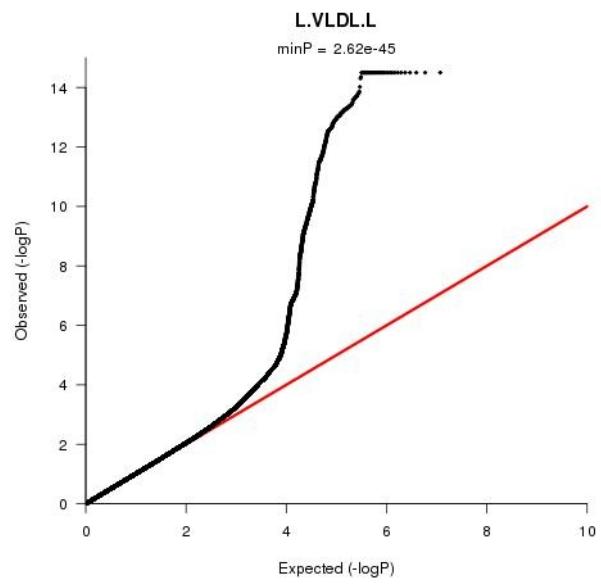
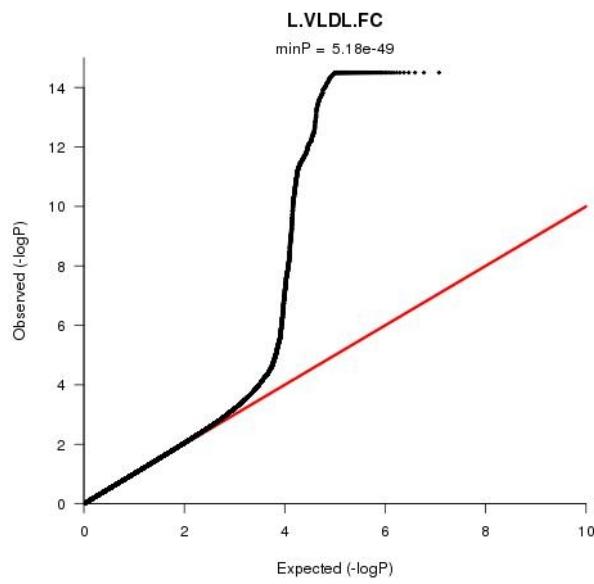
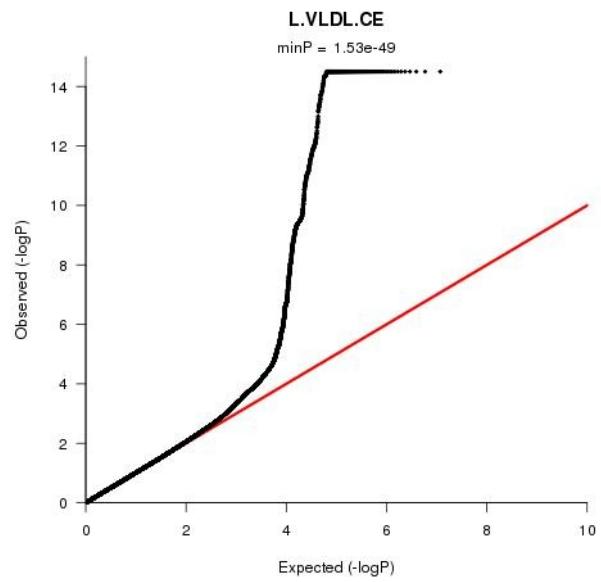
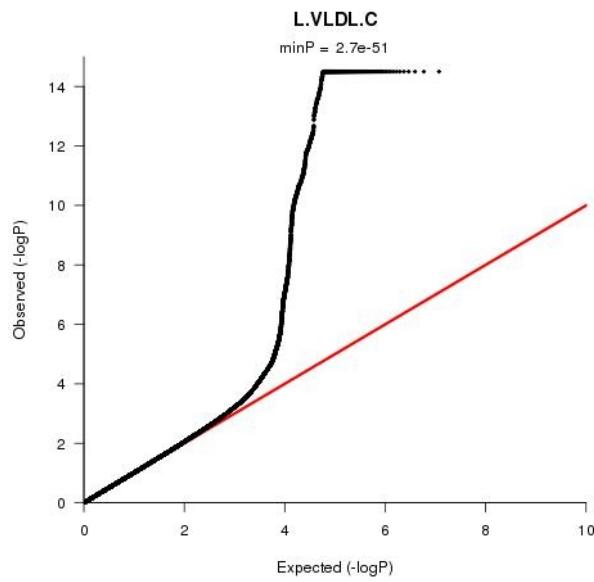


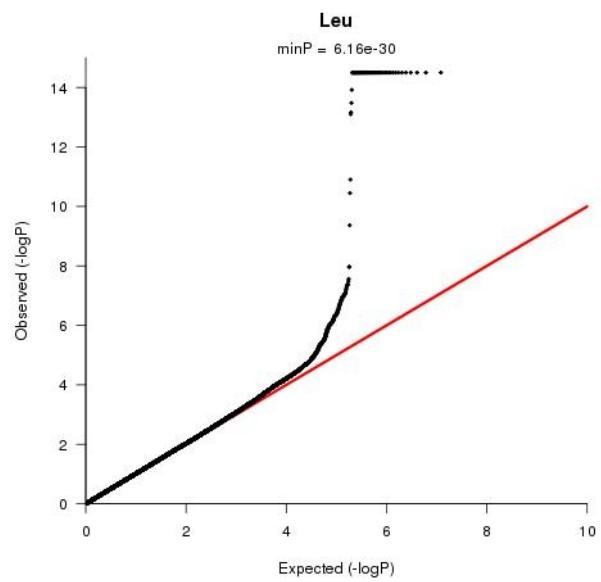
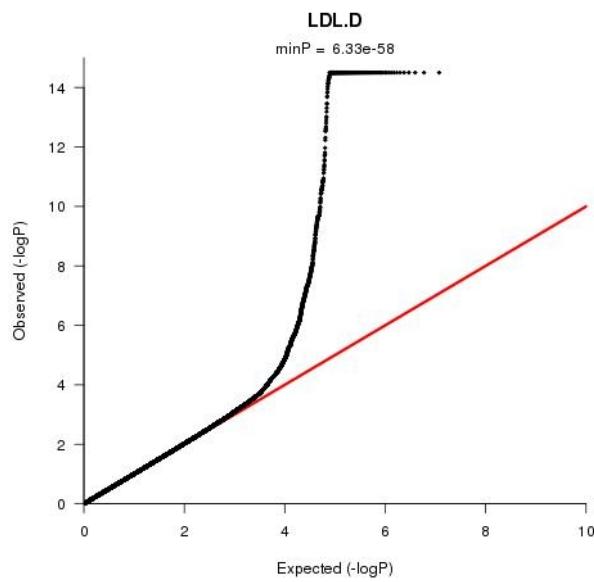
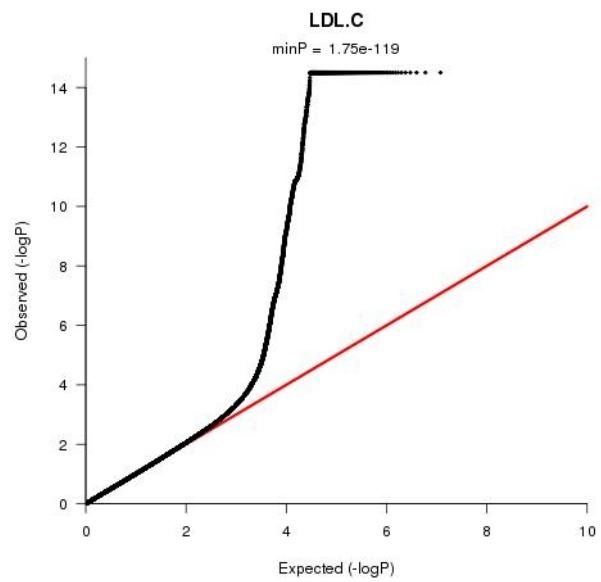
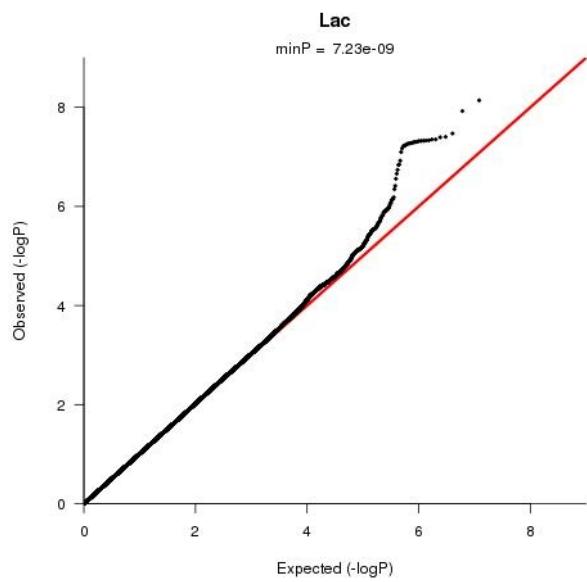
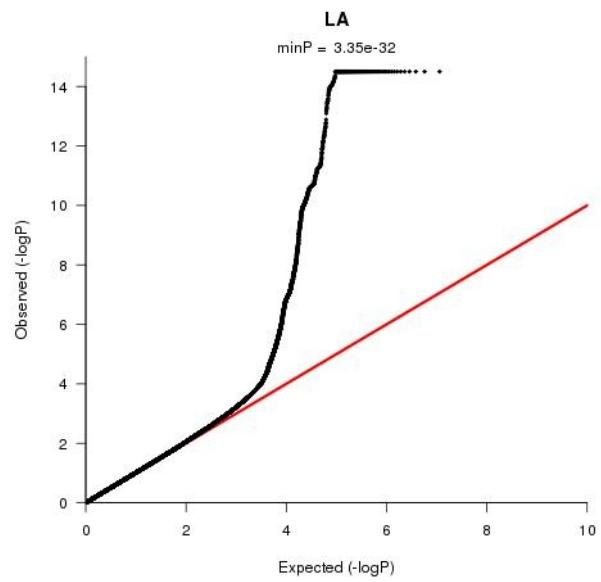
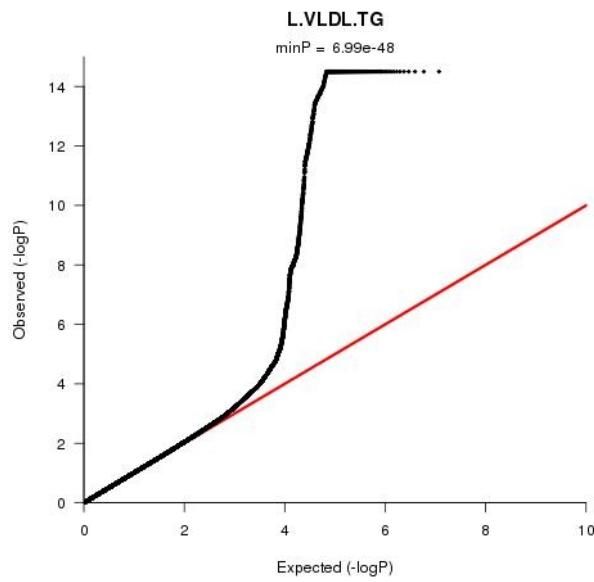


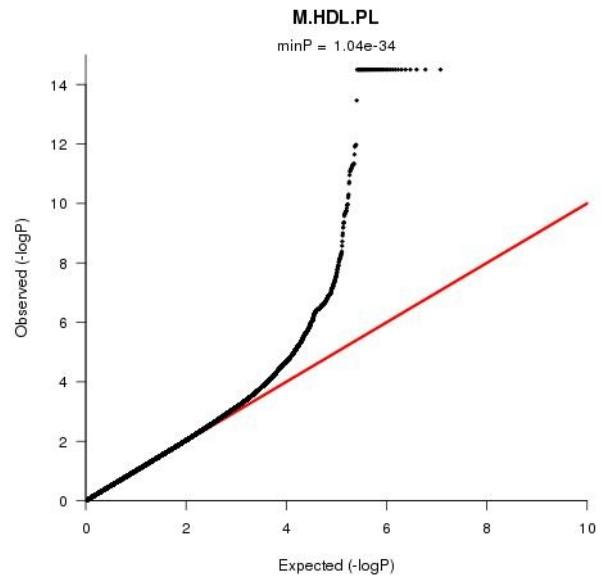
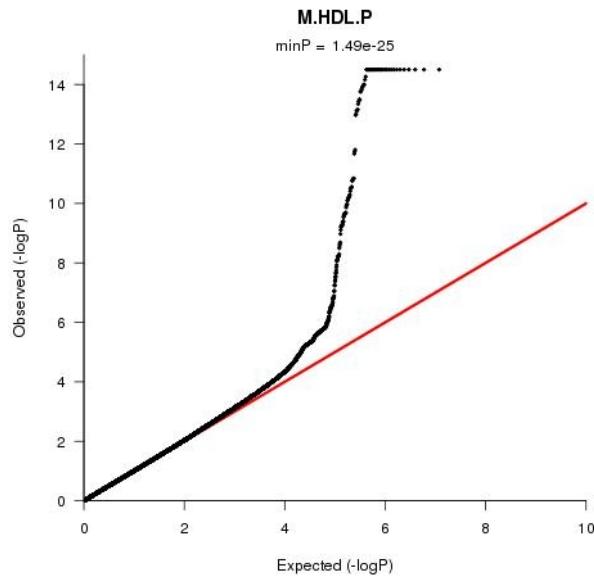
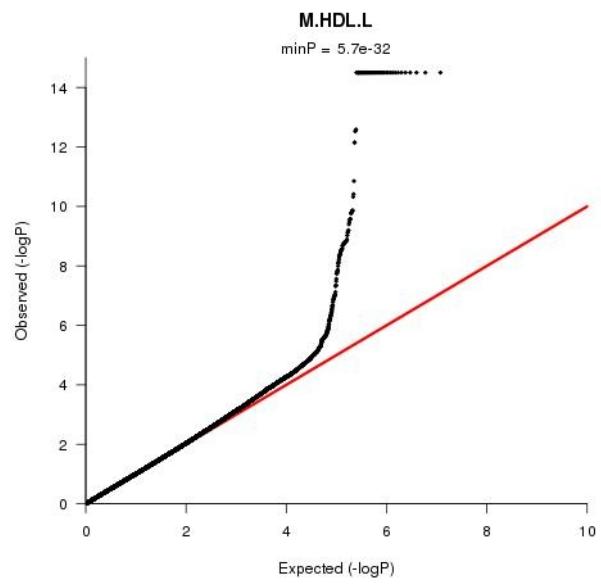
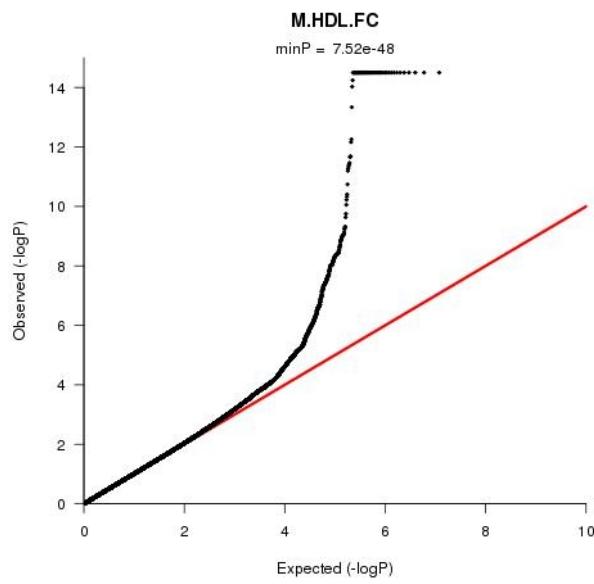
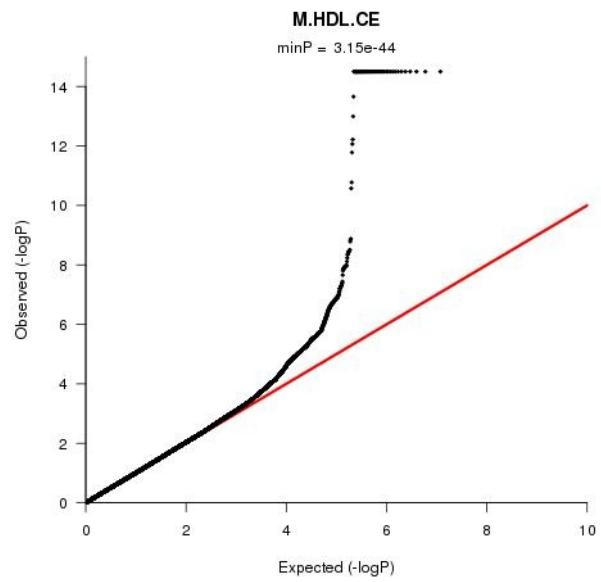
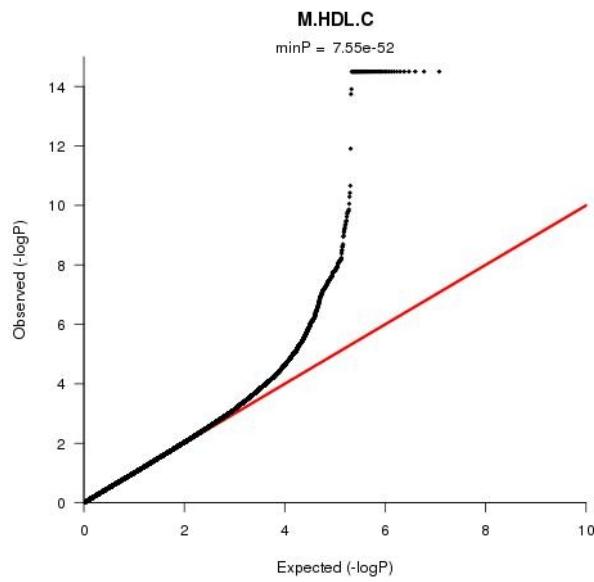


