nature genetics

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41588-019-0567-8

In the format provided by the authors and unedited.

Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion in humans

Pascal Schlosser[1,16], Yong Li[1,16], Peggy Sekula[1,16], Johannes Raffler[2], Franziska Grundner-Culemann[1], Maik Pietzner[3,4], Yurong Cheng[1], Matthias Wuttke[1,5], Inga Steinbrenner[1], Ulla T. Schultheiss[1,5], Fruzsina Kotsis[1,5], Tim Kacprowski[4,6,7], Lukas Forer[8], Birgit Hausknecht[9], Arif B. Ekici[10], Matthias Nauck[3,4], Uwe Völker[4,6], GCKD Investigators[11], Gerd Walz[5], Peter J. Oefner[12], Florian Kronenberg[8], Robert P. Mohney[13], Michael Köttgen[5], Karsten Suhre[14], Kai-Uwe Eckardt[9,15], Gabi Kastenmüller[2] and Anna Köttgen[1*]

[1]Institute of Genetic Epidemiology, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany. [2]Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. [3]Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany. [4]German Center for Cardiovascular Research (DZHK e.V.), partner site Greifswald, Greifswald, Germany. [5]Department of Medicine IV: Nephrology and Primary Care, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany. [6]Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany. [7]Research Group on Computational Systems Medicine, Chair of Experimental Bioinformatics, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany. [8]Department of Genetics and Pharmacology, Institute of Genetic Epidemiology, Medical University of Innsbruck, Innsbruck, Austria. [9]Department of Nephrology and Hypertension, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany. [10]Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany. [11]A list of members and affiliations appears in the Supplementary Note. [12]Institute of Functional Genomics, University of Regensburg, Regensburg, Germany. [13]Metabolon Inc., Durham, NC, USA. [14]Department of Physiology and Biophysics, Weill Cornell Medicine-Qatar, Education City, Doha, Qatar. [15]Department of Nephrology and Medical Intensive Care, Charité–Universitätsmedizin Berlin, Berlin, Germany. [16]These authors contributed equally: Pascal Schlosser, Yong Li, Peggy Sekula. *e-mail: anna.koettgen@uniklinik-freiburg.de

Genetic Studies of Urinary Metabolites Illuminate Mechanisms of Detoxification and Excretion in Humans

Pascal Schlosser*, Yong Li*, Peggy Sekula*, Johannes Raffler, Franziska Grundner-Culemann, Maik Pietzner, Yurong Cheng, Matthias Wuttke, Inga Steinbrenner, Ulla T. Schultheiss, Fruzsina Kotsis, Tim Kacprowski, Lukas Forer, Birgit Hausknecht, Arif B. Ekici, Matthias Nauck, Uwe Völker, GCKD Investigators**, Gerd Walz, Peter J. Oefner, Florian Kronenberg, Robert P. Mohney, Michael Köttgen, Karsten Suhre, Kai-Uwe Eckardt, Gabi Kastenmüller, Anna Köttgen

* these authors contributed equally
** a list of the GCKD Investigators is included in Supplementary Information

SUPPLEMENTARY INFORMATION

Table of Contents

Supplementary Note

Incorporation of existing biological knowledge into causal gene assignment

The workflow to assign potentially causal genes in GWAS loci was agnostic with respect to existing biological and biochemical knowledge. Upon evaluation after the gene had been assigned, the great majority of automatically assigned genes was also supported by existing biochemical knowledge and experimental studies. There were a few instances, however, in which incorporation of existing biochemical and biological knowledge into the causal gene assignment process would have supported another gene in the locus (see Table). At these loci, it may be that the automated gene assignment that was agnostic with respect to existing biological and biochemical knowledge did not prioritize the correct causal gene. As unbiased databases to facilitate automated gene assignment become more and more complete, for example through the generation of gene expression data in additional tissues and cell types, the assignment of the most likely gene in a given region may be subject to change. Regardless, sensitivity analyses repeating all enrichment analyses with the use of these genes instead of the automatically assigned ones at these eight loci yielded the same or almost identical enriched pathways, tissues, and cell types (data not shown).

Automatically assigned gene in locus (gene score)	**Biologically supported gene in locus (gene score)**	**Associated known metabolite(s) or module(s)**	**Background on biologically supported gene**
CASP9 (hremc)	*AGMAT* (hrem)	4-guanidinobutanoate, beta-guanidinopropanoate	The enzyme encoded by *AGMAT*, agmatinase, metabolizes N-(4-aminobutyl)guanidine [http://www.hmdb.ca/]
LRP8 (hep)	*CPT2* (h)	methylsuccinoylcarnitine	Succinoylcarnitine is a fatty acid. *CPT2* encodes carnitine palmitoyltransferase II, which acts on fatty acids.
ZKSCAN5 (he)	*CYP3A7* (NA)	16a-hydroxy DHEA 3-sulfate, andro steroid monosulfate C19H28O6S (1)*, tauro-beta-muricholate	The enzyme encoded by *CYP3A7* hydroxylates dehydroepiandrosterone 3-sulphate.
PPP2R4 (hrem)	*CRAT* (hre)	2-methylmalonylcarnitine	*CRAT* encodes carnitine O-acetyltransferase. This enzyme

		(C4-DC)	converts short- and medium-chain acyl-CoAs, to which 2-methylmalonylcarnitine belongs.
TRIM48 (NA)	*FOLH1* (NA)	N-acetyl-aspartyl-glutamate (NAAG)	*FOLH1* encodes folate hydrolase 1, which metabolizes NAAG [PMID: 9622670]. The index SNP rs61898064 that gives rise to the automated assignment of *TRIM48* is in LD (r^2=0.798) with another mQTL (rs55728336). This other mQTL is associated with NAAG, and its index SNP was automatically assigned to *FOLH1*. The two signals were not merged because the r^2 was not >0.8.
NUPR1 (he), *CCDC101* (he)	*SULT1A2* (he)	3-hydroxyindolin-2-one sulfate, furaneol sulfate	The enzyme encoded by *SULT1A2* catalyzes the sulfate conjugation of a wide variety of molecules.
TYMS (ohre)	*ENOSF1* (ohe)	ribonate	The enzyme encoded by *ENOSF1* plays a role in the catabolism of the deoxy sugar L-fucose. Ribonate is a sugar acid.
CABP5 (e)	*SULT2A1* (e)	androstenediol (3beta,17beta) disulfate (1)	The encoded enzyme, dehydroepiandrosterone sulfotransferase, catalyzes the sulfation of steroids and bile acids including DHEA, of which androstenediol is a direct metabolite.
BTN3A1 (hem)	*SLC17A1/A3/A4* (h)	ME41	The metabolites assigned to this module belong to a substrate call transported by the SLC17A transporter family. The individually significant metabolite in the cluster, indolelactate, is assigned to *SLC17A1*/*A3*/*A4*.
RAB11FIP5 (h)	*NAT8* (NA)	ME160, ME161, ME166	All known metabolites assigned to the listed clusters are N-acetylated compounds. N-acetylation is a key function of the enzyme encoded by *NAT8*.