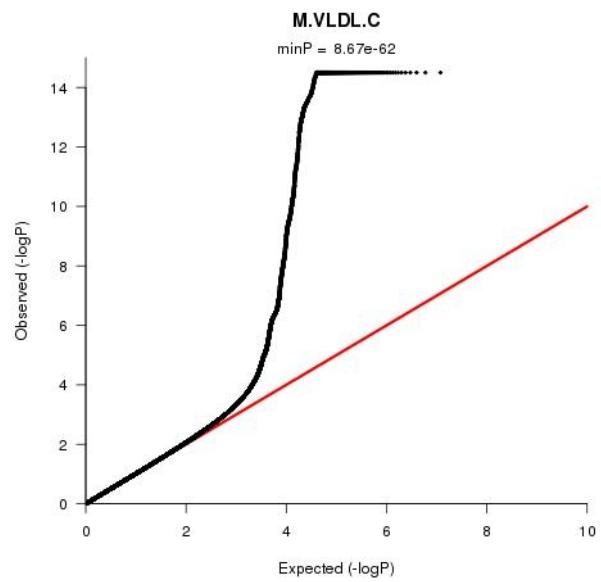
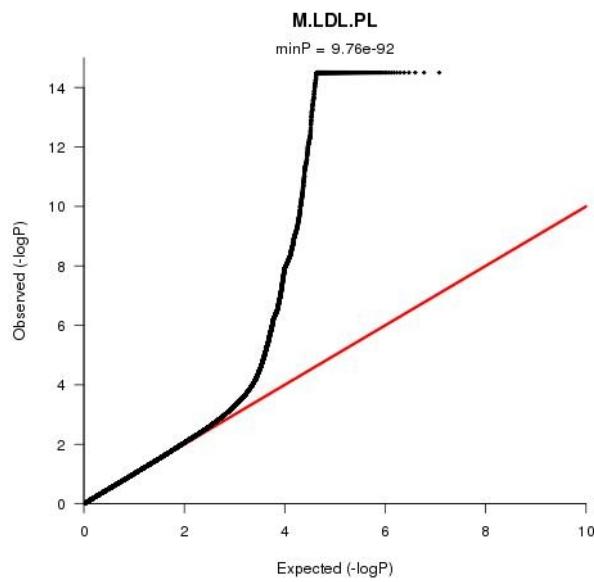
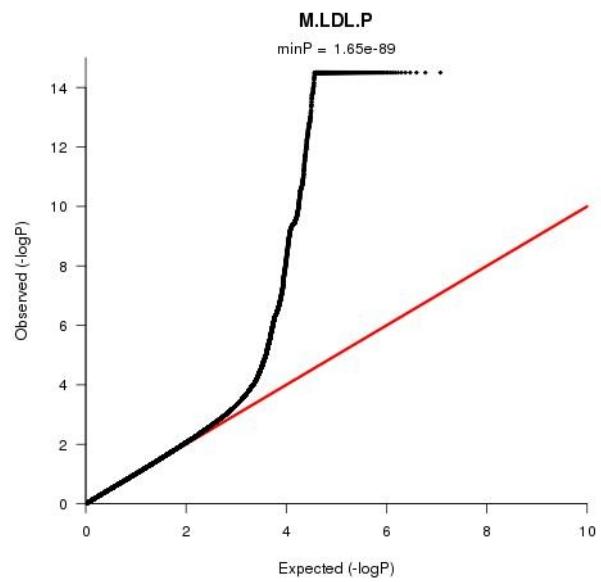
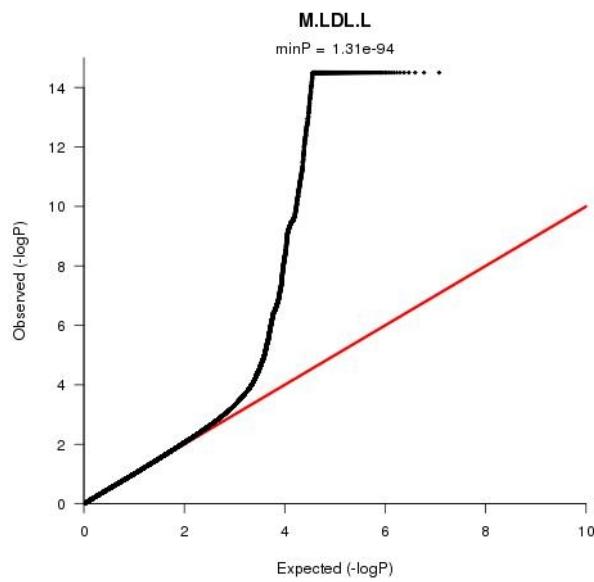
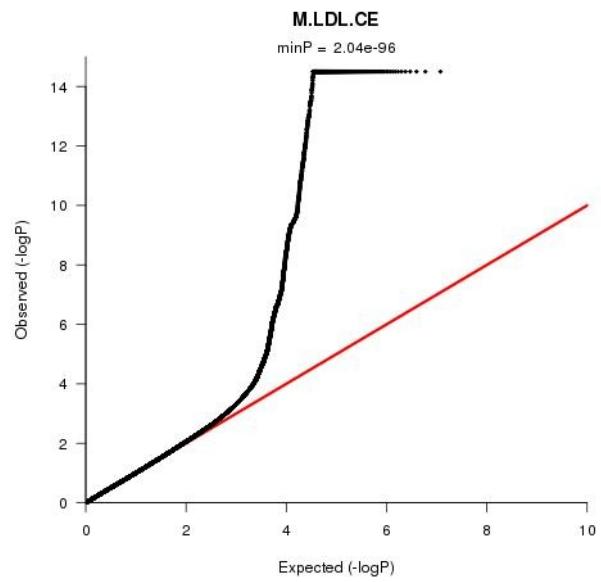
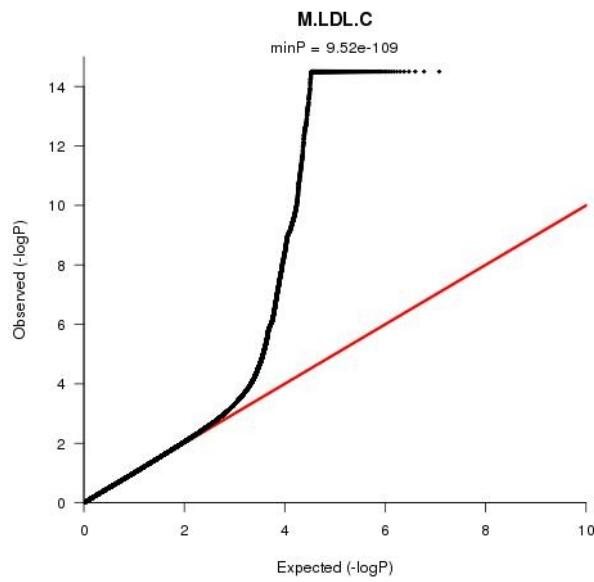


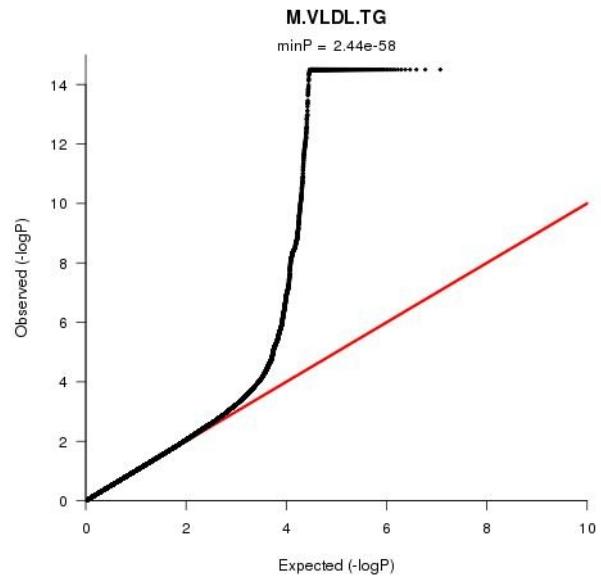
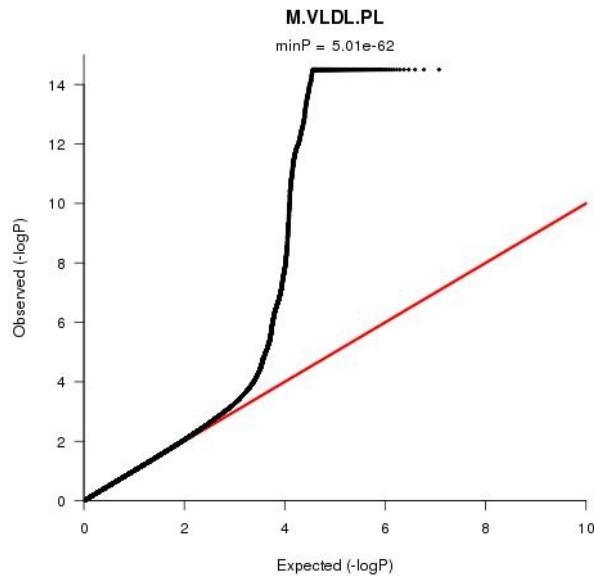
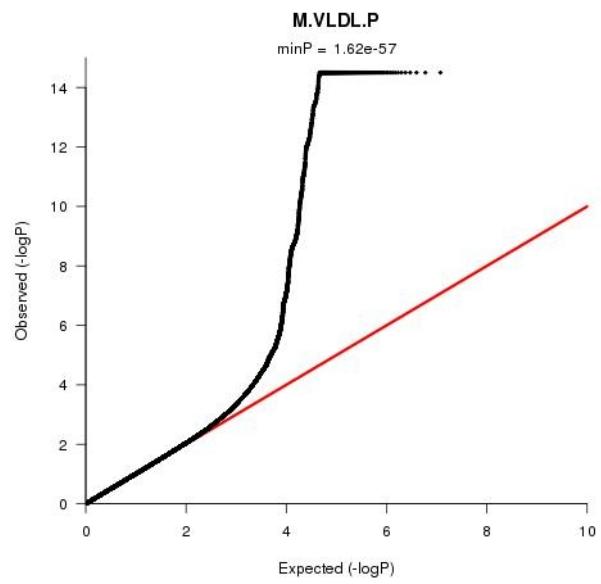
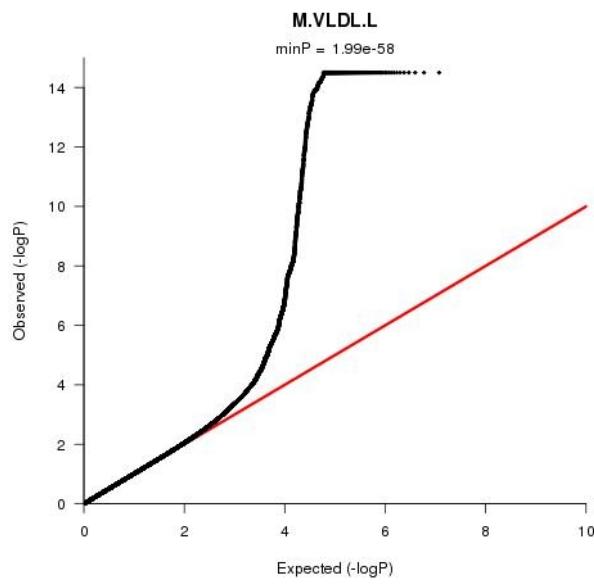
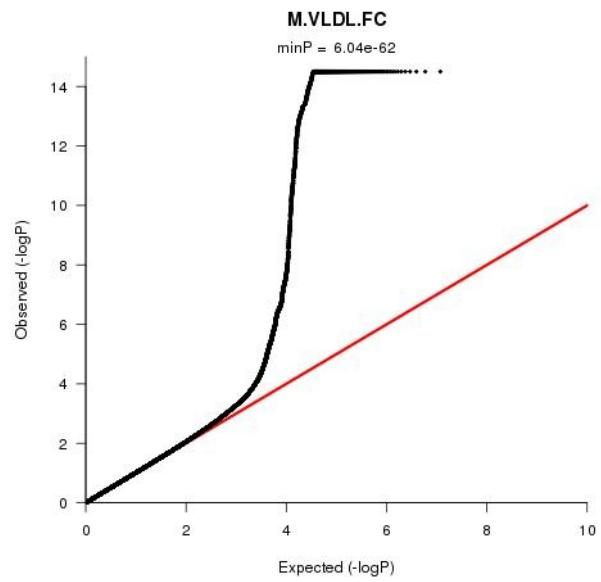
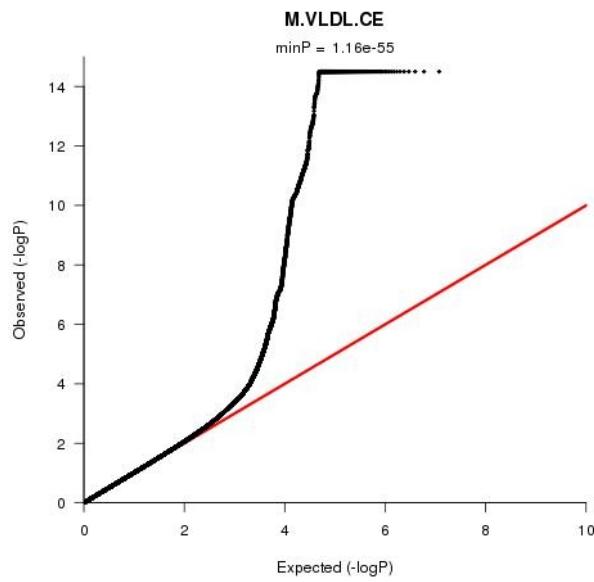


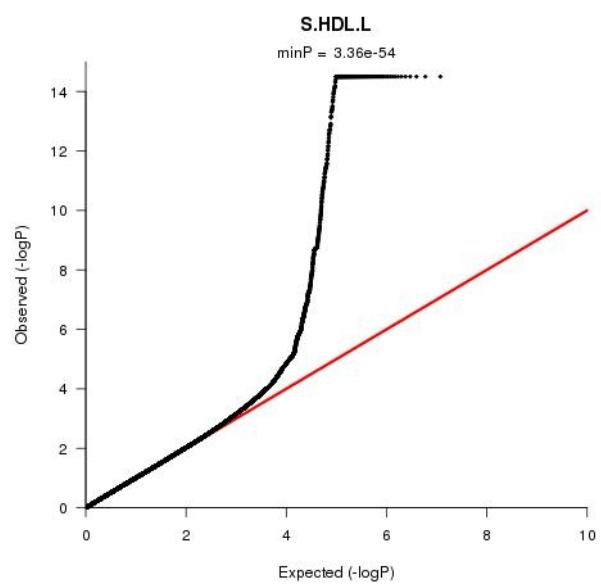
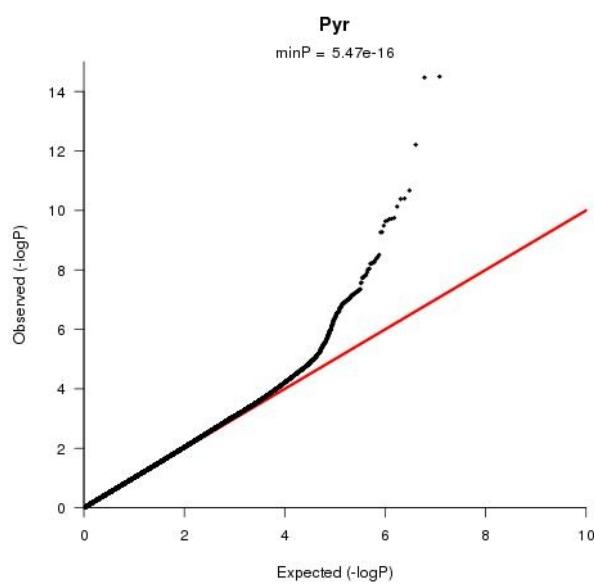
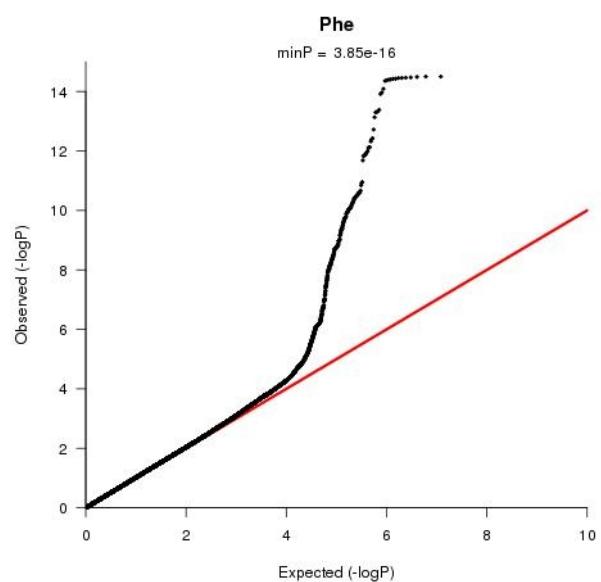
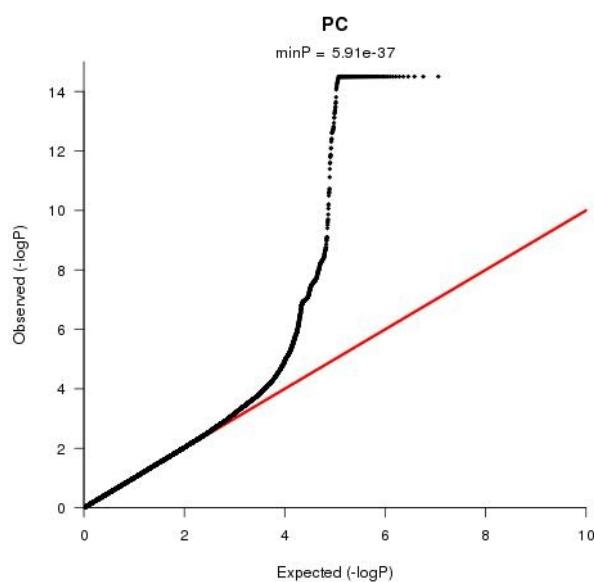
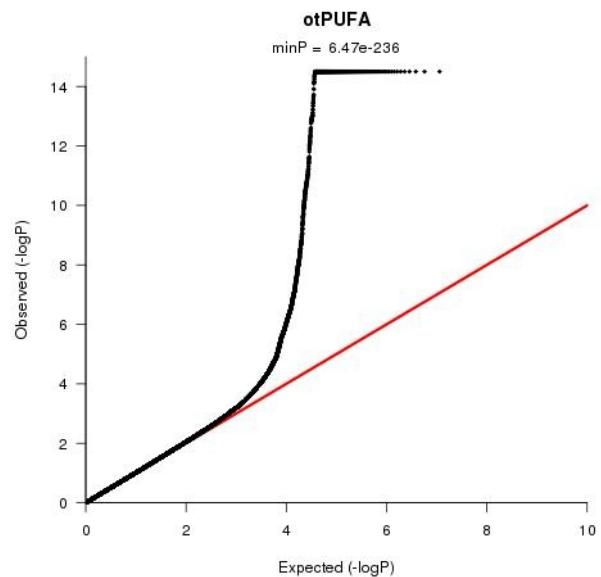
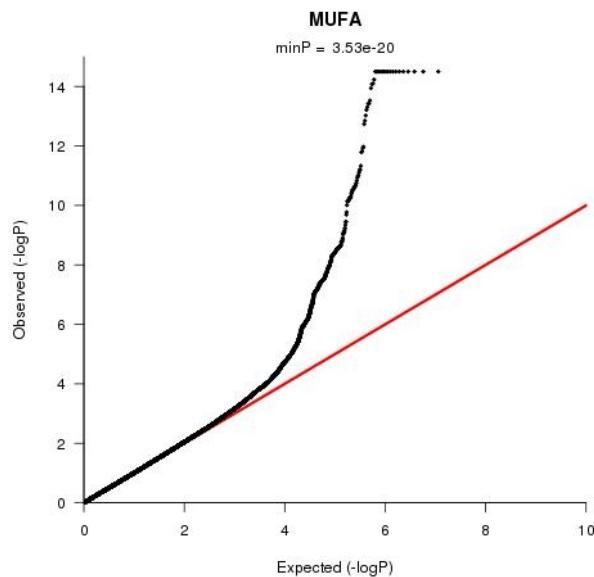


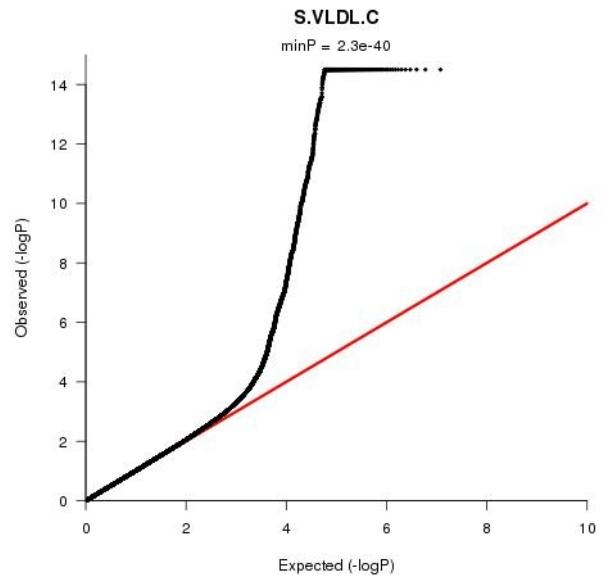
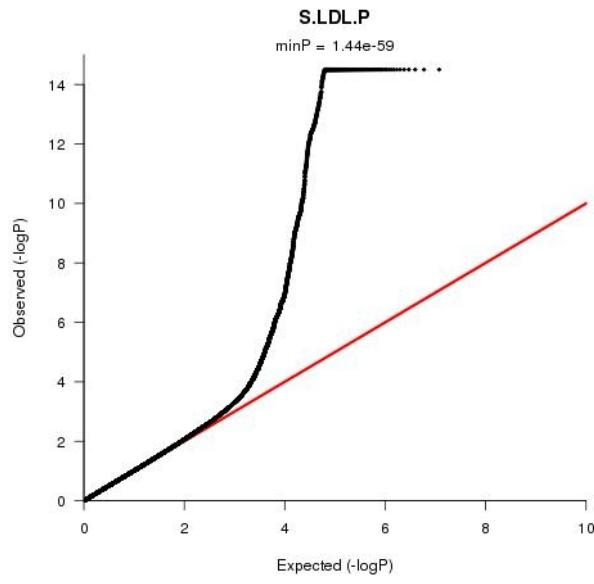
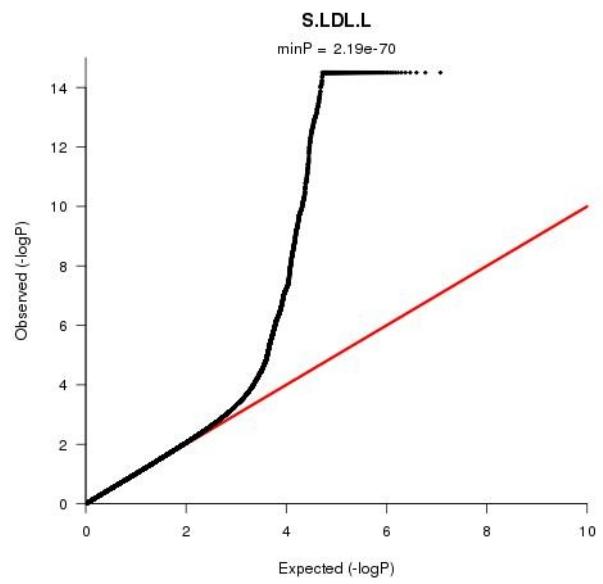
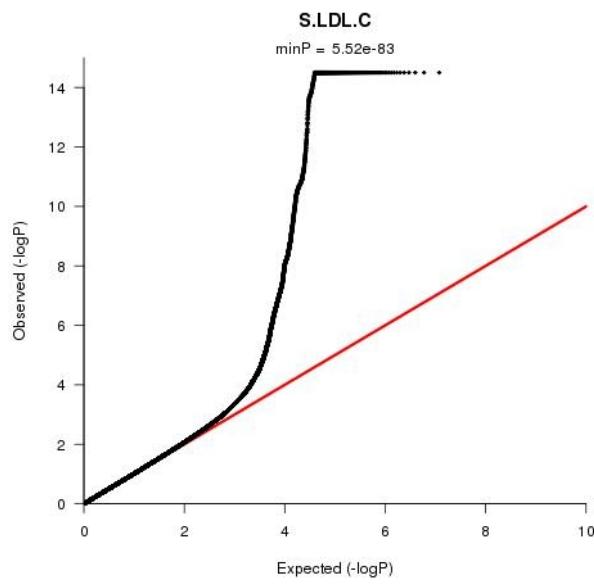
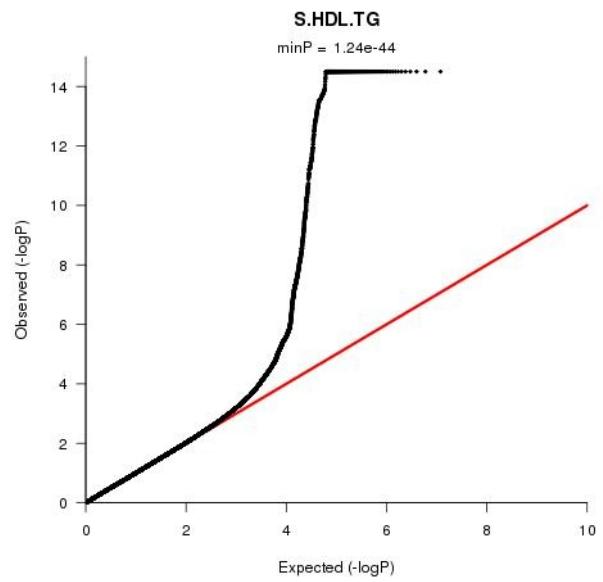
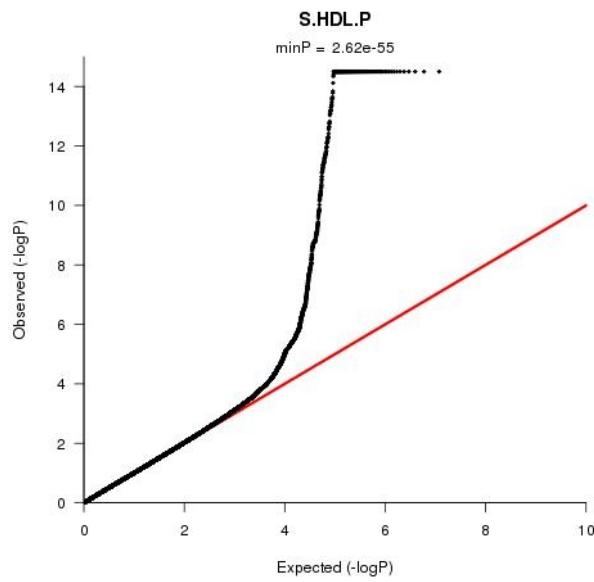


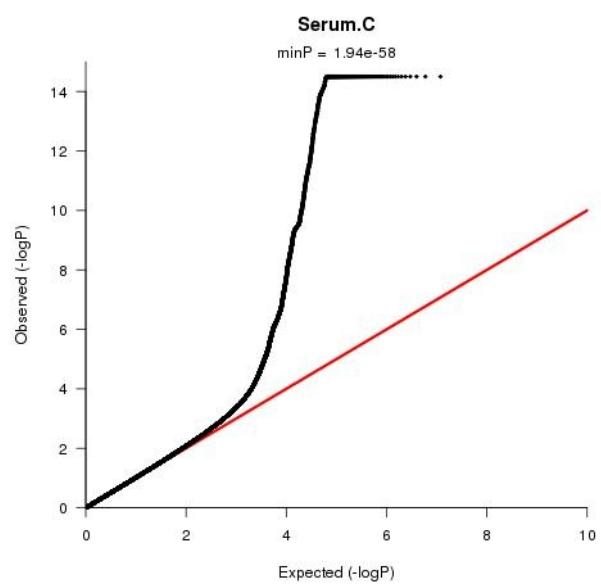
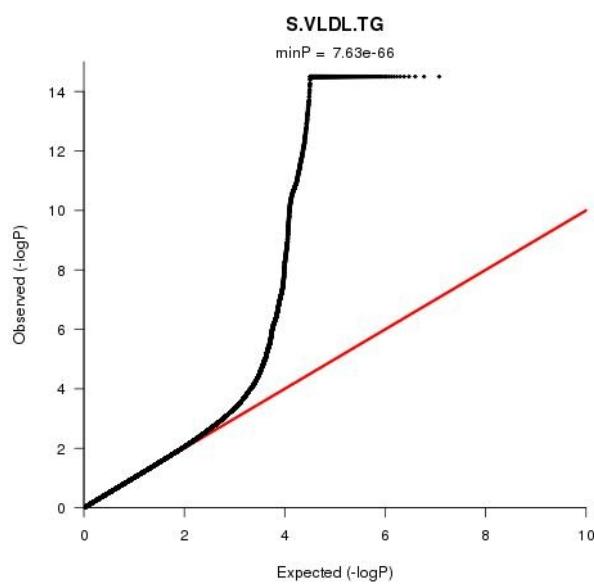
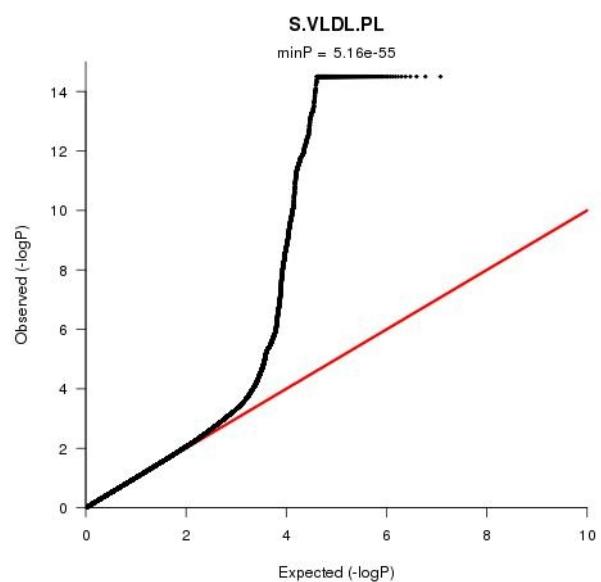
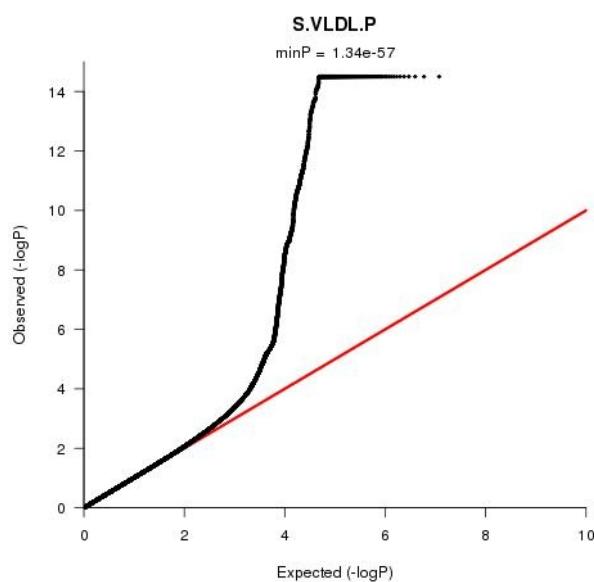
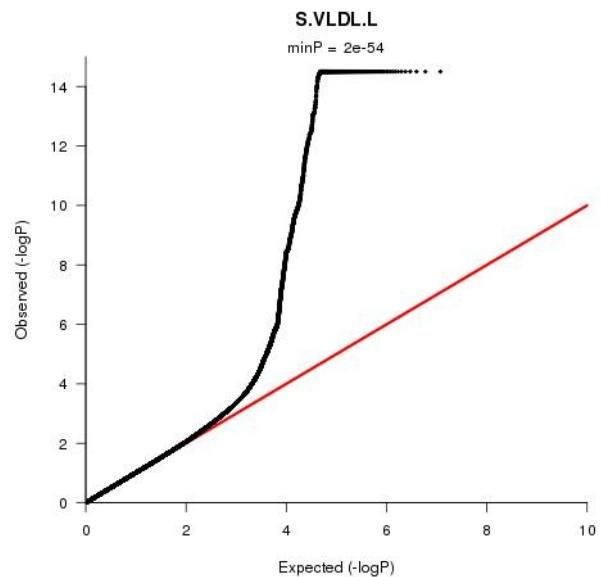
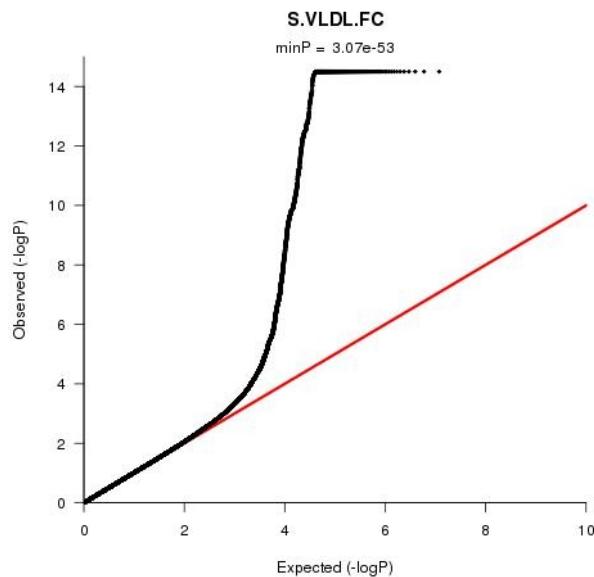