Metabolite clusters provide biological context for yet unnamed metabolites

Metabolites are intermediates of homeostatic reactions and as such inter-connected beyond pair-wise relationships. Groups of correlated metabolites ("modules") may reflect shared biochemical pathways or co-regulation. We used a weighted gene co-expression analysis-based approach, to construct 212 metabolite modules (Methods, **Extended Data Figure 6A**). GWAS of the modules' first principal component, the eigenmetabolite, identified 46 significant (P<2.4e-10 [5e-8/212]) and replicated associations between genetic variants and 38 unique metabolite modules (**Extended Data Figure 7, Supplementary Table 12**). In three instances, the gene scored as most likely to be causal was not part of the 90 genes identified in the single metabolite screen. One of them, *CPT2*, is also supported by biological evidence but did not receive the highest score within the locus in the single metabolite screen (see above). At the other four genes, biological evidence points toward a different gene than the one automatically assigned (see above, **Supplementary Table 12 + 15**).

Eigenmetabolites that showed particularly strong genetic associations originated from a module of five unknown metabolites (missense rs2147896 in *PYROXD2*, P=2.5e-917) and a module composed of N2-acetyllysine, N-alpha-acetylornithine, X-12124, X-12125, and X-15666 (missense rs13538 in *NAT8*, P=4.3e-635; **Extended Data Figure 6B, C**). Such associations are suggestive of a common function of the enzyme on metabolites in the module, implicating the unknown molecules in the *NAT8*-associated module as additional N-acetylated compounds or their precursors. Similarly, a module of the known vitamin E (tocopherol)-related metabolites also contained the two unknowns X-13689 and X-24359 (**Supplementary Table 12**) and was associated with rs55744319, which is in high LD with a missense variant in *CYP4F2*, encoding p.Val433Met. This variant has previously been identified in response to vitamin E supplementation[1], vitamin E levels[2], and warfarin

maintenance dose[2,3]. Investigation of the unknown metabolites based on their mass, retention time, spectral information and genetic evidence nominated the unknown molecules as structurally related to Vitamin E, with the glucuronide of alpha-CMBHC as a candidate for X-13689. We experimentally verified this prediction through the examination and comparison of retention times from ion chromatograms and the locations and intensities of the MS/MS fragmentation spectra between a standard of the glucuronide of alpha-CMBHC and X-13689 (**Extended Data Figure 7**). Thus, knowledge of an unknown's module membership and its genetic association can provide information beyond mass and retention time by restricting the search space of their possible identity for experimental verification.

While there were no genetic associations with modules that did not contain at least one metabolite also identified by mGWAS, screening of eigenmetabolites provided the important advantage of permitting a hypothesis-generating screen of higher order genetic associations, whereas already the assessment of all pair-wise metabolite ratios would have accumulated to 686,206 GWAS. Furthermore, 35 of 46 eigenmetabolite associations implicate additional metabolites that were not identified in mGWAS after correction for multiple testing (**Extended Data Figure 7, Supplementary Table 12**). We have made our software, used for identification of eigenmetabolites, publicly available (https://github.com/genepi-freiburg/Netboost), which may be of particular interest for emerging large-scale integrative Omics efforts.

Metabolite ratios capture insights into physiology and pharmacogenetics

The renal tubular amino acid exchanger *SLC7A9* is known to exchange dibasic amino acids such as lysine from urine against intracellular neutral amino acids in model systems[4,5]. We therefore screened all pair-wise metabolite ratios for this known exchanger. A *P*-gain threshold of 6,728,320 (672,832*10) was used to identify ratios that contributed information beyond their individual components, where 672,832 represents the number of tested ratios (1172*1171/2) after exclusion of ratios with less than 300 measurements. The index SNP rs12460876, associated with differential *SLC7A9* expression (**Supplementary Table 3**), was related to 83 informative metabolite ratios (*P*-gain >6.7e6, **Extended Data Figure 5**). Of these, all ratios that contained lysine also contained neutral amino acids such as phenylalanine, threonine, glutamine, or alanine, reflecting its known physiological function, amino acid exchange at the apical membrane of tubular epithelial cells[4]. In this screen, 5-hydroxy-lysine and the unknown metabolite X-24736 emerged as novel candidate substrates of this exchanger (**Extended Data Figure 5**). Based on spectral information, mass and retention time, X-24736 is likely an arginine-containing metabolite, consistent with the uptake of dibasic amino acids from urine. The identification of novel candidate substrates of known transport proteins is of high interest for pharmaceutical research, both with respect to a target's therapeutic potential but also to anticipate potential side effects.

Association between NAT8-associated metabolites and CKD progression and complications

Given the high and near-exclusive expression of *NAT8* in kidney, we tested whether the 30 *NAT8*-associated metabolites may carry information about the risk of CKD progression and CKD-related endpoints complementary to eGFR, for example by capturing detoxification capacity (**Supplementary Methods**). N-acetylation is an important reaction in a major route of detoxification, the generation of water-soluble mercapturic acids[6]. Spearman correlation coefficients of the 30 metabolites with eGFR, the main measure of kidney function in clinical practice, were weak (range -0.17 to 0.24). We assessed the association of the *NAT8*-associated metabolites with incident end-stage kidney disease (ESKD, n=61), incident major cardiovascular events (MACE, n=143) and all-cause mortality (n=129; Methods). In comparison to a model with clinical information alone, inclusion of metabolites significantly improved the model fit for ESKD (*P*=1.5e-4, Methods, **Supplementary Table 14**). While higher urinary concentrations of X-13698 and N-acetylglutamine were protective (hazard ratio [HR]=0.59 for both), N-acetylkynurenine, N-acetylcitrulline, N-delta-acetylornithine and X-12125 were associated with higher risk (HR range: 1.17-1.47). These metabolites therefore represent potential new biomarkers of ESKD, along with altered N-acetylation capacity as an implicated mechanism, for evaluation in larger studies of CKD progression.

Extended acknowledgements

List of GCKD Study Investigators

A list of nephrologists currently collaborating with the GCKD study is available at http://www.gckd.org.

University of Erlangen-Nürnberg	Kai-Uwe Eckardt, Heike Meiselbach, Markus Schneider, Thomas Dienemann, Hans-Ulrich Prokosch, Barbara Bärthlein, Andreas Beck, Thomas Ganslandt, André Reis, Arif B. Ekici, Susanne Avendaño, Dinah Becker-Grosspitsch, Ulrike Alberth-Schmidt, Birgit Hausknecht, Rita Zitzmann, Anke Weigel
University of Freiburg	Gerd Walz, Anna Köttgen, Ulla Schultheiß, Fruzsina Kotsis, Simone Meder, Erna Mitsch, Ursula Reinhard
RWTH Aachen University	Jürgen Floege, Georg Schlieper, Turgay Saritas, Sabine Ernst, Nicole Beaujean
Charité, University Medicine Berlin	Elke Schaeffner, Seema Baid-Agrawal, Kerstin Theisen
Hannover Medical School	Hermann Haller, Jan Menne
University of Heidelberg	Martin Zeier, Claudia Sommerer, Rebecca Woitke
University of Jena	Gunter Wolf, Martin Busch, Rainer Fuß
Ludwig-Maximilians University of München	Thomas Sitter, Claudia Blank
University of Würzburg	Christoph Wanner, Vera Krane, Antje Börner-Klein, Britta Bauer
Medical University of Innsbruck, Division of Genetic Epidemiology	Florian Kronenberg, Julia Raschenberger, Barbara Kollerits, Lukas Forer, Sebastian Schönherr, Hansi Weissensteiner
University of Regensburg, Institute of Functional Genomics	Peter Oefner, Wolfram Gronwald, Helena Zacharias
Department of Medical Biometry, Informatics and Epidemiology (IMBIE), University of Bonn	Matthias Schmid, Jennifer Nadal

Supplementary Methods

Non-targeted mass spectrometry analysis

Sample preparation at Metabolon, Inc was carried out as described previously[7]. Briefly, recovery standards were added prior to the first step in the extraction process for quality control purposes. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills Genogrinder 2000) followed by centrifugation. The resulting extract was divided into fractions and vacuum dried. For each sample, dried extracts were dissolved in injection solvent containing eight or more injection standards at fixed concentrations, depending on the platform, to assure injection and chromatographic consistency. Each sample was analyzed by four ultra-high performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) methods: 1) One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18-2.1x100 mm, 1.7 μm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). 2) A second aliquot was also analyzed using acidic positive ion conditions; however, it was chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same aforementioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an overall higher organic content. 3) A third aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, however with 6.5mM ammonium bicarbonate at pH 8. 4) The

fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 µm) using a gradient consisting of water and acetonitrile with 10mM ammonium formate, pH 10.8. 5) A fifth aliquot was reserved for backup.

Three types of controls were analyzed in concert with the experimental samples: samples generated from a pool of human urine extensively characterized by Metabolon, Inc. served as technical replicate throughout the data set; extracted water samples served as process blanks; and a cocktail of standards spiked into every analyzed sample allowed instrument performance monitoring. Instrument variability was determined by calculating the median relative standard deviation (RSD) for the standards that were added to each sample prior to injection into the mass spectrometers (median RSD=5-7%; n=31 standards). Overall process variability was determined by calculating the median RSD for all endogenous metabolites (i.e., non-instrument standards) present in 100% of the pooled human urine samples (median RSD=7-9%; n>1,000 metabolites). All RSDs for metabolites present in at least 90% of the pooled human urine samples are reported in **Supplementary Table 2.** Experimental samples and controls were randomized across the platform run. All methods utilized a Waters ACQUITY UPLC and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. Instruments were tuned and calibrated for mass resolution and mass accuracy daily. The MS analysis alternated between MS and data-dependent MSn scans using dynamic exclusion. The scan range varied slighted between methods but covered 70-1,000 m/z.

Metabolite identification and quantification

Metabolites were identified by automated comparison of the ion features in the experimental samples to a reference library of chemical standard entries that included retention time, molecular weight (m/z), preferred adducts, and in-source fragments as well as associated MS spectra, and curated by visual inspection for quality control using software developed at Metabolon. Identification of known chemical entities was based on comparison to Metabolon's spectral library of >4,500 purified chemical standards. Commercially available purified standard compounds have been acquired and registered into LIMS for distribution to the various UPLC-MS/MS platforms for determination of their detectable characteristics. Known metabolites reported in this study conformed to confidence Level 1 (the highest confidence level of identification) of the Metabolomics Standards Initiative[8,9], unless otherwise denoted with an asterisk. Additional mass spectral entries have been created for structurally unnamed biochemicals (>2,750 in the Metabolon library), which have been identified by virtue of their recurrent nature (both chromatographic and mass spectral). These compounds have the potential to be identified by future acquisition of a matching purified standard or by classical structural analysis, and were included in this study.

Peaks were quantified using area-under-the-curve. Raw area counts for each metabolite in each sample were normalized to correct for variation resulting from instrument inter-day tuning differences by the median value for each run-day, therefore, setting the medians to 1.0 for each run. This preserved variation between samples but allowed metabolites of widely different raw peak areas to be compared on a similar graphical scale.

GWAS of eigenmetabolites

Netboost was applied to the GCKD discovery sample, and the resulting clustering transferred to the replication sample. GWAS were conducted separately for discovery and replication eigenmetabolites and results were meta-analyzed using a fixed effects inverse variance model in METAL[10]. Statistical significance was defined as genome-wide significance (P<5e-8) in the discovery sample, a one-sided P<0.05 in the replication sample, and significance (P<2.4e-10) in the meta-analysis after correcting for multiple testing by a Bonferroni procedure (5e-8/212 eigenmetabolites). Significantly associated 1-Mb intervals were merged by overlap, and loci were merged into genetic regions across eigenmetabolites if their index SNPs were correlated (r^2>0.8). The extended MHC region (chromosome 6, 25.5-34 Mb) was considered as one region.

Associations with CKD progression and CKD-related endpoints

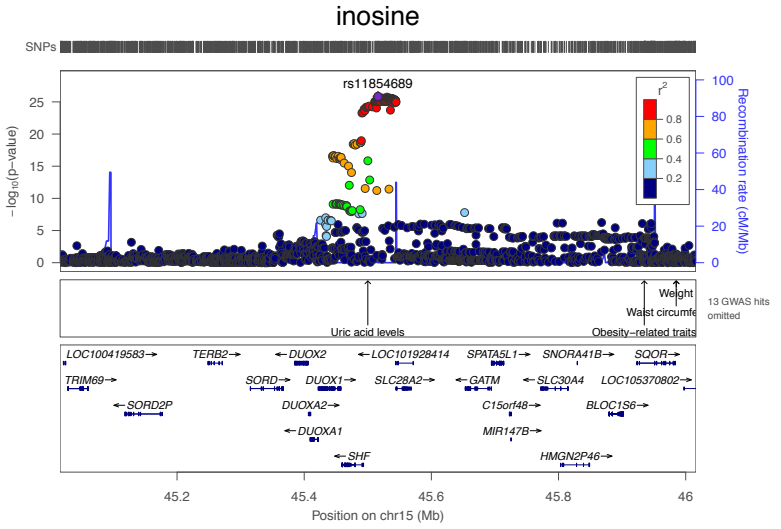
Prospective endpoints abstracted in the GCKD-study include ESKD, cardiovascular events and death[11]. For this project, endpoints over the first four years of follow-up were included (median follow-up time: 1,476 days). To evaluate the association of the 30 *NAT8*-associated metabolites, survival regression models were fitted for time from study entry to first event. Evaluated endpoints were overall mortality, ESKD (combined endpoint of dialysis, kidney transplantation and kidney-related death) and major adverse cardiac events (MACE, combined endpoint of myocardial infarction, stroke, peripheral vascular disorders and heart-related deaths). Besides $\log_2$-transformed metabolites, the four clinical variables age, sex, log(eGFR) and log(UACR) were considered; 1,619 GCKD participants had complete data. Three different Cox proportional hazard models were fitted: 1) clinical model (4 clinical variables), 2) metabolite model (selected metabolites using backward elimination based on Akaike's information criterion), and 3) combined model (repetition of backward elimination of metabolites selected for model 2 with forced inclusion of clinical variables). The deviance test based on the likelihood ratio statistic was used to evaluate significance of model improvement. The Cox proportional hazard model provided hazard ratio estimates for overall mortality and cause-specific hazard ratio estimates for ESKD and MACE in the presence of competing events[12]. No metabolites were associated with the competing events after correction for multiple testing (data not shown). No major violations to the proportional hazard assumption were found based on scaled Schoenfeld residuals and graphical assessment[13].