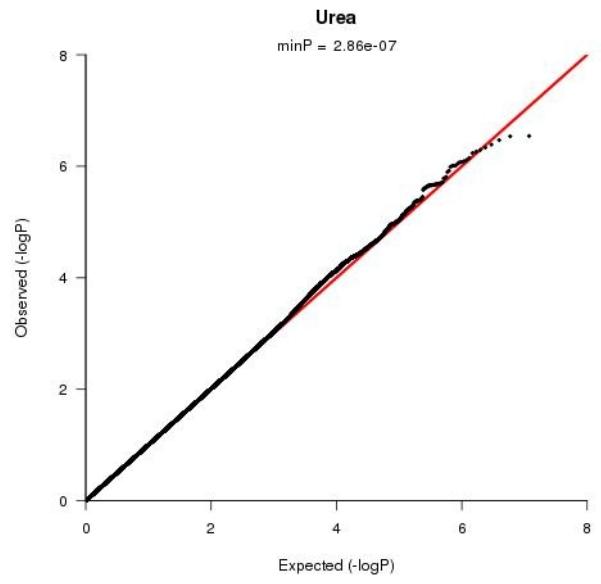
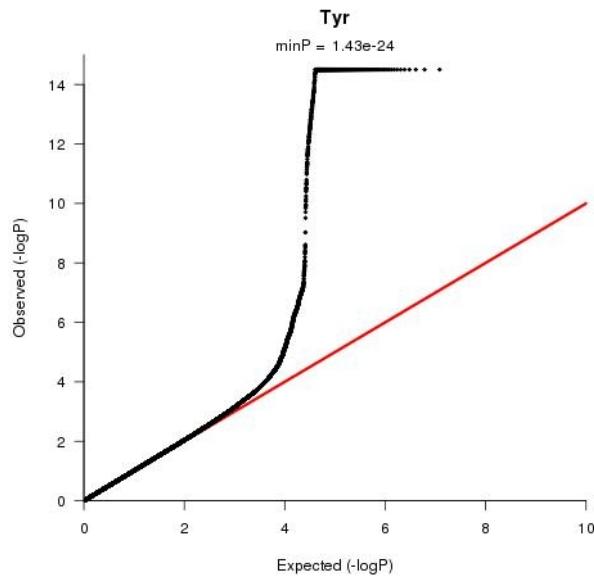
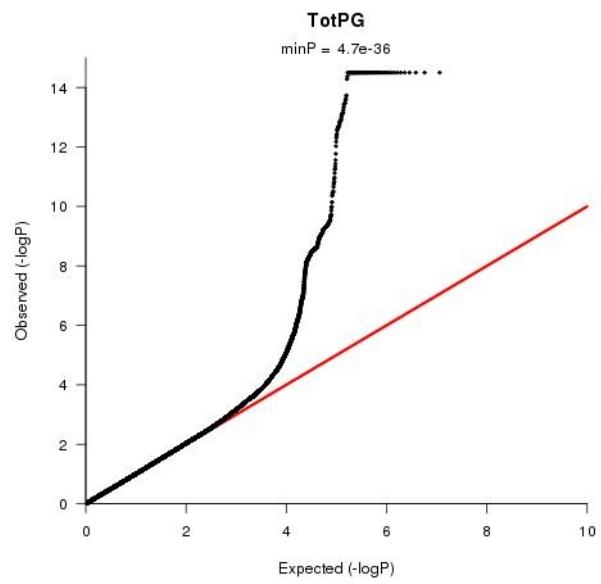
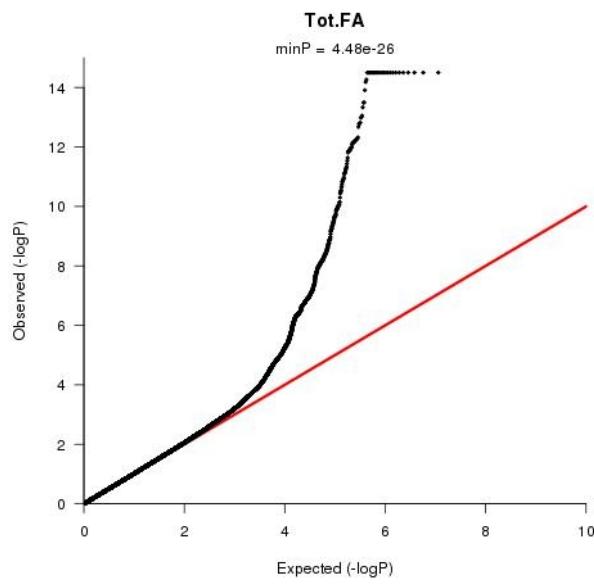
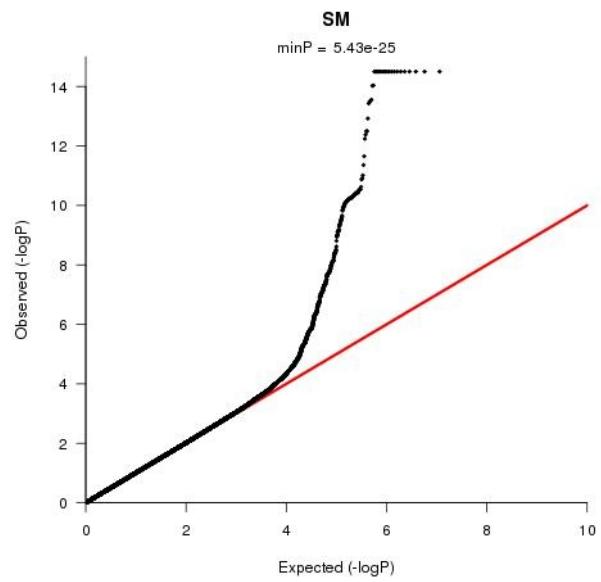
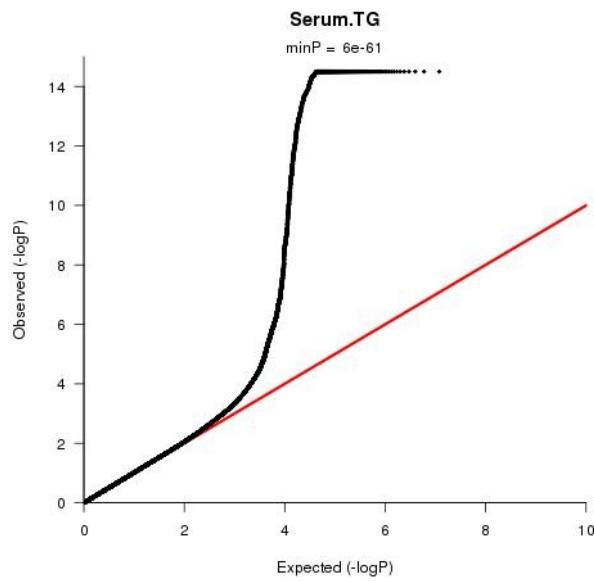


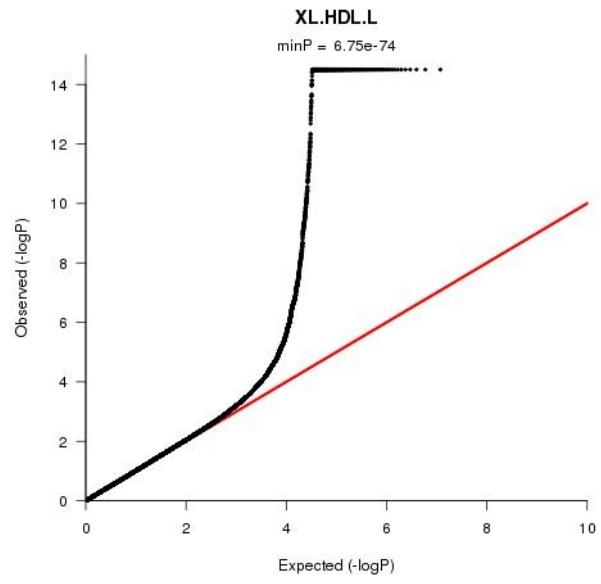
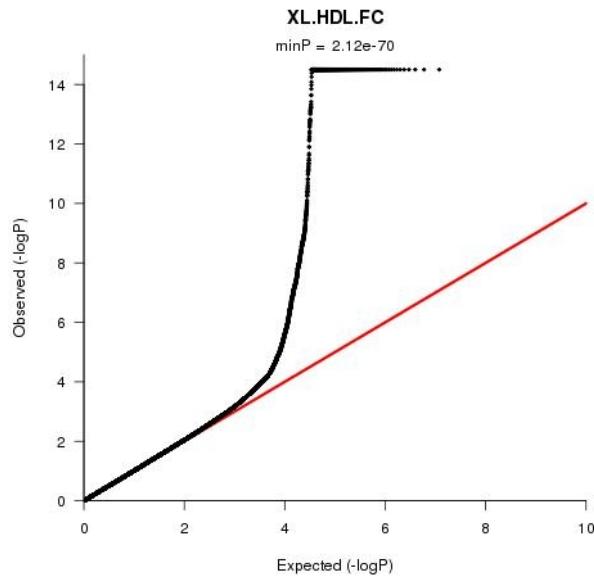
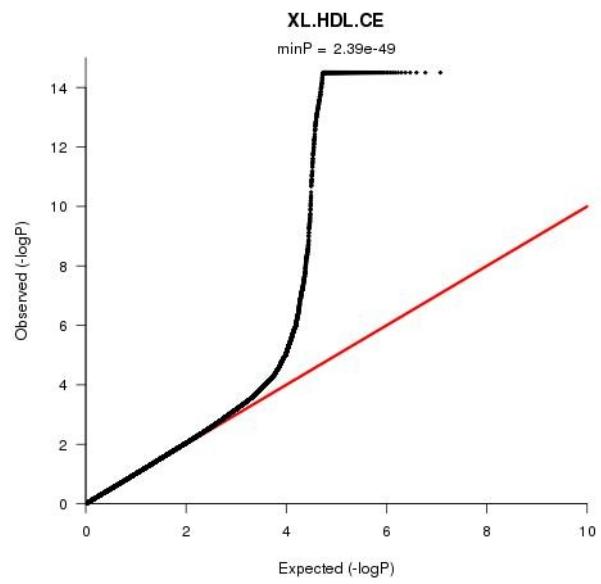
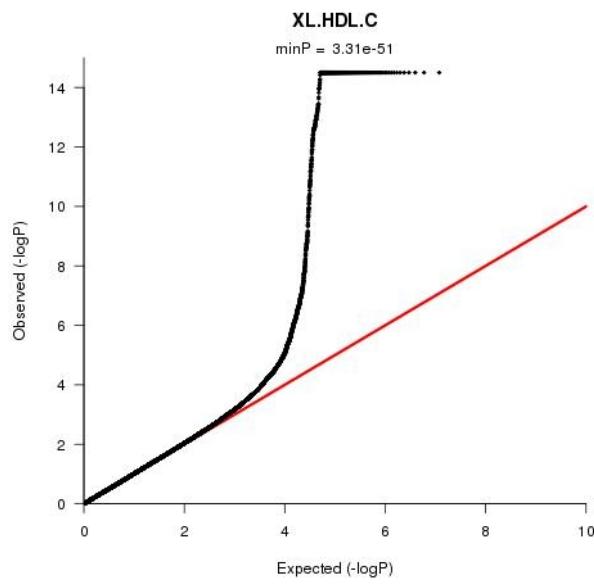
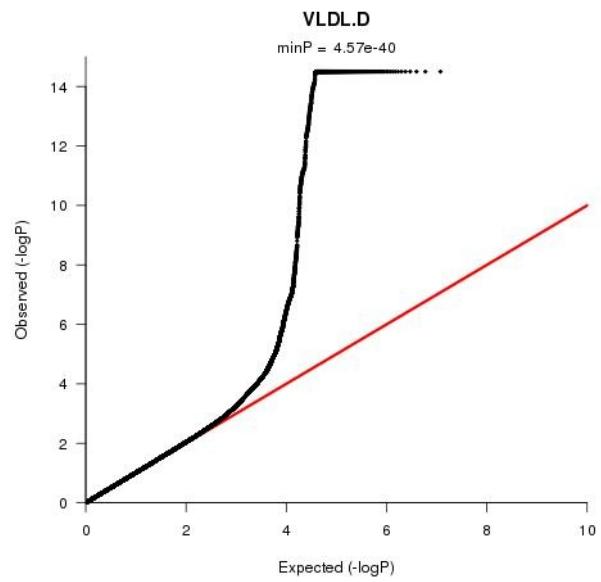
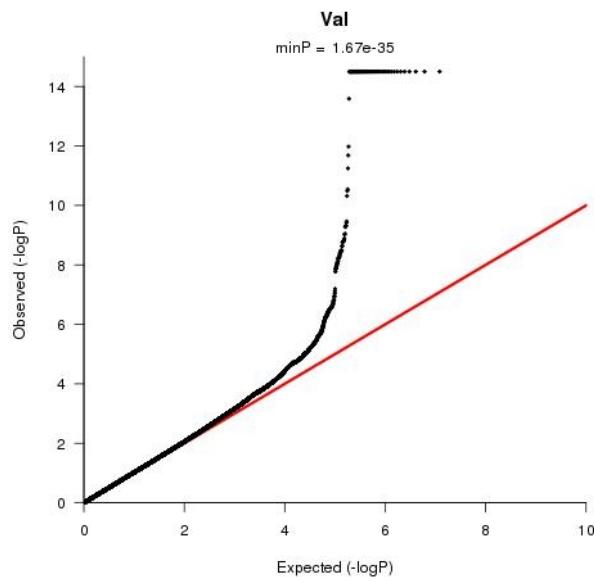


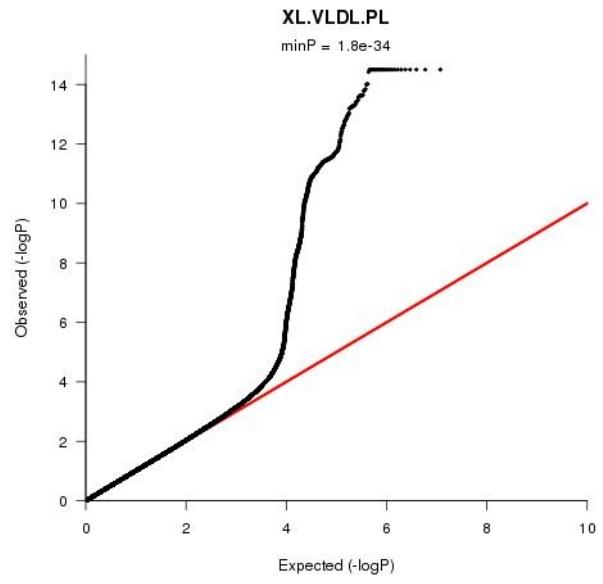
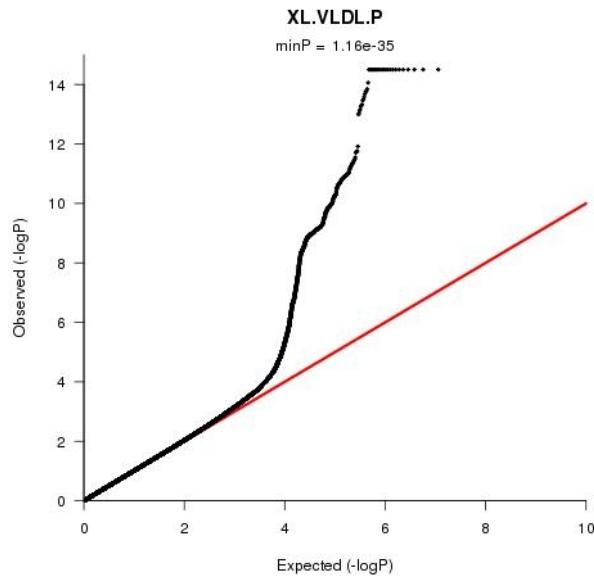
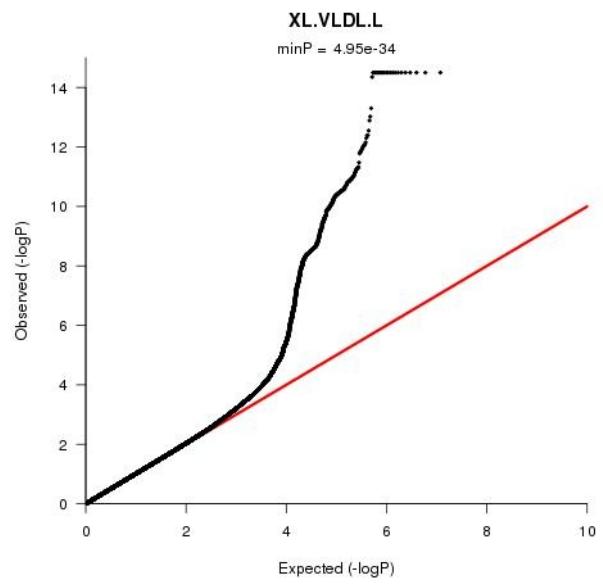
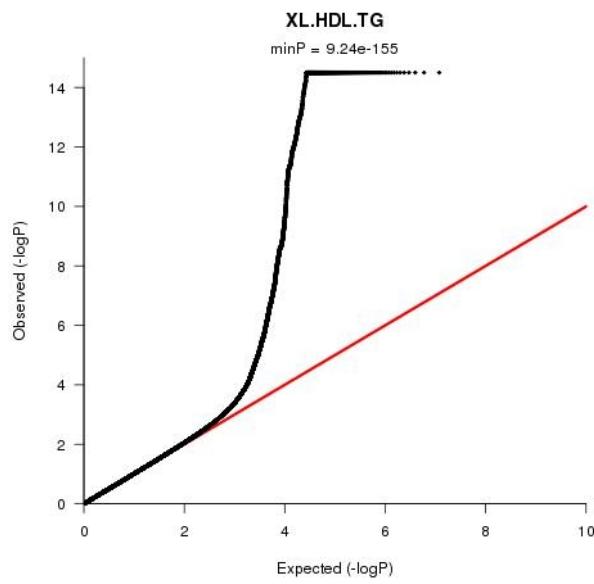
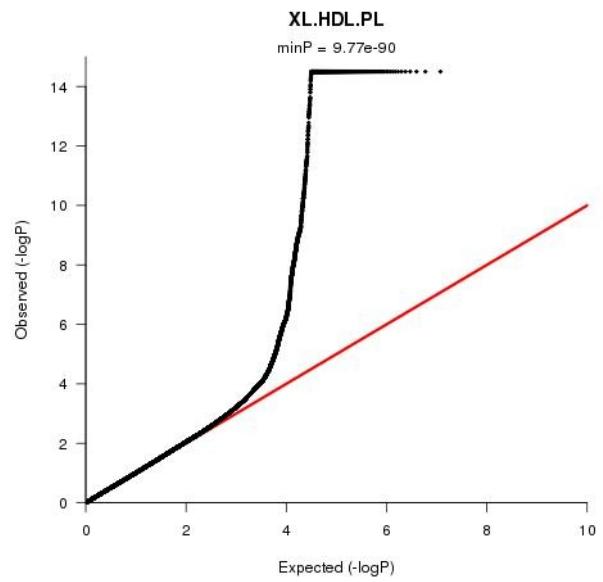
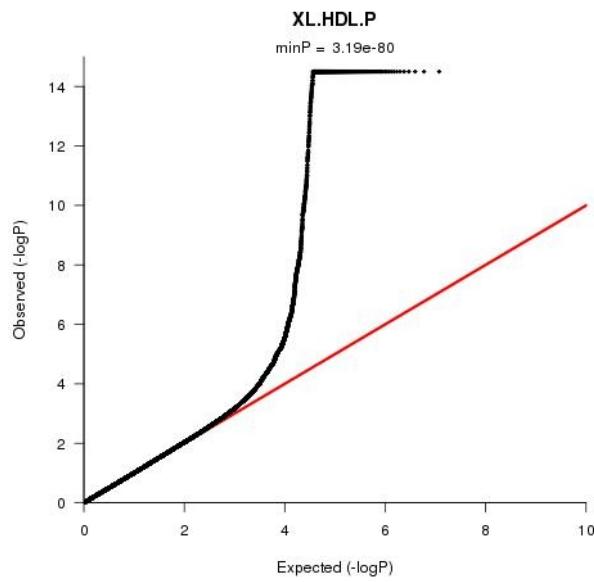