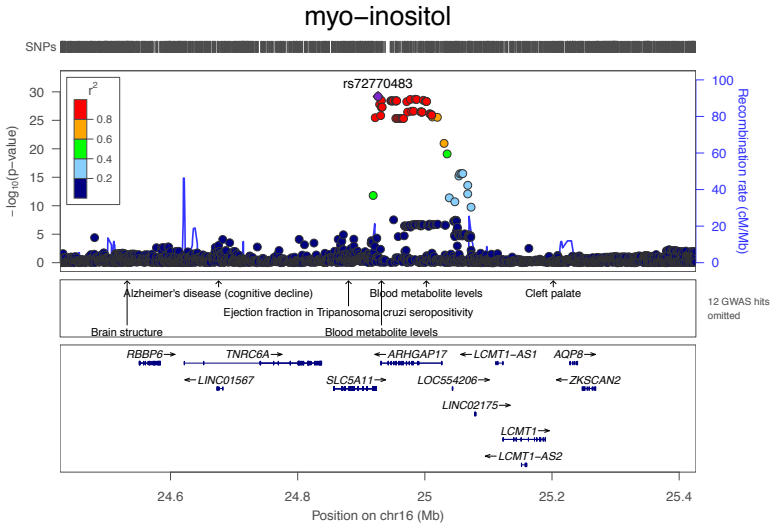
References:

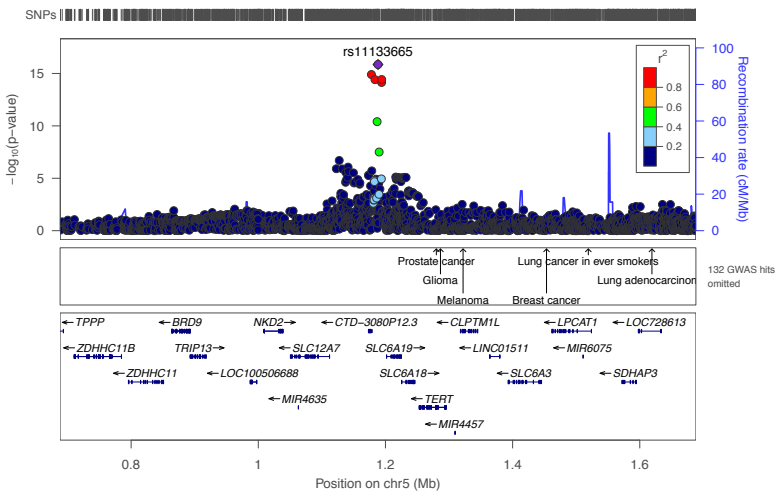
1. Major, J.M. *et al.* Genome-wide association study identifies three common variants associated with serologic response to vitamin E supplementation in men. *J Nutr* **142**, 866-71 (2012).
2. Major, J.M. *et al.* Genome-wide association study identifies common variants associated with circulating vitamin E levels. *Hum Mol Genet* **20**, 3876-83 (2011).
3. Takeuchi, F. *et al.* A genome-wide association study confirms VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. *PLoS Genet* **5**, e1000433 (2009).
4. Li, Y. *et al.* Genome-Wide Association Studies of Metabolites in Patients with CKD Identify Multiple Loci and Illuminate Tubular Transport Mechanisms. *J Am Soc Nephrol* **29**, 1513-1524 (2018).
5. Bertran, J. *et al.* Expression cloning of a cDNA from rabbit kidney cortex that induces a single transport system for cystine and dibasic and neutral amino acids. *Proc Natl Acad Sci U S A* **89**, 5601-5 (1992).
6. Veiga-da-Cunha, M. *et al.* Molecular identification of NAT8 as the enzyme that acetylates cysteine S-conjugates to mercapturic acids. *J Biol Chem* **285**, 18888-98 (2010).
7. Evans, A.M. *et al.* High Resolution Mass Spectrometry Improves Data Quantity and Quality as Compared to Unit Mass Resolution Mass Spectrometry in High-Throughput Profiling Metabolomics. *Metabolomics* **4**, 7 (2014).
8. Sumner, L.W. *et al.* Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **3**, 211-221 (2007).
9. Schrimpe-Rutledge, A.C., Codreanu, S.G., Sherrod, S.D. & McLean, J.A. Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *J Am Soc Mass Spectrom* **27**, 1897-1905 (2016).
10. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
11. Eckardt, K.U. *et al.* The German Chronic Kidney Disease (GCKD) study: design and methods. *Nephrol Dial Transplant* **27**, 1454-60 (2012).
12. Austin, P.C., Lee, D.S. & Fine, J.P. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. *Circulation* **133**, 601-9 (2016).
13. Grambsch, P.M. & Therneau, T.M. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika* **81**, 515-526 (1994).

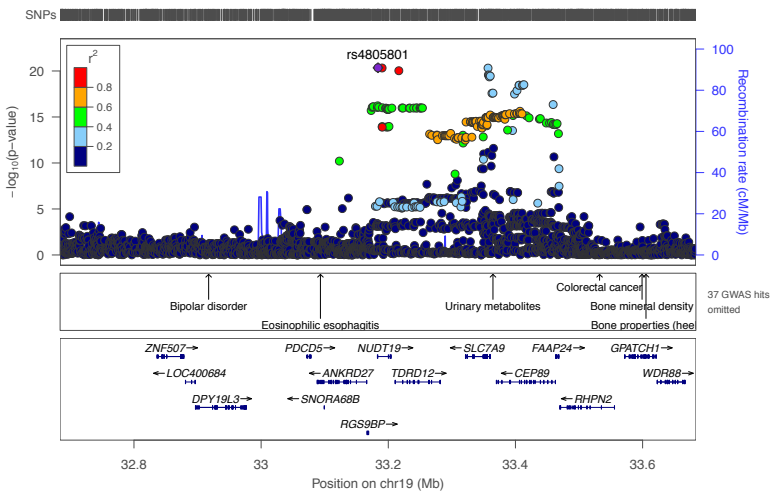
Supplementary Figure 1: Regional association plots for mQTLs identified in mGWAS of urinary metabolite concentrations

For each of the 240 mQTLs, the region for plotting was selected as the outer borders of merged overlapping 1-Mb windows. The extended MHC region was treated as one region. The index SNP with the lowest p-value is indicated. P-values were calculated based on linear regression. The metabolite that gave rise to the association is included in the title. Linkage disequilibrium information, used to color-code correlation with the index SNP, was calculated from the analyzed subsample of the GCKD study. For additional detail on index SNP association see **Supplementary Table 3**.