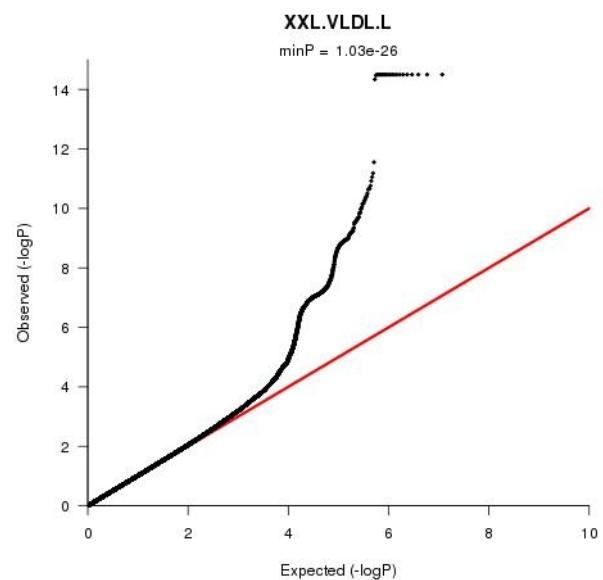
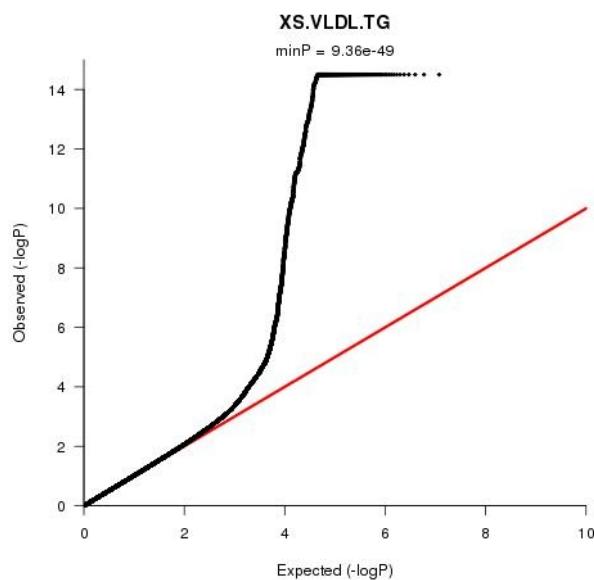
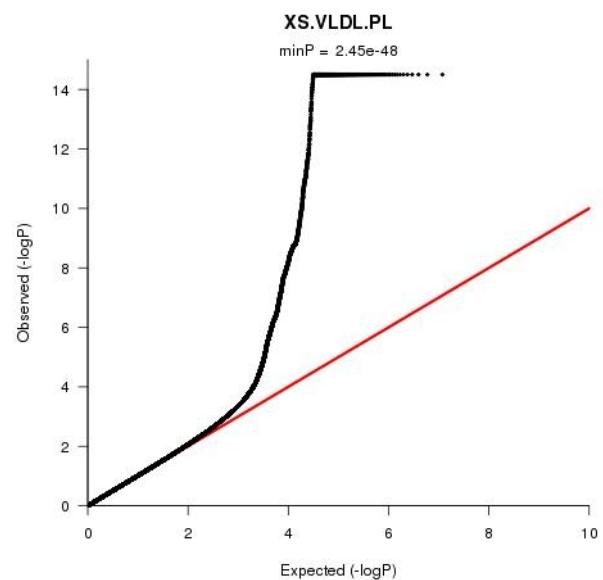
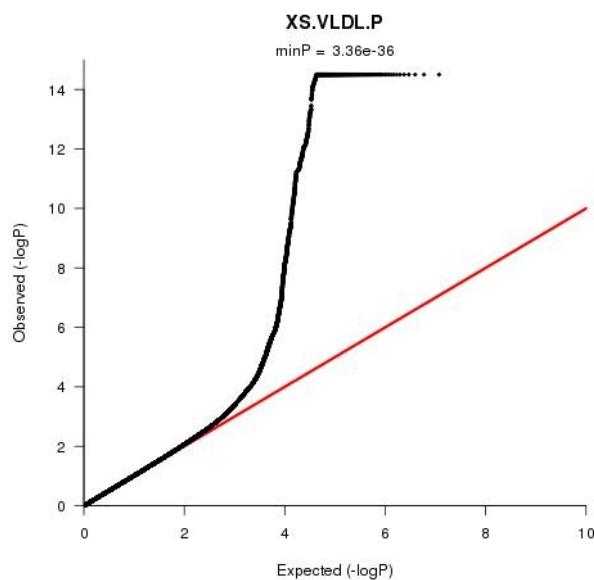
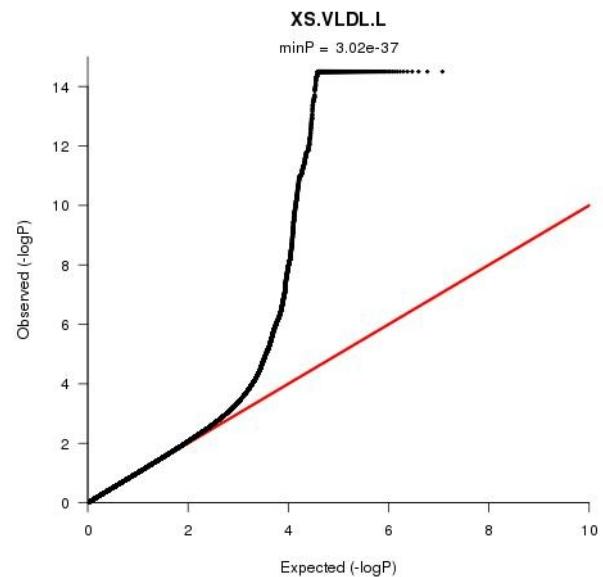
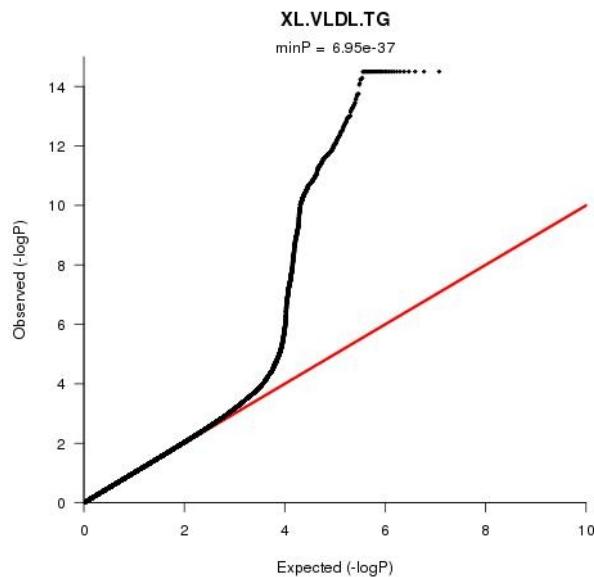


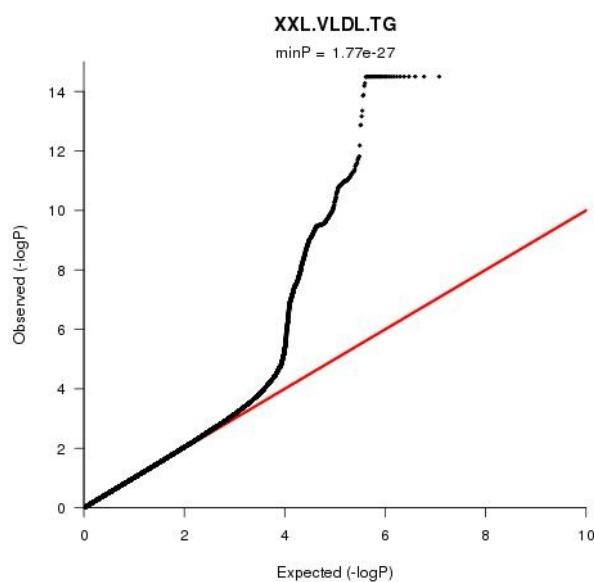
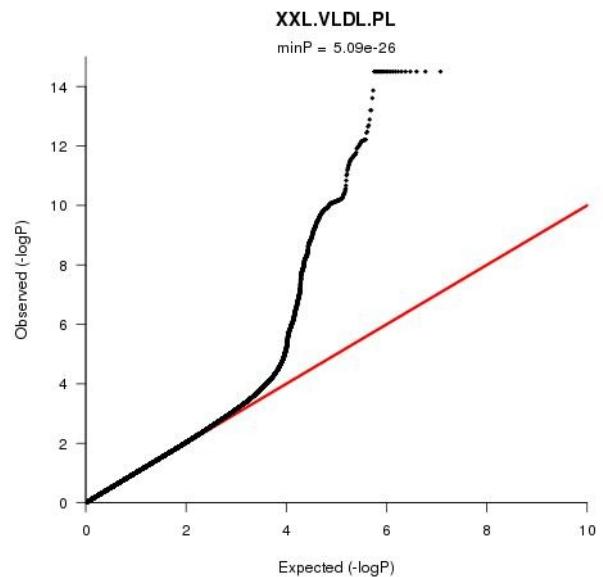
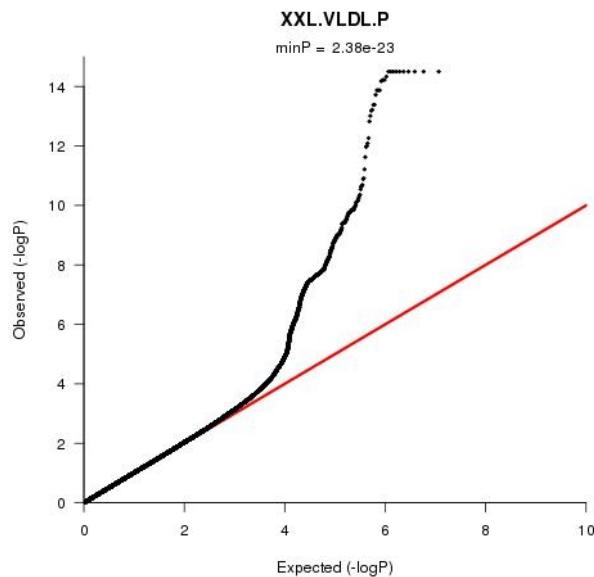












Supplementary Table 1. Metabolite abbreviations and names, number of individuals analyzed and proportion of variance explained.

Abbreviation	Name	N	Proportion of variance explained
AcAce	Acetoacetate	19262	0.225 %
Ace	Acetate	24748	NA
Ala	Alanine	24796	1.326 %
Alb	Albumin	18960	0.984 %
ApoA1	ApoA1	20687	5.005 %
ApoB	ApoB	20690	8.609 %
Bis.DB.ratio	Ratio of bis-allylic bonds to double bonds in lipids	13524	7.392 %
Bis.FA.ratio	Ratio of bis-allylic bonds to total fatty acids in lipids	13171	6.324 %
bOHBut	3-hydroxybutyrate	24154	NA
CH2.in.FA	CH2 groups in fatty acids	19021	2.697 %
CH2.DB.ratio	CH2 groups to double bonds ratio	13532	5.122 %
Cit	Citrate	24770	1.197 %
Crea	Creatinine	24810	0.930 %
DB.in.FA	Double bonds in fatty acids	15728	3.066 %
DHA	22:6, docosahexaenoic acid (DHA)	13499	1.677 %
Est.C	Esterified cholesterol	13497	7.582 %
FALen	Fatty acid length	13476	NA
FAw3	Omega-3 fatty acids	13544	2.873 %
FAw6	Omega-6 fatty acids	13506	5.036 %
FAw79S	Omega-7 and -9 and saturated fatty acids	13506	2.500 %
Free.C	Free cholesterol	13497	8.347 %
Glc	Glucose	24679	1.182 %
Gln	Glutamine	24462	1.772 %
Glol	Glycerol	20235	NA
Gly	Glycine	18734	12.488 %
Gp	Glycoprotein acetyls, mainly a1-acid glycoprotein	19270	2.410 %
HDL.C	Total cholesterol in HDL	21555	6.115 %
HDL.D	HDL diameter	19273	10.323 %
His	Histidine	19244	1.108 %
IDL.C	Total cholesterol in IDL	19273	10.892 %
IDL.FC	Free cholesterol in IDL	21559	10.894 %
IDL.L	Total lipids in IDL	19273	11.267 %
IDL.P	Concentration of IDL particles	19273	11.173 %
IDL.PL	Phospholipids in IDL	21559	11.205 %
IDL.TG	Triglycerides in IDL	19273	9.943 %
Ile	Isoleucine	24776	0.557 %
L.HDL.C	Total cholesterol in large HDL	21558	8.227 %
L.HDL.CE	Cholesterol esters in large HDL	19273	8.386 %
L.HDL.FC	Free cholesterol in large HDL	21559	7.686 %
L.HDL.L	Total lipids in large HDL	19273	8.425 %
L.HDL.P	Concentration of large HDL particles	19273	8.436 %
L.HDL.PL	Phospholipids in large HDL	19273	8.030 %
L.LDL.C	Total cholesterol in large LDL	21552	10.442 %
L.LDL.CE	Cholesterol esters in large LDL	19273	10.559 %
L.LDL.FC	Free cholesterol in large LDL	21555	11.078 %
L.LDL.L	Total lipids in large LDL	19273	10.715 %
L.LDL.P	Concentration of large LDL particles	19273	11.131 %
L.LDL.PL	Phospholipids in large LDL	21550	10.506 %
L.VLDL.C	Total cholesterol in large VLDL	21235	3.239 %
L.VLDL.CE	Cholesterol esters in large VLDL	18960	3.391 %
L.VLDL.FC	Free cholesterol in large VLDL	21238	2.907 %
L.VLDL.L	Total lipids in large VLDL	18960	2.441 %
L.VLDL.P	Concentration of large VLDL particles	18960	2.841 %
L.VLDL.PL	Phospholipids in large VLDL	21239	2.393 %
L.VLDL.TG	Triglycerides in large VLDL	21239	2.586 %
LA	18:2, linoleic acid (LA)	13527	6.584 %
Lac	Lactate	24871	NA
LDL.C	Total cholesterol in LDL	21559	10.720 %
LDL.D	LDL diameter	19273	3.124 %
Leu	Leucine	24728	0.797 %
M.HDL.C	Total cholesterol in medium HDL	21558	2.378 %
M.HDL.CE	Cholesterol esters in medium HDL	19273	2.066 %
M.HDL.FC	Free cholesterol in medium HDL	21559	2.380 %
M.HDL.L	Total lipids in medium HDL	19273	1.764 %
M.HDL.P	Concentration of medium HDL particles	19273	1.537 %

M.HDL.PL	Phospholipids in medium HDL	21558	1.901 %
M.LDL.C	Total cholesterol in medium LDL	21559	9.903 %
M.LDL.CE	Cholesterol esters in medium LDL	19273	10.090 %
M.LDL.L	Total lipids in medium LDL	19273	9.933 %
M.LDL.P	Concentration of medium LDL particles	19273	9.708 %
M.LDL.PL	Phospholipids in medium LDL	21558	8.684 %
M.VLDL.C	Total cholesterol in medium VLDL	21551	5.119 %
M.VLDL.CE	Cholesterol esters in medium VLDL	19273	5.926 %
M.VLDL.FC	Free cholesterol in medium VLDL	21240	4.285 %
M.VLDL.L	Total lipids in medium VLDL	19273	3.649 %
M.VLDL.P	Concentration of medium VLDL particles	19273	3.387 %
M.VLDL.PL	Phospholipids in medium VLDL	21240	4.283 %
M.VLDL.TG	Triglycerides in medium VLDL	21241	3.190 %
MUFA	Mono-unsaturated fatty acids	13535	2.067 %
otPUFA	Other polyunsaturated fatty acids than 18:2	13549	10.162 %
PC	Phosphatidylcholine and other cholines	13542	4.258 %
Phe	Phenylalanine	22663	1.105 %
Pyr	Pyruvate	24756	0.523 %
S.HDL.L	Total lipids in small HDL	19273	3.063 %
S.HDL.P	Concentration of small HDL particles	19273	2.494 %
S.HDL.TG	Triglycerides in small HDL	21558	4.242 %
S.LDL.C	Total cholesterol in small LDL	21556	8.964 %
S.LDL.L	Total lipids in small LDL	19273	8.689 %
S.LDL.P	Concentration of small LDL particles	19273	8.504 %
S.VLDL.C	Total cholesterol in small VLDL	21557	6.716 %
S.VLDL.FC	Free cholesterol in small VLDL	21559	6.426 %
S.VLDL.L	Total lipids in small VLDL	19273	6.258 %
S.VLDL.P	Concentration of small VLDL particles	19273	5.495 %
S.VLDL.PL	Phospholipids in small VLDL	21551	6.322 %
S.VLDL.TG	Triglycerides in small VLDL	21558	4.910 %
Serum.C	Serum total cholesterol	21491	8.671 %
Serum.TG	Serum total triglycerides	21545	4.331 %
SM	Sphingomyelins	13476	3.918 %
Tot.FA	Total fatty acids	13505	3.346 %
TotPG	Total phosphoglycerides	13519	4.242 %
Tyr	Tyrosine	24925	0.444 %
Urea	Urea	18814	NA
Val	Valine	24900	1.008 %
VLDL.D	VLDL diameter	19273	4.652 %
XL.HDL.C	Total cholesterol in very large HDL	21540	5.051 %
XL.HDL.CE	Cholesterol esters in very large HDL	19273	5.174 %
XL.HDL.FC	Free cholesterol in very large HDL	21542	6.875 %
XL.HDL.L	Total lipids in very large HDL	19273	7.718 %
XL.HDL.P	Concentration of very large HDL particles	19273	7.258 %
XL.HDL.PL	Phospholipids in very large HDL	19273	9.376 %
XL.HDL.TG	Triglycerides in very large HDL	21536	10.760 %
XL.VLDL.L	Total lipids in very large VLDL	19273	2.047 %
XL.VLDL.P	Concentration of very large VLDL particles	18960	2.009 %
XL.VLDL.PL	Phospholipids in very large VLDL	21237	2.109 %
XL.VLDL.TG	Triglycerides in very large VLDL	21548	2.409 %
XS.VLDL.L	Total lipids in very small VLDL	19273	9.862 %
XS.VLDL.P	Concentration of very small VLDL particles	19273	9.447 %
XS.VLDL.PL	Phospholipids in very small VLDL	19273	10.552 %
XS.VLDL.TG	Triglycerides in very small VLDL	19273	7.125 %
XXL.VLDL.L	Total lipids in chylomicrons and extremely large VLDL	18960	1.567 %
XXL.VLDL.P	Concentration of chylomicrons and extremely large VLDL particles	18960	1.401 %
XXL.VLDL.PL	Phospholipids in chylomicrons and extremely large VLDL	21542	1.572 %
XXL.VLDL.TG	Triglycerides in chylomicrons and extremely large VLDL	21540	1.710 %

NA Those metabolic measures where no genome-wide associations were found

Supplementary Table 2. Genotyping details, genotype quality control and metabolomics measurement details.