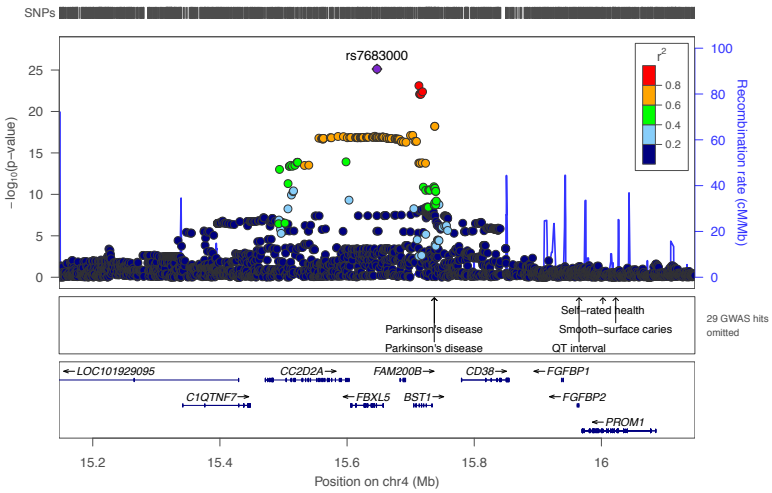

inosine
SNPs
rs11854689
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
Uric acid levels
Weight
Waist circumfe
Obesity−related traits
13 GWAS hits omitted
LOC100419583
TERB2
DUOX2
LOC101928414
SPATA5L1
SNORA41B
SQOR
TRIM69
SORD
DUOX1
SLC28A2
GATM
SLC30A4
LOC105370802
SORD2P
DUOXA2
C15orf48
BLOC1S6
DUOXA1
MIR147B
SHF
HMGN2P46
45.2
45.4
45.6
45.8
46
Position on chr15 (Mb)


myo−inositol
SNPs
rs72770483
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
Alzheimer's disease (cognitive decline)
Blood metabolite levels
Cleft palate
Ejection fraction in Tripanosoma cruzi seropositivity
Brain structure
Blood metabolite levels
12 GWAS hits omitted
RBBP6
TNRC6A
ARHGAP17
LCMT1−AS1
AQP8
LINC01567
SLC5A11
LOC554206
ZKSCAN2
LINC02175
LCMT1
LCMT1−AS2
24.6
24.8
25
25.2
25.4
Position on chr16 (Mb)

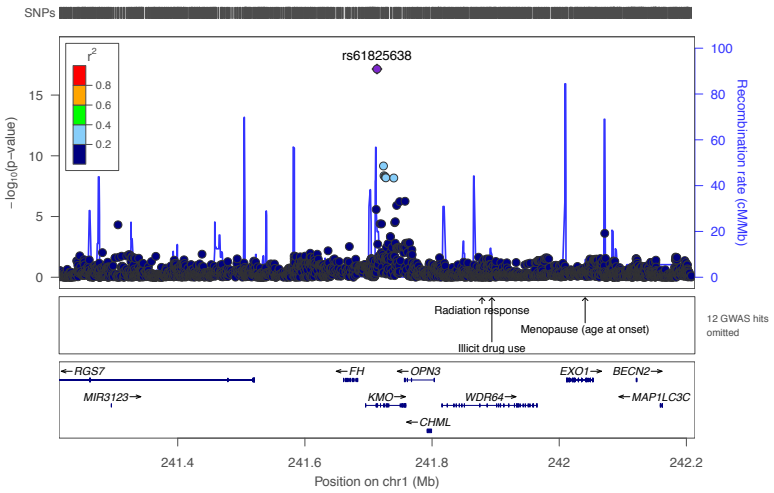

tyrosine
SNPs
rs11133665
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
Prostate cancer
Glioma
Melanoma
Lung cancer in ever smokers
Breast cancer
Lung adenocarcinor
132 GWAS hits omitted
TPPP
BRD9
NKD2
CTD−3080P12.3
CLPTM1L
LPCAT1
LOC728613
ZDHHC11B
TRIP13
SLC12A7
SLC6A19
LINC01511
MIR6075
ZDHHC11
LOC100506688
SLC6A18
SLC6A3
SDHAP3
MIR4635
TERT
MIR4457
0.8
1
1.2
1.4
1.6
Position on chr5 (Mb)


lysine
SNPs
rs4805801
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
Bipolar disorder
Eosinophilic esophagitis
Urinary metabolites
Colorectal cancer
Bone mineral density
Bone properties (heel
37 GWAS hits omitted
ZNF507
PDCD5
NUDT19
SLC7A9
FAAP24
GPATCH1
LOC400684
ANKRD27
TDRD12
CEP89
WDR88
DPY19L3
SNORA68B
RHPN2
RGS9BP
32.8
33
33.2
33.4
33.6
Position on chr19 (Mb)

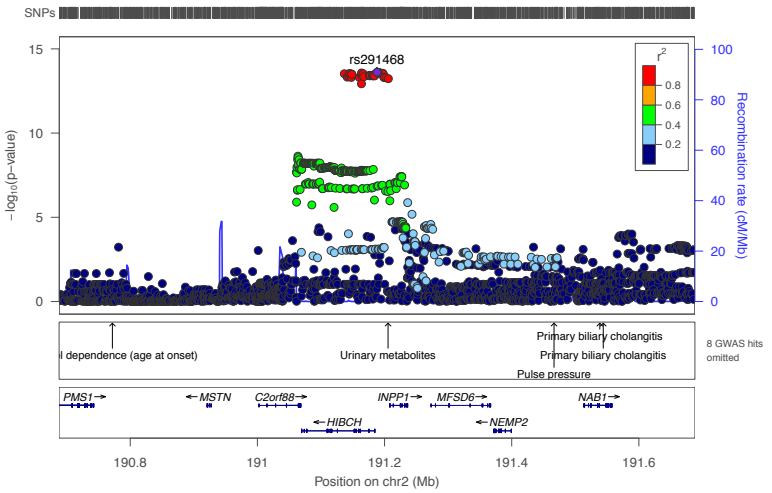
2'−deoxyguanosine



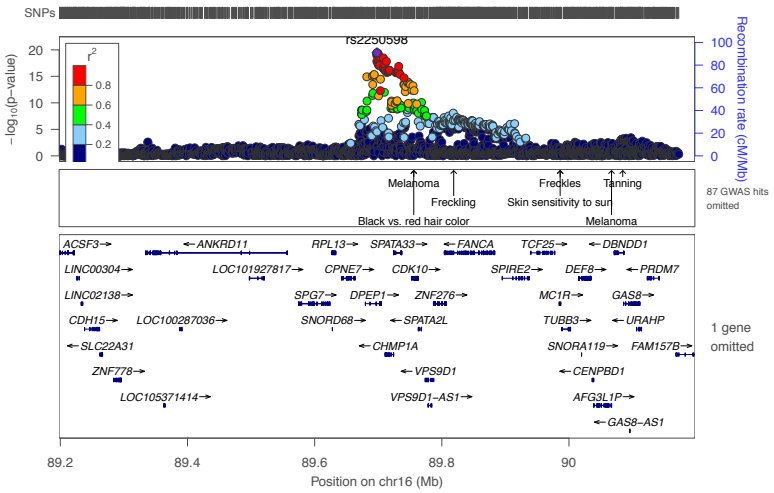
kynurenate



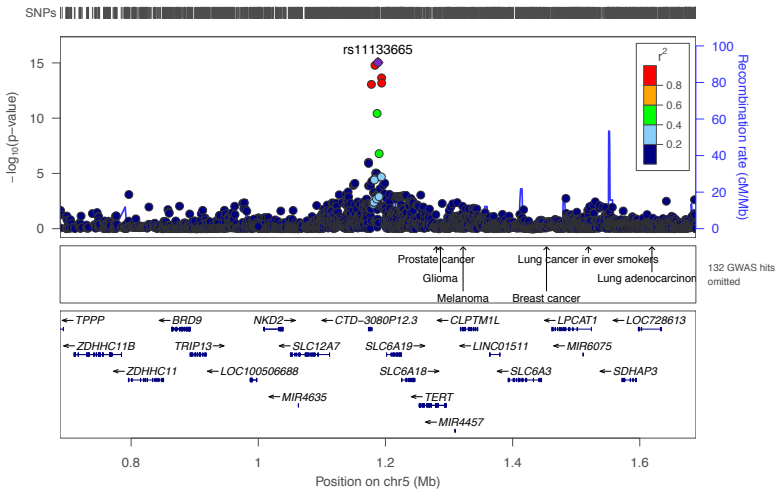
methylmalonate (MMA)



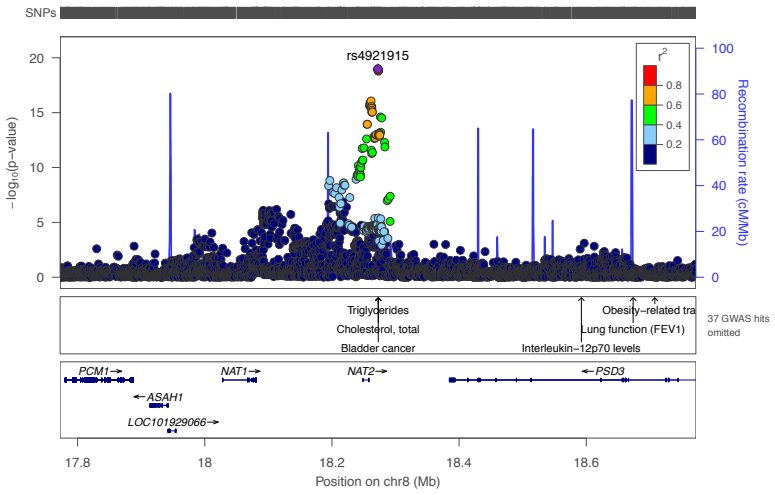
picolinate



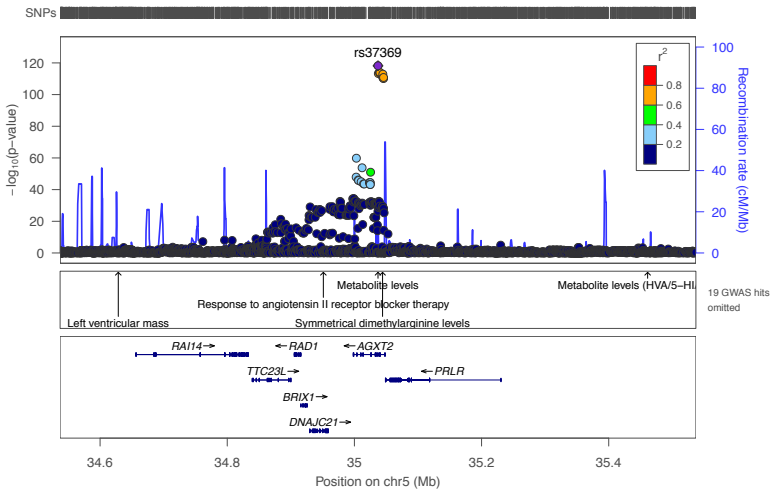
kynurenine



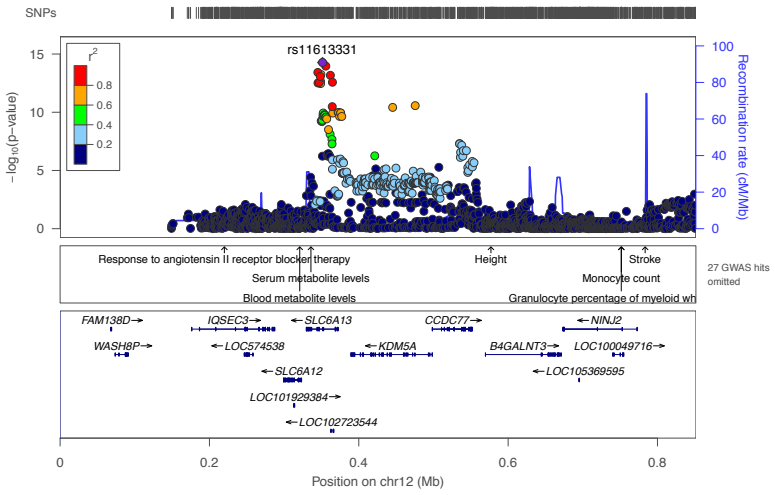
4−acetamidobutanoate



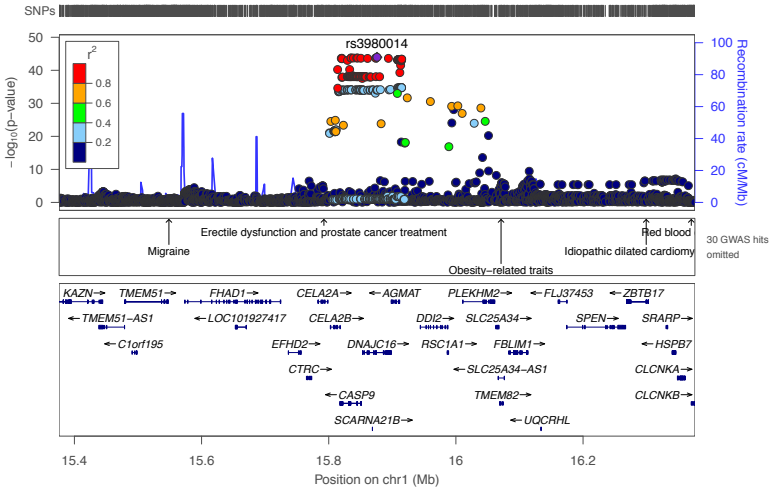
3−aminoisobutyrate



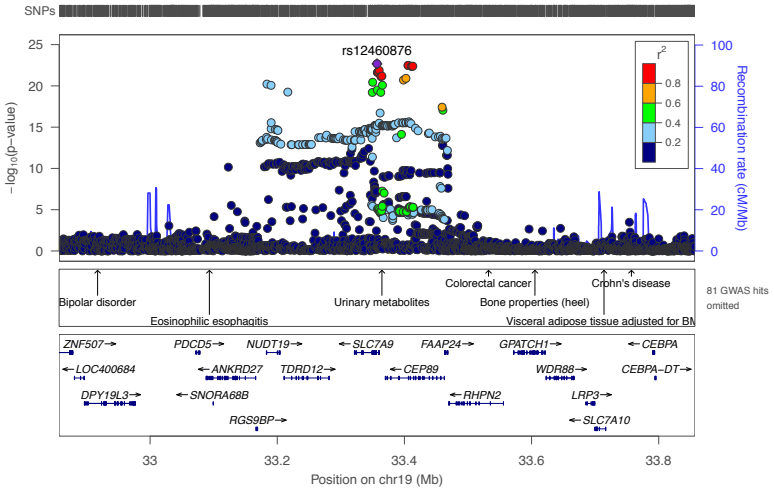
3−aminoisobutyrate



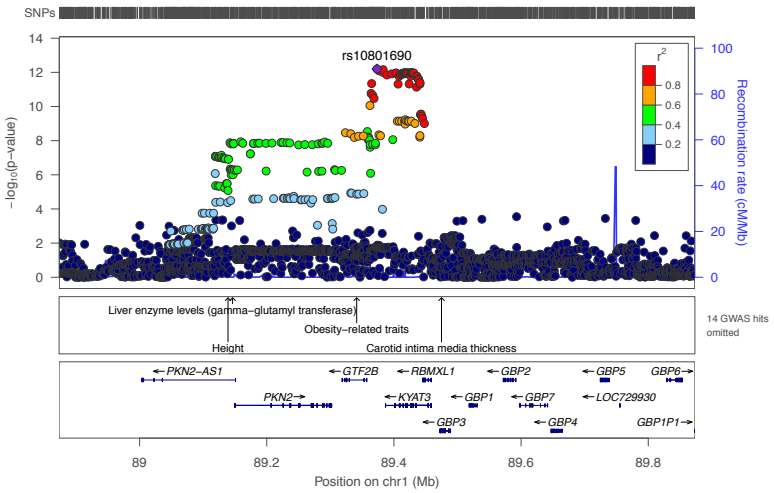