Study	Short name	COROGENE	EGCUT	ERF	FR97	GenMets	HBCS	KORA-F4	LLS	NFBC 1966	NTR	Predict CVD	Prote	Twins	YFS
	<b>Full name</b>	Genetic Predisposition of Coronary Heart Disease in Patients Verified with Coronary Angiogram	Estonian Genome Center, University of Tartu	Erasmus Rucphen Family Study	FINRISK 1997	Genetics of metabolic syndrome	Helsinki birth cohort study	kooperative gesundheitsforschung in der region augsburg	Leiden Longevity Study	Norther Finland Birth Cohort Study 1966	Netherlands Twin Register				The Cardiovascular Risk in Young Finns Study
Study design		Case-control study, only controls used	Population-based cohort	Family-based	Population-based cohort	Case-control study	Birth cohort	Population based cohort	Family-based	Birth cohort	Population based twin study	Cohort study	Population-based	Population based twin study	Follow up study in children
Ethnicity		Finnish	Estonian	Dutch	Finnish	Finnish	Finnish	German	Dutch	Finnish	Dutch	Finnish	Estonian	Finnish	Finnish
Total genotyped sample size (N)		4122	7734	2917	5288	2118	1676	1996	2415	5104	14003	1881	1397	664	2443
Sample QC	<b>Call rate*</b>	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	SNP missing rate < 10% in subsets	≥ 95%	≥ 95%	≥ 95%	≥ 95%
	<b>other exclusions</b>	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	Additional exclusions: IBD > 0.1, sex mismatches, duplicates	Mendelian inconsistencies	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	Relatedness, Mendelian inconsistencies	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	Heterozygosity F < -0.10 or > 0.10; mismatch between X-chromosomal genotype and phenotypic sex; failure of expected IBS / IBD relations between samples; Mendel errors > 2%; different DNA for the same sample or between confirmed MZ co-twins over merged SNP genotype-imputed subsets	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	Additional exclusions: IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution	IBD > 0.1, sex mismatches, duplicates, heterozygosity outliers by eye from distribution
Assessment NMR	<b>Machine</b>	500 MHz Bruker AVANCE III spectrometer	500 MHz Bruker AVANCE III spectrometer	600 MHz Bruker AVANCE II spectrometer	500 MHz Bruker AVANCE III spectrometer	500 MHz Bruker AVANCE III spectrometer	500 MHz Bruker AVANCE III spectrometer	500 MHz Bruker AVANCE III spectrometer	600 MHz Bruker AVANCE II spectrometer	500 MHz Bruker AVANCE III spectrometer	600 MHz Bruker AVANCE II spectrometer	500 MHz Bruker AVANCE III spectrometer	500 MHz Bruker AVANCE III spectrometer	500 MHz Bruker AVANCE III spectrometer	500 MHz Bruker AVANCE III spectrometer
	<b>Sample material</b>	Serum	Plasma	Serum	Serum	Serum	Serum	Serum	Plasma	Serum	Plasma	Serum	Plasma	Serum	Serum
	<b>Processing</b>	In-house algorithm	In-house algorithm	in-house deconvolution algorithm	In-house algorithm	In-house algorithm	In-house algorithm	In-house algorithm	In-house deconvolution algorithm	In-house algorithm	In-house deconvolution algorithm	In-house algorithm	In-house algorithm	In-house algorithm	In-house algorithm
Genotyping		illumina 610k	Illumina HumanOmniExpress, Illumina HumanCNV370	Illumina 318K, Illumina 370K and Affymetrix 250K	Illumina core-exome	Illumina 610k	Illumina 670k	Affmetrix 6.0	Illumina 660W, Illumina OmniExpress	Illumina 370k	Perlegen-Affymetrix 500K; Illumina 660K; Illumina 370K; Affymetrix 6.0 907K; Illumina Omni 1M	Illumina omni express	Illumina HumanOmniExpress, Illumina HumanCNV370	Illumina 670k	Illumina 670k
	<b>Genotype calling algorithm</b>	Illuminus	Genome Studio	Illumina Bead Studio, BRLMM	Illuminus	Illuminus	Birdseed2	GenomeStudio	Illuminus	Genotyper, Beadstudio, Birdseed	Illuminus	Genome Studio	Illuminus	Illuminus	Illuminus
	<b>MAF</b>	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%	≥ 1%
	<b>Call rate*</b>	≥ 95%	≥ 95%	≥ 98%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%	≥ 95%
	<b>p for HWE</b>	> 10 <sup>-6</sup>	> 10 <sup>-6</sup>	> 10 <sup>-6</sup>	> 10 <sup>-6</sup>	> 10 <sup>-6</sup>	> 10 <sup>-6</sup>	> 10 <sup>-7</sup>	> 10 <sup>-4</sup>	> 10 <sup>-6</sup>	> 1E-05	> 10 <sup>-6</sup>	> 10 <sup>-6</sup>	> 10 <sup>-6</sup>	> 10 <sup>-6</sup>
	<b>SNPs that met QC criteria</b>	554988	38.02 M in final list	659174	273113	555388	546814	692637	522801-645804	335118	289598-1139672	645088	38.02 M in final list	552894	546677
	<b>Imputation software</b>	IMPUTE2	IMPUTE2	Mach/Minimac	IMPUTE2	IMPUTE2	IMPUTE2	IMPUTE	IMPUTE version 2.2 (beta)	IMPUTE2	Mach phasing, Minimac Imputation	IMPUTE2	IMPUTE2	IMPUTE2	IMPUTE2
Association analyses	<b>Analysis software</b>	SNPTest 2.4.1	SNPTest 2.4.1	ProbABEL	SNPTest 2.4.1	SNPTest 2.4.1	SNPTest 2.4.1	SNPTTEST	QT-assoc	SNPTest 2.4.1	PLINK 1.07	SNPTest 2.4.1	SNPTest 2.4.1	GEMMA	SNPTest 2.4.1
	<b>Cohort specific adjustments</b>			Family structure					Familial relations		Relatedness; genotyping sample			Family structure	

Supplementary Table 3. Conditional variants

Round 1: conditioned with the lead variant in the locus given in Tables 2 and 3

Phenotype	rsid	Chromosome	Position	Beta	SE	p-value
L.LDL.FC	rs11591147	1	55505647	-0.5335034	0.03230012	2.76E-61
L.LDL.FC	rs13011615	2	21274167	-0.1247424	0.01401459	5.54E-19
Val	rs34226052	4	89163523	-0.0967459	0.01401841	5.15E-12
HDL.C	rs75835816	8	19885513	-0.279236	0.0368885	3.74E-14
His	rs2072510	12	96403199	-0.0689075	0.01048573	4.98E-11
XL.HDL.TG	rs261334	15	58726744	-0.2686627	0.0112669	1.13E-125
HDL.C	rs11076175	16	57006378	-0.1511545	0.0135707	8.17E-29
otPUFA	rs190121281	19	19252779	-0.3228719	0.03264413	4.57E-23
S.HDL.L	rs6032614	20	44609595	-0.1017956	0.01388914	2.32E-13

Round 2: conditioned with the lead variant in the locus given in Tables 2 and 3 as well as round 1 variant above

Phenotype	rsid	Chromosome	Position	Beta	SE	p-value
XS.VLDL.TG	chr2.211488	2	21148888	-0.1263201	0.01389021	9.53E-20
XL.HDL.TG	rs56050415	15	58574966	-0.1198553	0.01366415	1.76E-18

Round 3: conditioned with the lead variant in the locus given in Tables 2 and 3 as well as round 1 and 2 variants above

Phenotype	rsid	Chromosome	Position	Beta	SE	p-value
HDL.D	chr15.58927	15	58927492	-0.1706685	0.02565108	2.86E-11

**Supplementary Table 4.** Previously identified loci and SNPs significant in this GWAS. Beta refers to one copy addition of the effect allele in standard deviation units. Metabolic traits are grouped coarsely by phenotypic similarity.

Lipoprotein measures													
Trait	Variant identifier	Chr	Position	ea/nea	Eaf	Beta	SE	P-value	Q P-value	N samples	Candidate gene	eQTL	Function
I.LDL.FC	rs191448950	1	55584844	A/G	0.02	-0.5	0.03	4.8×10 <sup>-56</sup>	0.31	21549	PCSK9	—	Intron;USP24
I.LDL.FC	chr1:10981815	1	109818158	C/T	0.78	0.14	0.01	1.0×10 <sup>-21</sup>	0.15	19265	SORT1	CELSR2	S'UTR/TBFS;CELSR2
XL.HDL.PL	rs450368	1	161190250	C/T	0.53	-0.07	0.01	3.2×10 <sup>-11</sup>	0.34	19270	USF1	—	—
M.HDL.FC	rs590820	1	230309619	G/A	0.44	-0.06	0.01	1.6×10 <sup>-7</sup>	0.2	21553	GALNT2	—	Intron
IDL.FC	rs952275	2	212139970	G/T	0.43	0.11	0.01	1.5×10 <sup>-21</sup>	0.18	21556	APOB	—	TFBS
LDL.C	rs6756629#	2	44065090	A/G	0.08	-0.14	0.02	1.2×10 <sup>-14</sup>	0.15	21559	ABCGB5	—	Missense
S.LDL.C	rs4703667	5	74613906	C/G	0.39	0.09	0.01	2.6×10 <sup>-18</sup>	0.76	21551	HMGCR	—	—
VLDL.D	rs10455872	6	161010118	G/A	0.04	-0.2	0.03	1.3×10 <sup>-12</sup>	0.03	19272	LP4	—	Intron
I.LDL.PL	rs73066442	7	21592973	G/A	0.25	0.08	0.01	9.7×10 <sup>-11</sup>	0.71	21546	DNAHII	DNAHII	Intron
VLDL.D	rs17145750	7	73026378	T/C	0.15	-0.13	0.01	4.3×10 <sup>-20</sup>	0.5	19268	MLXIPL	MLXIPL	Intron
S.VLDL.TG	rs115849089	8	19912370	A/G	0.11	-0.18	0.02	1.6×10 <sup>-23</sup>	0.4	21554	LPL	—	—
XS.VLDL.TG	rs2954029	8	126490972	T/A	0.48	-0.09	0.01	2.8×10 <sup>-16</sup>	0.051	19270	TRIB1	—	—
XL.HDL.FC	rs686030	9	15304782	A/C	0.87	0.09	0.01	2.0×10 <sup>-17</sup>	0.91	21541	ITC39B	—	TFBS
XL.HDL.C	rs2575876	9	107665739	A/G	0.19	-0.1	0.01	2.7×10 <sup>-15</sup>	0.16	21540	ABCA1	—	TFBS/Intron
IDL.FC	chr9:13614970	9	136149709	A/A/C	0.19	0.09	0.01	1.0×10 <sup>-11</sup>	0.068	19269	APOB	—	Intron
S.VLDL.TG	rs964184	11	116648917	C/G	0.86	-0.24	0.01	7.6×10 <sup>-66</sup>	4.0×10 <sup>-10</sup>	21556	ZNF259	—	TFBS
HDL.D	rs67053123	12	125353810	A/T	0.15	0.11	0.01	9.8×10 <sup>-13</sup>	0.67	19270	SCARB1	—	TFBS
XL.HDL.TG	rs1532085	15	58683366	G/A	0.6	-0.26	0.01	9.2×10 <sup>-13</sup>	0.59	21536	LIPC	—	—
HDL.C	rs247617	16	56990716	A/C	0.3	0.23	0.01	5.8×10 <sup>-97</sup>	0.26	21550	CETP	—	TFBS
HDL.D	rs6507939	18	471176261	C/A	0.84	0.11	0.01	8.3×10 <sup>-16</sup>	0.46	19268	LIPG	—	—
VLDL.D	rs116843064#	19	8429323	A/G	0.03	-0.22	0.03	2.8×10 <sup>-19</sup>	0.016	19268	4NGPTL4	—	Missense
I.LDL.FC	rs142130958	19	11190652	A/G	0.11	-0.24	0.02	6.5×10 <sup>-11</sup>	0.11	21554	LDLR	—	—
M.VLDL.C	rs72999033	19	19366632	T/C	0.06	-0.18	0.02	1.0×10 <sup>-16</sup>	0.6	21548	CILP	—	TFBS/UTR;HAPLN4
I.LDL.FC	rs7412	19	45412079	T/C	0.06	-0.59	0.03	2.0×10 <sup>-13</sup>	0.15	18263	APOE	—	Missense
XS.VLDL.PL	rs1883711	20	39179822	C/G	0.06	0.17	0.02	2.3×10 <sup>-11</sup>	0.16	19268	MABP	—	TFBS
S.HDL.P	rs6073958	20	44551855	C/T	0.2	0.2	0.01	2.6×10 <sup>-53</sup>	0.35	19269	PLTP	—	—
Lipid measures													
Serum.TG	rs1168041	1	62960250	C/T	0.72	0.1	0.01	3.4×10 <sup>-17</sup>	0.62	21539	4NGPTL3	—	Intron;DOCK7
oPUFAs	rs174547	11	161570783	C/T	0.4	-0.4	0.01	6.5×10 <sup>-21</sup>	3.8×10 <sup>-10</sup>	13547	FADS2	—	Intron
Bis.DB.ratio	rs11646401	16	15172118	C/T	0.28	0.1	0.01	7.7×10 <sup>-13</sup>	0.16	13517	PDXCI	NPIP45	Intron
Sphingomyelin	rs181807530	17	4774814	G/C	0.04	-0.37	0.04	5.4×10 <sup>-28</sup>	0.25	11439	MINK1	—	Intron
Small molecule and protein measures													
Alanine	rs1260326#	2	27730940	C/T	0.64	-0.1	0.01	7.4×10 <sup>-26</sup>	0.19	22569	GCKR	—	Missense
Valine	rs10211524	2	65208074	A/G	0.41	0.09	0.01	5.2×10 <sup>-20</sup>	0.1	24898	SLC1A4	—	—
Glucose	rs560887	2	169763148	C/T	0.7	0.12	0.01	3.4×10 <sup>-32</sup>	0.0016	22452	G6PC2	—	Intron
Glycine	rs1047891#	2	211540507	A/C	0.33	0.49	0.01	<1×10 <sup>-30</sup>	5.3×10 <sup>-10</sup>	18730	CPSI	—	Missense
Albumin	rs184650103	4	74850649	T/C	0.01	-0.6	0.05	1.2×10 <sup>-30</sup>	0.23	17461	ALB	—	TFBS
Creatinine	rs1986734	4	77420787	T/C	0.56	0.06	0.01	3.0×10 <sup>-10</sup>	0.62	21918	SHROOM3	—	TFBS
Valine	rs9637599	4	89206230	C/A	0.47	0.11	0.01	1.7×10 <sup>-31</sup>	0.12	24899	PPMIK	lncRNA	TFBS
Histidine	rs3733404#	4	187158034	A/G	0.57	0.08	0.01	9.1×10 <sup>-19</sup>	0.013	19241	KLKB1	—	Missense
Citrate	rs2921604	5	14867948	C/T	0.44	0.1	0.01	3.7×10 <sup>-29</sup>	0.43	24765	4NKH	ANHK	Intron
Phenylalanine	rs2731672	5	176842474	C/T	0.74	0.09	0.01	3.9×10 <sup>-16</sup>	4.0×10 <sup>-10</sup>	20436	F12	LMAN2	—
Tyrosine	rs14399	6	111543944	A/C	0.39	-0.1	0.01	1.4×10 <sup>-24</sup>	0.086	24918	SLC16A10	—	3'UTR
Glucose	rs878521	7	44255643	A/G	0.22	0.09	0.01	8.0×10 <sup>-14</sup>	0.016	21020	GCK	—	TFBS
Creatinine	rs10265221	7	151414329	C/T	0.24	0.07	0.01	7.5×10 <sup>-11</sup>	0.32	24805	PRKAG2	—	Intron
Glycine	rs2169387	8	98118395	G/A	0.87	-0.13	0.02	1.3×10 <sup>-18</sup>	0.99	18729	PPPIR3B	Processed pseudogene	—
Glutamine	rs7078003	10	99359412	T/C	0.2	0.07	0.01	3.0×10 <sup>-10</sup>	0.014	24459	HOGAI	—	Intron/3'UTR
Glucose	rs10466351	11	92697981	T/C	0.38	0.07	0.01	3.3×10 <sup>-13</sup>	0.042	24676	MTNR1B	—	—
Glutamine	rs2657879#	12	56865338	G/A	0.18	-0.22	0.01	3.3×10 <sup>-79</sup>	0.11	24458	GLS2	—	Missense
Phenylalanine	rs1718309	12	103242396	G/A	0.6	-0.08	0.01	2.5×10 <sup>-15</sup>	0.07	22660	PAH	—	Intron
Glycoprotein acetyls	rs28929474#	14	94844947	T/C	0.02	-0.29	0.04	2.7×10 <sup>-13</sup>	0.17	19267	SERPINAI	—	Missense
Creatinine	rs61524473	15	45646283	C/T	0.26	0.09	0.01	1.2×10 <sup>-15</sup>	0.51	24805	GATM	GATM, SPATSL1, LOC101928414	—
Glycoprotein acetyls	rs77303550	16	72079657	T/C	0.19	0.15	0.01	4.0×10 <sup>-39</sup>	0.86	19267	HP	—	Intron;TXNL4B
Citrate	rs172642	17	5945398	C/A	0.48	0.08	0.01	4.8×10 <sup>-18</sup>	0.5	24767	SLC13A5	DVL2	Intron
Creatinine	rs2079742	17	59465697	C/T	0.18	0.1	0.01	8.6×10 <sup>-14</sup>	0.46	22575	BCAS3	—	Intron
Citrate	rs2040771	22	19161935	T/C	0.48	-0.09	0.01	1.3×10 <sup>-22</sup>	0.36	24769	SLC25A1	SLC25A1	TFBS