4−guanidinobutanoate



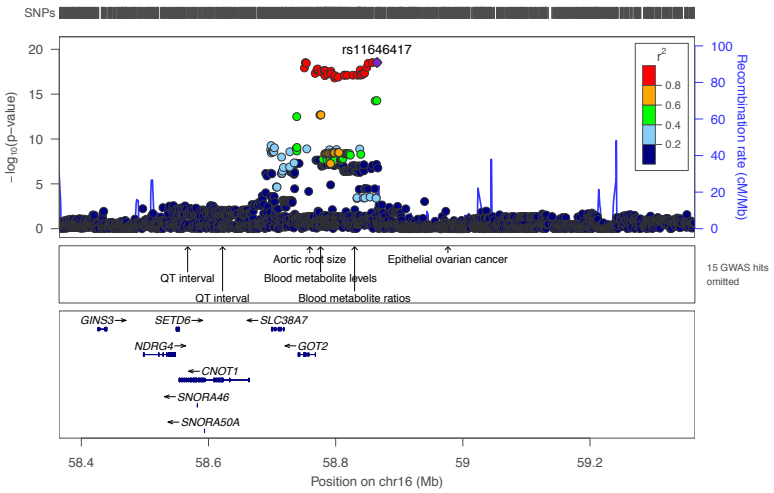
5−hydroxylysine



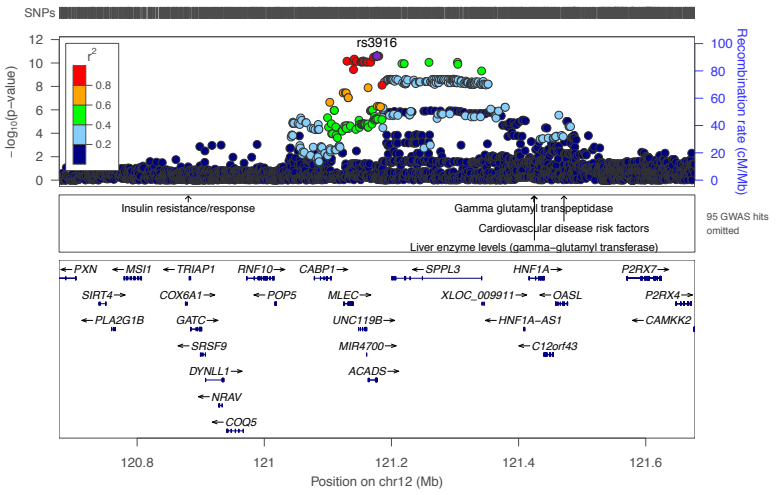
imidazole lactate



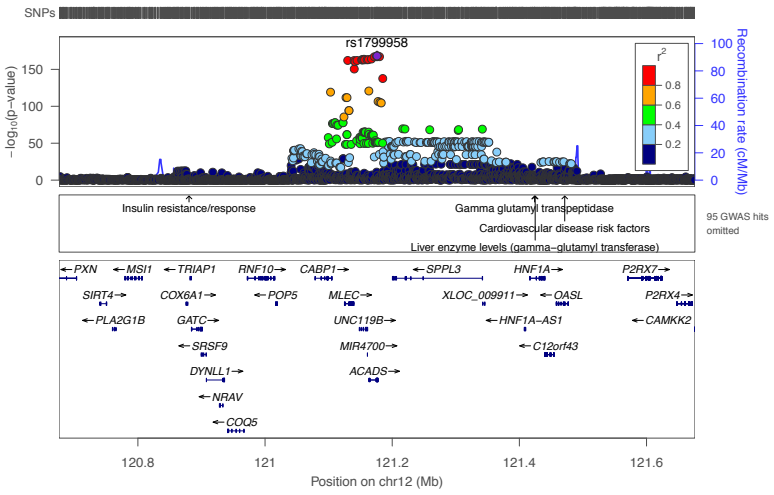
imidazole lactate



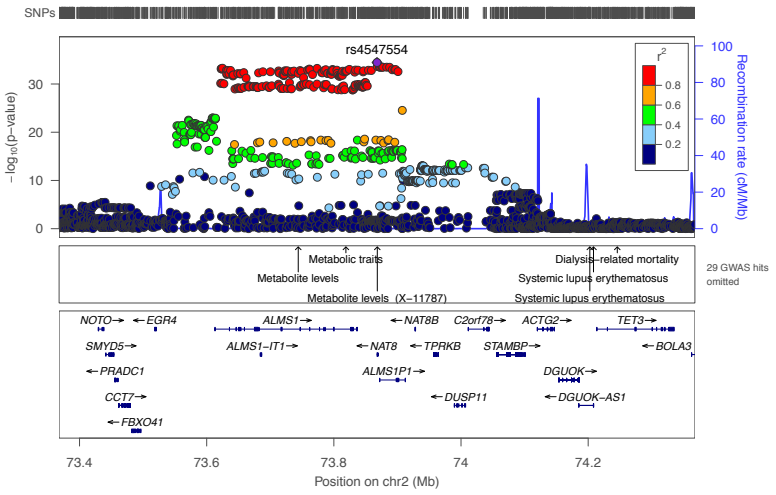
methylsuccinate



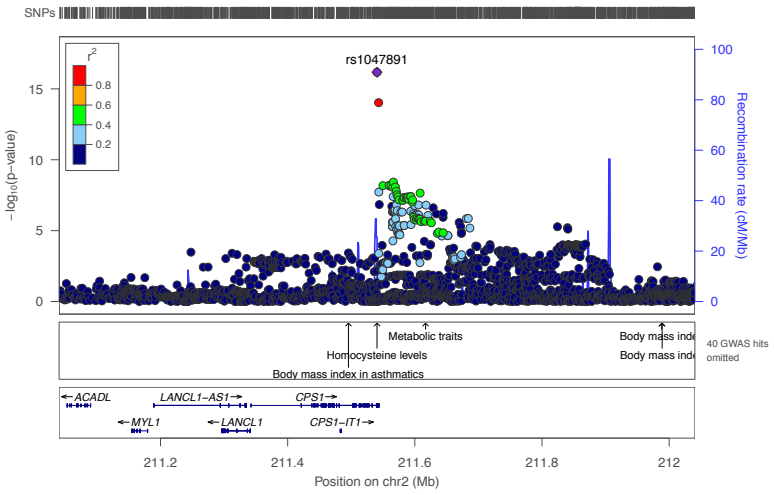