In lipoprotein abbreviations, first is given the size category if applicable: XL=extra large, L=large, M=medium, S=small and XS=extra small. Then the lipoprotein particle is given: VLDL=very-low-density lipoprotein particle, IDL=intermediate-density lipoprotein particle, LDL=low-density lipoprotein particles and HDL=high-density lipoprotein particle and finally the lipid measure of the particle: C=total cholesterol, D=the mean diameter of the particle, FC=free cholesterol, L=total lipids, P=particle concentration, PL=phospholipids, TG=triglycerides. In lipids and small molecules: Bis.DB.ratio = Ratio of bis-allylic bonds to double bonds in lipids, oPUFAs= Other polyunsaturated fatty acids than 18:2, ea= effect allele, nea=non-effect allele, Eaf=effect allele frequency, Beta=effect estimate, SE=standard error, eQTL=expression quantitative trait locus from GTEx, # = missense variant, TFBS: transcription factor binding site and Q: Heterogeneity statistics. If the SNP is located in an intron of a different gene than the candidate, then the gene is presented in the Function column after semicolon. Full list of abbreviations can be found from **Supplementary Table 1**

Supplementary Table 5. GTEx eQTLs

Tissue	SNP	Ensembl Gene ID	Gene type	UniProt Gene Name	beta	t-stat	p-value	FDR	pheno	chromosomes	r <sup>2</sup>	position	gene	novelty	function	MAF
Muscle_skeletal	chr1:109818158::1	ENSG000000143126	protein_coding	CELSR2	-0.4642675	-5.624193	1.29E-07	0.000128	LDLFC	1 chr1:109818158	0.99818158	SORT1	old	TFBS_Ins	0.215503	
Heart_Left_Ventricle	rs9637599	ENSG000000246375	lncRNA	101929118	0.9322468	9.5884052	7.38E-14	1.43E-09	Val	4 rs9637599	89206230	PPM1K	old	TFBS	0.46958	
Heart_Left_Ventricle	rs9637599	ENSG000000207480	misc_RNA		0.62357991	5.8398098	2.07E-07	0.00020307	Val	4 rs9637599	89206230	PPM1K	old	TFBS	0.46958	
Thyroid	rs9637599	ENSG000000246375	lncRNA	101929118	0.56862444	9.4538002	3.72E-06	0.00140644	Val	4 rs9637599	89206230	PPM1K	old	TFBS	0.46958	
Artery_Tibial	rs9637599	ENSG000000246375	lncRNA	101929118	0.6217407	6.17461482	1.82E-08	1.33E-05	Val	4 rs9637599	89206230	PPM1K	old	TFBS	0.46958	
Nerve_Tibial	rs9637599	ENSG000000246375	lncRNA	101929118	1.02111658	10.7500861	3.16E-16	6.81E-13	Val	4 rs9637599	89206230	PPM1K	old	TFBS	0.46958	
Nerve_Tibial	rs9637599	ENSG000000207480	misc_RNA		0.57966837	4.98798945	4.58E-06	0.00197816	Val	4 rs9637599	89206230	PPM1K	old	TFBS	0.46958	
Whole_blood	rs2912604	ENSG000000154122	protein_coding	ANKH	-0.2541829	-4.280844	3.51E-05	0.01774489	Cit	5 rs2912604	14867948	ANKH	old	intron	0.444373	
Heart_Left_Ventricle	rs2731672	ENSG000000169223	protein_coding	LMAN2	0.25785696	4.18379027	9.19E-05	0.02537323	Phe	5 rs2731672	176842474	F12	old	NA	0.259219	
Thyroid	rs73066442	ENSG000000105877	protein_coding	DNAH11	-0.8147978	6.2407472	1.70E-08	1.28E-05	LDLPL	7 rs73066442	21592973	DNAH11	old	intron	0.245779	
Heart_Left_Ventricle	rs17145750	ENSG000000099590	protein_coding	MLXIP	0.79120012	3.97243974	0.00018794	0.04036603	VLDL_D	7 rs17145750	73026378	MLXIP	old	intron	0.148692	
Thyroid	rs17145750	ENSG000000099590	protein_coding	MLXIP	1.0881626	6.865265984	1.07E-09	1.21E-06	VLDL_D	7 rs17145750	73026378	MLXIP	old	intron	0.148692	
Skin_sun_exposed_lower_leg	rs17145750	ENSG000000099590	protein_coding	MLXIP	0.77159537	4.99376257	3.75E-06	0.00412804	VLDL_D	7 rs17145750	73026378	MLXIP	old	intron	0.148692	
Thyroid	rs2169387	ENSG000000253697	processed_pseudogene		-0.4845005	-5.5209556	1.12E-06	0.00050933	Gly	8 rs2169387	9181395	PPP1R3B	old	NA	0.13332	
Heart_Left_Ventricle	rs174547	ENSG000000134824	protein_coding	FADS2	0.63735921	5.51966948	7.09E-07	0.00045697	oTPUFA	11 rs174547	61570783	FADS1	old	Intron-FADS1	0.402707	
Heart_Left_Ventricle	rs174547	ENSG000000256443	lncRNA		0.60464182	4.25482232	7.20E-05	0.02318245	oTPUFA	11 rs174547	61570783	FADS1	old	Intron-FADS1	0.402707	
Muscle_skeletal	rs174547	ENSG000000134824	protein_coding	FADS2	0.3522459	4.66225052	8.36E-06	0.00331788	oTPUFA	11 rs174547	61570783	FADS1	old	Intron-FADS1	0.402707	
Nerve_Tibial	rs174547	ENSG000000134824	protein_coding	FADS2	0.49602567	4.93549184	5.58E-06	0.00200845	oTPUFA	11 rs174547	61570783	FADS1	old	Intron-FADS1	0.402707	
Whole_blood	rs174547	ENSG000000134824	protein_coding	FADS2	0.5078831	6.33037234	3.36E-09	6.80E-06	oTPUFA	11 rs174547	61570783	FADS1	old	Intron-FADS1	0.402707	
Heart_Left_Ventricle	rs7954638	ENSG000000139344	protein_coding	AMHD1	0.5876378	4.14013139	0.00010668	0.02577536	His	12 rs7954638	96314795	HAL	new	intron-CCDC38	0.482356	
Thyroid	rs7954638	ENSG000000139344	protein_coding	AMHD1	0.55002097	4.52452761	1.97E-05	0.00640156	His	12 rs7954638	96314795	HAL	new	intron-CCDC38	0.482356	
Artery_Tibial	rs7954638	ENSG000000139344	protein_coding	AMHD1	0.48807644	4.28572368	4.52E-05	0.01595644	His	12 rs7954638	96314795	HAL	new	intron-CCDC38	0.482356	
Muscle_skeletal	rs7954638	ENSG000000139344	protein_coding	AMHD1	0.73453433	7.84045274	2.31E-12	4.58E-09	His	12 rs7954638	96314795	HAL	new	intron-CCDC38	0.482356	
Heart_Left_Ventricle	rs61524473	ENSG000000259520	antisense	LOC101928414	-0.731142	-4.7842257	1.10E-05	0.00531831	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Thyroid	rs61524473	ENSG000000259520	antisense	LOC101928414	-0.853445	-7.8694479	1.11E-11	2.51E-08	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Thyroid	rs61524473	ENSG000000171763	protein_coding	SPATAS5L	0.57982168	4.46326719	2.49E-05	0.00706289	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Artery_Tibial	rs61524473	ENSG000000259520	antisense	LOC101928414	-0.6890699	-6.1617995	1.93E-08	1.33E-05	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Artery_Tibial	rs61524473	ENSG000000259520	processed_pseudogene	SPATAS5L	-0.494388	-4.8112902	5.92E-06	0.00306431	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Artery_Tibial	rs61524473	ENSG000000171763	protein_coding	SPATAS5L	0.45453552	4.62624266	1.23E-05	0.00508763	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Lung	rs61524473	ENSG000000259520	antisense	LOC101928414	-0.6693657	-6.202932	3.02E-08	6.74E-05	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Muscle_skeletal	rs61524473	ENSG000000171766	protein_coding	GATM	0.44907279	5.0064056	1.98E-06	0.00130904	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Muscle_skeletal	rs61524473	ENSG000000259520	antisense	LOC101928414	-0.5024898	-4.9373027	2.66E-06	0.00131378	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Nerve_Tibial	rs61524473	ENSG000000259520	antisense	LOC101928414	-0.7743297	-6.3368808	2.28E-08	2.46E-05	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Nerve_Tibial	rs61524473	ENSG000000259433	processed_pseudogene		-0.6584493	-5.5152584	6.04E-07	0.00043483	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Whole_blood	rs61524473	ENSG000000171763	protein_coding	SPATAS5L	0.37890372	5.87204221	3.18E-08	3.21E-05	Crea	15 rs61524473	45646283	GATM	old	NA	0.272263	
Skin_sun_exposed_lower_leg	rs11644601	ENSG000000183793	protein_coding	NPIAP5	0.50805101	4.72138515	1.07E-05	0.00782577	Bis.DR.ratio	16 rs11644601	15172118	PDXC1	old	intron	0.277297	
Muscle_skeletal	rs10083777	ENSG000000166454	protein_coding	ATMIN	0.37275208	4.38231689	2.58E-05	0.00851883	Gly	16 rs10083777	81065282	GCSH	new	TFBS	0.172891	
Nerve_Tibial	rs10083777	ENSG000000140905	protein_coding	GCSH	-0.6454039	-4.6893123	1.39E-05	0.00429704	Gly	16 rs10083777	81065282	GCSH	new	TFBS	0.172891	
Whole_blood	rs10083777	ENSG000000261061	sense_intronic	LOC102724325	-0.4315857	-4.022807	9.53E-05	0.03856528	Gly	16 rs10083777	81065282	GCSH	new	TFBS	0.172891	
Skin_sun_exposed_lower_leg	rs172642	ENSG000000040475	protein_coding	DVL2	-0.3277284	-4.331814	4.52E-05	0.02491013	Crea	17 rs172642	6595398	SLC13A5	new	intron	0.480953	
Heart_Left_Ventricle	rs2040771	ENSG000000100075	protein_coding	SLC25A1	-0.3678193	-4.3533879	5.11E-05	0.01975487	Cit	22 rs2040771	19161935	SLC25A1	old	TFBS	0.476266	
Thyroid	rs2040771	ENSG000000100075	protein_coding	SLC25A1	-0.5097302	-5.5108741	3.84E-07	0.00021771	Cit	22 rs2040771	19161935	SLC25A1	old	TFBS	0.476266	
Artery_Tibial	rs2040771	ENSG000000100075	protein_coding	SLC25A1	-0.4771556	-6.5316593	3.65E-09	7.56E-06	Cit	22 rs2040771	19161935	SLC25A1	old	TFBS	0.476266	
Lung	rs2040771	ENSG000000100075	protein_coding	SLC25A1	-0.3534833	-5.5393401	5.55E-07	0.00061845	Cit	22 rs2040771	19161935	SLC25A1	old	TFBS	0.476266	
Nerve_Tibial	rs2040771	ENSG000000100075	protein_coding	SLC25A1	-0.5105157	-5.1753537	2.25E-06	0.00121939	Cit	22 rs2040771	19161935	SLC25A1	old	TFBS	0.476266	
Skin_sun_exposed_lower_leg	rs2040771	ENSG000000100075	protein_coding	SLC25A1	-0.4374226	-5.4723843	5.63E-07	0.00120438	Cit	22 rs2040771	19161935	SLC25A1	old	TFBS	0.476266	
Whole_blood	rs2040771	ENSG000000100075	protein_coding	SLC25A1	-0.1892798	-5.1889422	7.58E-07	0.00051143	Cit	22 rs2040771	19161935	SLC25A1	old	TFBS	0.476266	

Supplementary Table 6. Genetic risk score for Lp(a)  
log(LPA) SNP associations in FINRISK97

ID	NEA	EA	Estimate	SE	P
rs10455872	A	G	1.29083958	0.03524657	7.66E-260
rs112092923	A	G	0.39071088	0.04693155	1.08E-16
rs117052562	G	A	0.57207416	0.04576285	2.53E-35
rs11756060	C	T	0.316227	0.03716715	2.32E-17
rs117898561	T	A	0.40094462	0.0577513	4.35E-12
rs190981730	G	A	0.75985504	0.07922578	1.35E-21
rs192776464	G	A	-0.5744705	0.04947457	9.07E-31
rs200319086	TA	T	-0.1480954	0.01382495	1.75E-26
rs201673688	G	GA	0.21289994	0.02607088	4.01E-16
rs2504921	T	G	-0.2855298	0.0141451	4.09E-87
rs41266385	A	C	0.5008693	0.03393023	2.76E-48
rs41267809	A	G	-0.3995413	0.04080005	1.94E-22
rs56393506	C	T	0.45768525	0.02010824	4.12E-109
rs71033586	AC	A	2.32579404	0.06692013	7.87E-237
rs783147	G	A	-0.1803628	0.01403455	3.34E-37
rs79018195	T	C	1.48259637	0.10269944	2.68E-46
rs79601080	A	G	-0.6400112	0.0758658	4.27E-17
rs9355839	G	A	0.22495482	0.02378786	4.79E-21
Sex			0.14549432	0.01666602	3.44E-18
Age			0.00313083	0.00063022	7.00E-07

EA Effect allele  
NEA Non-effect allele  
SE Standard error  
P P-value

log(LPA) SNP associations in YFS

ID	Estimate	SE	P
rs10455872	1.0902468	0.08211059	1.27E-38
rs112092923	0.67176931	0.10657425	3.57E-10
rs117052562	0.93784151	0.09257205	1.47E-23
rs11756060	0.3144094	0.07943694	7.82E-05
rs117898561	0.61401742	0.13035034	2.64E-06
rs190981730	0.7208449	0.19191886	1.78E-04
rs192776464	-1.0936299	0.10528971	1.19E-24
rs200319086	-0.2222792	0.02979609	1.28E-13
rs201673688	0.32991772	0.05630624	5.42E-09
rs2504921	-0.4711327	0.0298661	1.53E-43
rs41266385	1.01076523	0.07129889	2.54E-09
rs41267809	-0.723179	0.12080486	8.06E-80
rs56393506	0.83603685	0.04222552	3.21E-73
rs71033586	2.59944904	0.13776413	7.67E-29
rs783147	-0.3016177	0.02899995	2.39E-13
rs79018195	1.78164975	0.15738207	3.46E-07
rs79601080	-1.1458333	0.1553635	6.26E-53
rs9355839	0.25186722	0.04925064	1.04E-24
Sex	-0.0123919	0.03570944	7.29E-01
Age	0.00254736	0.00355268	4.73E-01

**Supplementary Table 7.** For the Dutch NMR method the most representative chemical shift values for the metabolic measures in the NMR spectra are shown (many metabolites have a complicated signal with components at multiple chemical shift locations). The representative coefficients of variation (in percent) are calculated for each metabolic measure based on 174 repeated quality control serum samples that are processed identically with the actual study samples; distributed randomly among the boxes represent the averaged situation for the total 7628 samples analysed.

Measure	CS / ppm	CV / %
Leucine	0.95	3.91
Isoleucine	1.00	4.45
Valine	1.03	3.96
3-Hydroxybutyrate	1.19	6.49
Lipids (CH <sub>2</sub> )	1.27	6.17
Alanine	1.47	4.38
Acetate	1.91	8.26
Pyruvate / Oxaloacetate	2.36	5.52
Glutamine	2.47	11.00
Citrate	2.53	5.88
Glycine	3.55	5.97
Glycerol	3.56	18.67
Creatinine	4.05	5.08
Lactate	4.10	4.68
Glucose	5.23	3.33
Lipids (CH=CH*CH <sub>2</sub> CH <sub>2</sub> )	5.30	12.01
Tyrosine	6.89	6.41
Phenylalanine	7.42	10.53

## **Supplementary Note 1: Sample descriptions**

### **1. FINRISK**

FINRISK surveys are cross-sectional, population-based studies conducted every five years since 1972 to monitor the risk of chronic diseases. For each survey, a representative random sample was selected from 25- to 74-year-old inhabitants of different regions in Finland. The survey included a questionnaire and a clinical examination, at which a blood sample was drawn, with linkage to national registers of cardiovascular and other health outcomes. The study protocol has been described elsewhere<sup>1</sup>. Study participants were followed up through December 31, 2012. The current study included eligible individuals from FINRISK surveys conducted in 1992, 1997, 2002, and 2007 (total n=27 838). FINRISK samples were genotyped in three separate batches: DILGOM, PredictCVD and FINRISK. All three batches were analyzed separately to avoid batch effects and meta-analyzed.

### **2. Helsinki Birth Cohort Study (HBCS)**

The primary aim of the Helsinki Birth Cohort Study was to assess how growth and environmental factors acting during the fetal period and childhood are related to health in adult life. A particular focus has been the study of the early life origins of cardiovascular disease and its risk factors. Participants were born between 1934 and 1944 in Helsinki. An essential component of the project was a detailed clinical examination of over 2500 volunteers, parts of which provided data included in this study<sup>2</sup>.

### **3. Health2000 GenMets Study (GenMets)**

The GenMets sample has been described in detail previously by Perttilä et al.<sup>3</sup>. Individuals in GenMets are metabolic syndrome cases and matched controls drawn from the Finnish Health2000 study; a population health survey. Persons with known diabetes were excluded. Participants included in the present study had participated in an additional survey, which included drawing fasting serum samples.

### **4. Northern Finland Birth Cohort 1966 (NFBC1966)**

The NFBC1966 has been described in detail previously by Rantakallio<sup>4</sup>. The original study design focused on factors affecting pre-term birth, low birth weight, and subsequent morbidity and mortality. Mothers living in the two northern-most provinces of Finland were invited to participate if they had expected delivery dates during 1966. Individuals still living in the Helsinki area or Northern Finland (N = 5923) were asked at age 31 to participate in a detailed biological and medical examination as well as a questionnaire.

### **5. Cardiovascular Risk in Young Finns Study (YF)**

The Cardiovascular Risk in Young Finns Study (YF) is a population based prospective cohort study. It was conducted at five medical schools in Finland (Turku, Helsinki, Kuopio, Tampere and Oulu), with the aim of studying the levels of cardiovascular risk factors in children and adolescents in different parts of the country. The latest follow-up was conducted in 2007. The serum samples for this metabolomics study were collected at this latest follow up. The study and data collection protocols have been described in detail by Raitakari et al.<sup>5</sup>.

### **6. Finnish twin cohort**

Individuals in this study were derived from the FinnTwin12 (FT12) and FinnTwin16 (FT16) cohort studies. In brief, the FT12 is a population-based cohort longitudinal study of five consecutive birth cohorts of Finnish twins born between 1983 and 1987. All twins (and their parents) were initially contacted and invited to participate by mail in the autumn of the year in which their birth cohort

reached 11 years of age. Subsequent follow-up assessments were made when the twins were aged 14, 17 and ~22 years. The FT16 is a population-based longitudinal study of five consecutive birth cohorts of Finnish twins born between 1975 and 1979. Each pair was initially approached and invited to participate by mail in the 1-2 months following the twins' 16th birthday. The baseline data collection started in 1991 and was completed in 1996, with participation of 2733 twin pairs (response rate ~ 88%). Subsequent follow-up assessments were made when the twins were 17, 18.5, and ~25 years, following a similar approach. For both the FT12 and FT16, the baseline and follow-up assessments included surveys of health habits and attitudes, symptom checklists, personality scales, and social relationships<sup>6</sup>. In addition, blood samples were taken from all twins during a visit to the twin research clinic in Helsinki (Finland) at the last follow-up (young adulthood) for DNA and biochemistry analyses. Data collection and analysis were approved by the ethics committees of the Department of Public Health of the University of Helsinki, the Helsinki and Uusimaa Hospital District and the IRB of Indiana University. Written informed consent was obtained from all participating twins.

## 7. ERF

The Erasmus Rucphen Family ([http://www.erasmusmc.nl/klinische\\_genetica/research/intro/genepi/](http://www.erasmusmc.nl/klinische_genetica/research/intro/genepi/)) genetic isolate study is a family based study in a genetically isolated population in the Southwest of the Netherlands. This young genetic isolate was founded in the mid - 18th century and minimal immigration and marriages have occurred between surrounding settlements due to social and religious reasons. The ERF population includes 3,465 individuals that are living descendants of 22 couples with at least six children baptized<sup>7</sup>. The study protocol was approved by the medical ethics board of the Erasmus MC Rotterdam, the Netherlands.

## 8. EGCUT

The Estonian Biobank is the population-based biobank of the Estonian Genome Center of the University of Tartu (EGCUT). The project is conducted according to the Estonian Gene Research Act and all participants have signed broad informed consent<sup>8</sup>. The cohort size is currently 51535 people aged 18 years and older. All subjects are volunteers and were recruited randomly by general practitioners (GP) and physicians in hospitals. A Computer Assisted Personal Interview is conducted at the doctor's office to record personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history spanning four generations), educational and occupational history, lifestyle data (physical activity, dietary habits - FFQ, smoking, alcohol consumption, women's health, quality of life). Metabolite concentrations (over 110; NMR and MS) are available for about 12000 subjects; metabolic follow-up has been conducted for more than 2000 donors. The EGCUT database has been linked with the national registries and hospital databases for obtaining up-to-date phenotypic information. Medical history and current health status are recorded according to the ICD-10, medication according to the ATC. Anthropometric measurements, blood pressure (sitting position at the end of the interview), and resting heart rate are measured; 30-50 mL of venous blood are collected into EDTA Vacutainers. These are transported to the central laboratory of the EGCUT at +4...+6 C (in 6 to 36h) where DNA, plasma and WBC are immediately isolated and kept in aliquots in MAPI straws in liquid N<sub>2</sub>. All procedures are run according to ISO 9000-2008 and LIMS.

## 9. KORA

KORA (Cooperative Health Research in the Region of Augsburg; <http://www.helmholtz-muenchen.de/en/kora-en/kora-homepage/index.html>) exists since 1996 in the region of Augsburg in the southwest of Germany, and builds on the MONICA (Monitoring of trends and determinants in cardiovascular disease) project initiated in 1984<sup>9</sup>. KORA is a regional research platform for population-based surveys and a cohort of more than 18,000 subjects are actively followed up until

the present date. Four cross-sectional health surveys have been performed in five-year intervals, each containing independent random samples of residents in the city of Augsburg and the two adjacent counties in the age-range between 25 to 74 years at baseline examination. The study followed the recommendations of the Declaration of Helsinki and was approved by the local ethical committees and informed consent was given by all participants. Blood was drawn in the morning between 8 and 10am after a period of overnight fasting. Material was immediately horizontal shaken (10 min), followed by 40 min resting at 4°C to obtain complete coagulation. The material was then centrifuged (2000 × g; 4°C). Serum was aliquoted and stored for 2–4 hours at 4°C, after which it was deep frozen to -80°C until sampling.

## 10. LLS

In the Leiden Longevity Study (<http://www.molepi.nl/research/longevity>), nonagenarian sibling pairs were included when aged older than 89 years for men and 91 years for women. Additional inclusion criteria were that participants needed to have at least one sister or a brother fulfilling these age criteria, and who was also willing to participate<sup>10</sup>. Because proper controls are lacking at very high ages, the offspring of the nonagenarian siblings were asked to be included in the study as well. The partners thereof were included in the study to serve as a control group, representing the general population at an age comparable to the offspring. The total study population, excluding the nonagenarian siblings, consisted of 2,415 participants (1,671 offspring; 744 partners). The Medical Ethical Committee of the Leiden University Medical Centre approved the study and informed consent was obtained from all participants. Blood samples were taken at baseline for extraction of DNA and the determination of non-fasted serum parameters. For metabolomics measurements, blood was drawn using a safety lock butterfly needle. Within two hours tubes with clotted blood were centrifuged for 15 minutes at 2,800 × g and serum was obtained and divided into subsamples, snap-frozen and stored at -80° C.

## 11. NTR

The Netherlands Twin Register (NTR; <http://www.tweelingenregister.org/>) recruits twins and their family members to study the causes of individual differences in health, behavior and lifestyle. Participants are followed longitudinally; details about the cohort have been published previously<sup>11</sup>. A subsample of unselected twins and their family members has taken part in the NTR-Biobank project. Genotyping was performed on the Affymetrix 6.0 ( $N=298$ ), Affymetrix Perlegen 5.0 ( $N=3,697$ ), Illumina 370 ( $N=290$ ), Illumina 660 ( $N=1,439$ ) and Illumina Omni Express 1M ( $N=455$ ) platforms. Calls were made with the platform specific software (Genotyper, Beadstudio). Per platform the quality control thresholds for SNPs were minor allele frequency (MAF) > 1%, Hardy-Weinberg equilibrium (HWE)  $P$  value > 0.00001, call rate >95%, and  $0.30 < \text{heterozygosity} < 0.35$ . Samples were excluded from the data if their expected sex and identity by descent (IBD) status did not match, or if the genotype missing rate was above 10%. For each platform all SNPs were aligned to the positive strand of the Hapmap 2 Build 36 release 24 CEU reference set. The alignment was checked using individuals and family members tested on multiple platforms. SNPs were excluded per platform if allele frequencies differed more than 15% with the reference set and/or the other platforms. The data of the platforms were subsequently merged into a single dataset ( $N=5,856$ ). This merged set was imputed against the reference set using IMPUTE v2. After imputation, genotype dosage was calculated if the highest genotype probability was above 90%. Badly imputed SNPs were removed based on HWE  $P$  value < 0.00001, proper info < 0.40, MAF < 1%, allele frequency difference > 0.15 against reference. From the total set of 5,856 individuals with imputed GWA data, a subset of unrelated participants were included in the present study. For metabolomics measurements, blood was collected in the morning using a safety lock butterfly needle.

## **Supplementary Note 2: Additional acknowledgements**

*The Estonian Genome Center of University of Tartu (EGCUT)* received support from EU FP7 grant BBMRI-LPC 313010, targeted financing from Estonian Government IUT20-60, IUT24-6, Estonian Research Roadmap through the Estonian Ministry of Education and Research (3.2.0304.11-0312), Center of Excellence in Genomics (EXCEGEN), Development Fund from the University of Tartu (SP1GVARENG) and from an EFSD New Horizons grant. EGCUT thanks Viljo Soo Estonian Genome Center for his help on this project. This work was also supported by the US National Institute of Health [R01DK075787].

*The Leiden Longevity Study (LLS)* has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 259679. This study was supported by a grant from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology, and the Netherlands Consortium for Healthy Ageing (grant 050-060-810), all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO), Unilever Colworth and by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007).

*Netherlands Twin Registry (NTR)* Research was funded by the Netherlands Organization for Scientific Research (NWO: MagW/ZonMW grants 904-61-090, 985-10-002, 904-61-193, 480-04-004, 400-05-717, Addiction-31160008 Middelgroot-911-09-032, Spinozapremie 56-464-14192), Center for Medical Systems Biology (CSMB, NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007), the VU University's Institute for Health and Care Research (EMGO+ ), the European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-F4-2007-201413) and the European Science Council (ERC - 230374 and ERC-284167). HHMD is supported by an EMGO+ Fellowship as part of the Mental Health research program of the EMGO Institute for Health and

Care Research. NTR thanks all participants in the Netherlands Twin Register. This work was carried out on the Dutch national e-infrastructure with the support of SURF Foundation.

*Erasmus Rucphen Family (ERF)* was supported by the Consortium for Systems Biology (NCSB), both within the framework of the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO). ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254) as well as FP7 project EUROHEADPAIN (nr 602633). High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organisation for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). Exome sequencing analysis in ERF was supported by the ZonMw grant (2010/31471/ZONMW). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscripts. ERF study is grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection.

*The Cardiovascular Risk in Young Finns Study (YFS)* The Young Finns Study has been financially supported by the Academy of Finland: grants 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001 for T.L.); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research (T.L.);

Finnish Cultural Foundation; Tampere Tuberculosis Foundation (T.L.); Emil Aaltonen Foundation (T.L.); and Yrjö Jahnsson Foundation (T.L.).

*Helsinki Birth Cohort Study (HBCS)* has been supported by grants from the Academy of Finland (120386, 125876 and 126775 to J.E.), the Finnish Diabetes Research Society, the Novo Nordisk Foundation, the European Science Foundation (EuroSTRESS), the Wellcome Trust (89061/Z/09/Z and 089062/Z/09/Z), the Samfundet Folkhälsan and the Finska Läkaresällskapet.

*Northern Finland Birth Cohort 1966 (NFBC1966)* received financial support from the Academy of Finland (104781, 120315, 129269, 1114194, 139900 and SALVE to M.-R.J.) University Hospital Oulu, Biocenter, University of Oulu (75617 to M.-R.J. and M.S.), the European Commission EURO-BLCS Framework 5 award (QLG1-CT-2000-01643 to M.-R.J.), the US National Heart, Lung, and Blood Institute (NHLBI) (5R01HL087679), the US National Institute of Mental Health (NIMH) (1RL1MH083268), European Network for Genetic and Genomic Epidemiology (ENGAGE) (HEALTH-F4-2007-201413 to L.P. and M.-R.J.), the MRC UK (G0500539, G0600705 and PrevMetSyn/Salve to M.-R.J.) and the Wellcome Trust (GR069224).

*Finnish Twin Cohort Study* Data collection for FinnTwin12 and FinnTwin16 were supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) (AA-12502, AA-09203 and AA-08315 to R.J.Rose) and the Academy of Finland (100499, 205585, 118555, 265240, 263278 and 141054 (Skidi-Kids) to J. Ka). L.H.B was supported by 1.3 milj. klubbi-klubben. Genotyping was performed at the Wellcome Trust Sanger Institute.

*FINRISK* The FINRISK/DILGOM study was supported by the Academy of Finland (118081) and by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. HEALTH-F2-2011-278913 (BiomarCaRE). Abbott Diagnostics provided test reagents for Lp(a). Academy of Finland supported V.S. (grant #139635), M.P. (grant #269517) and J.K. (grant #283045) and V.S. was supported by Finnish Foundation for Cardiovascular Research. M.P. was supported also through EU FP7 under grant agreements nr. 313010 (BBMRI-LPC), nr. 305280

(MIMOMics), and nr. 261433 (BioSHaRE-EU), the Yrjö Jahnsson Foundation and Juho Vainio Foundation.

KORA The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. This work was financed by the European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project (HEALTH-F4-2007-201413), and the German National Genome Research Network (NGFN; NGFNPlus, project number 01GS0834). We thank Julia Scarpa and Katharina Sckell for metabolomics measurements performed at the Helmholtz Centrum München, Genome Analysis Center, Metabolomics Core Facility. Karsten Suhre is supported by Biomedical Research Program funds at Weill Cornell Medical College in Qatar, a program funded by the Qatar Foundation. Christian Gieger is supported by the European Union's Seventh Framework projects MIMOMics (FP7-Health-F5-2012-305280) and PAIN-OMICS (project nr. 602736), by the BMBF project e:AtheroSysMed (project nr.: 01ZX1313G) and by the Helmholtz-Russia joint research group HRJRG-310.

The quantitative NMR metabolomics platform and its development has been supported by the Academy of Finland, TEKES - the Finnish Funding Agency for Technology and Innovation, the Sigrid Juselius Foundation, the Novo Nordisk Foundation, the Finnish Diabetes Research Foundation, the Paavo Nurmi Foundation, and the strategic and infrastructural research funding from the University of Oulu, Finland, as well as by the British Heart Foundation, the Wellcome Trust, and the Medical Research Council, UK.

GTEX: The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health. Additional funds were provided by the NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. Donors were enrolled at Biospecimen

Source Sites funded by NCI\SAIC-Frederick, Inc. (SAIC-F) subcontracts to the National Disease Research Interchange (10XS170), Roswell Park Cancer Institute (10XS171), and Science Care, Inc. (X10S172). The Laboratory, Data Analysis, and Coordinating Center (LDACC) was funded through a contract (HHSN268201000029C) to The Broad Institute, Inc. Biorepository operations were funded through an SAIC-F subcontract to Van Andel Institute (10ST1035). Additional data repository and project management were provided by SAIC-F (HHSN261200800001E). The Brain Bank was supported by a supplements to University of Miami grants DA006227 & DA033684 and to contract N01MH000028. Statistical Methods development grants were made to the University of Geneva (MH090941 & MH101814), the University of Chicago (MH090951, MH090937, MH101820, MH101825), the University of North Carolina - Chapel Hill (MH090936 & MH101819), Harvard University (MH090948), Stanford University (MH101782), Washington University St Louis (MH101810), and the University of Pennsylvania (MH101822). The data used for the analyses described in this manuscript were obtained from: the GTEx Portal on 12/01/2014 and/or dbGaP accession number phs000424.v3.p1 on 08/03/2013.

## Supplementary References

1. Vartiainen, E. et al. Thirty-five-year trends in cardiovascular risk factors in Finland. *Int J Epidemiol* **39**, 504-18 (2010).
2. Yliharsila, H. et al. Body mass index during childhood and adult body composition in men and women aged 56-70 y. *Am J Clin Nutr* **87**, 1769-75 (2008).
3. Perttila, J. et al. OSBPL10, a novel candidate gene for high triglyceride trait in dyslipidemic Finnish subjects, regulates cellular lipid metabolism. *J Mol Med (Berl)* **87**, 825-35 (2009).
4. Rantakallio, P. Groups at risk in low birth weight infants and perinatal mortality. *Acta Paediatr Scand* **193**, Suppl 193:1+ (1969).
5. Raitakari, O.T. et al. Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol* **37**, 1220-6 (2008).
6. Kaprio, J., Pulkkinen, L. & Rose, R.J. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. *Twin Res* **5**, 366-71 (2002).
7. Aulchenko, Y.S. et al. Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* **12**, 527-34 (2004).
8. Leitsalu, L. et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *Int J Epidemiol* (2014).

9. Holle, R., Happich, M., Lowel, H., Wichmann, H.E. & Group, M.K.S. KORA--a research platform for population based health research. *Gesundheitswesen* **67 Suppl 1**, S19-25 (2005).
10. Schoenmaker, M. et al. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. *Eur J Hum Genet* **14**, 79-84 (2006).
11. Willemsen, G. et al. The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. *Twin Res Hum Genet* **13**, 231-45 (2010).