ethylmalonate



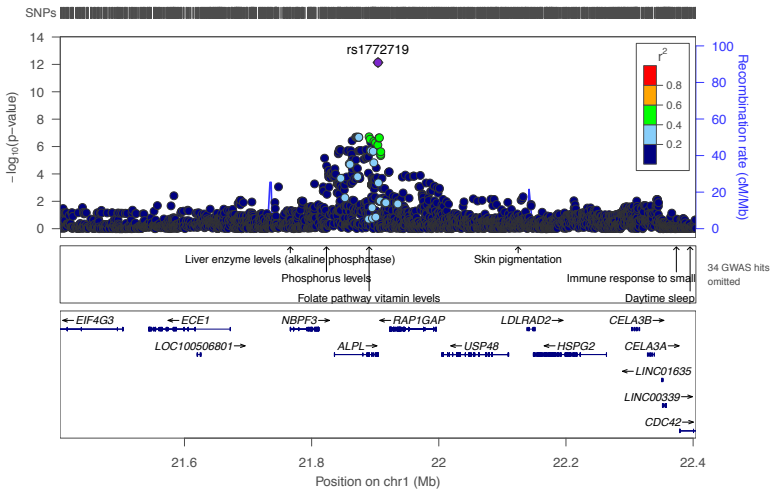
N−acetylleucine



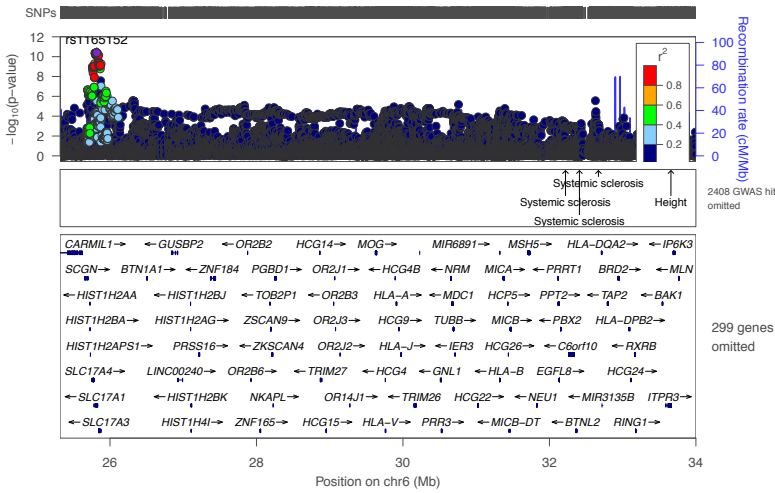
tigloylglycine



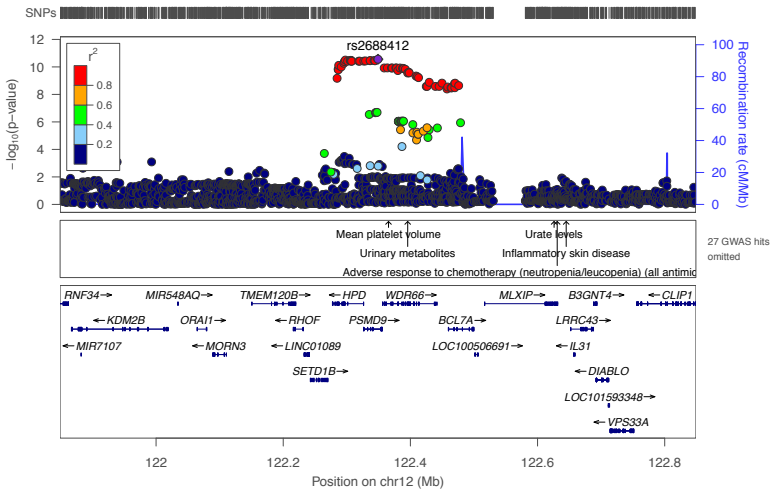
phosphoethanolamine



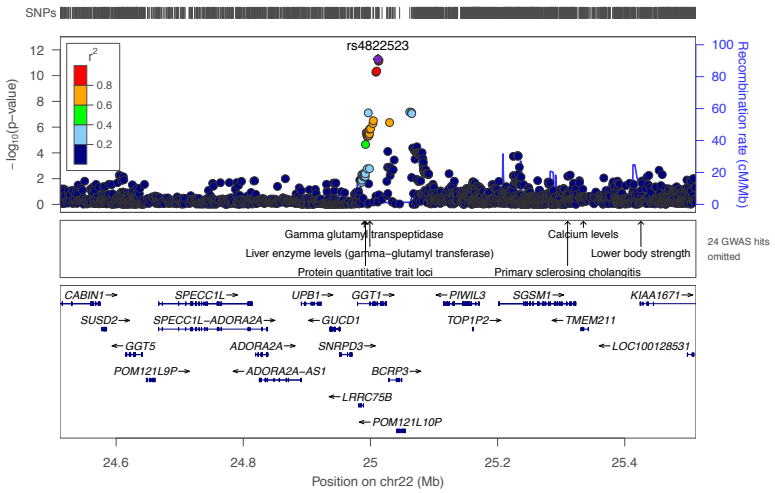
pyridoxal



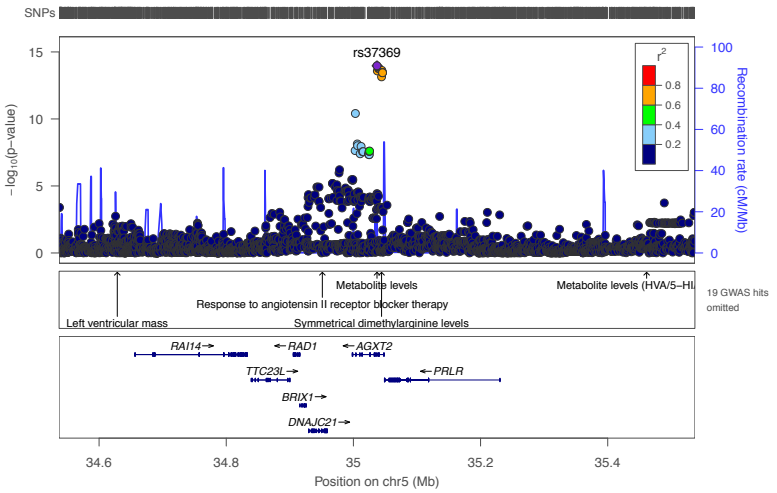
4−hydroxyphenylpyruvate



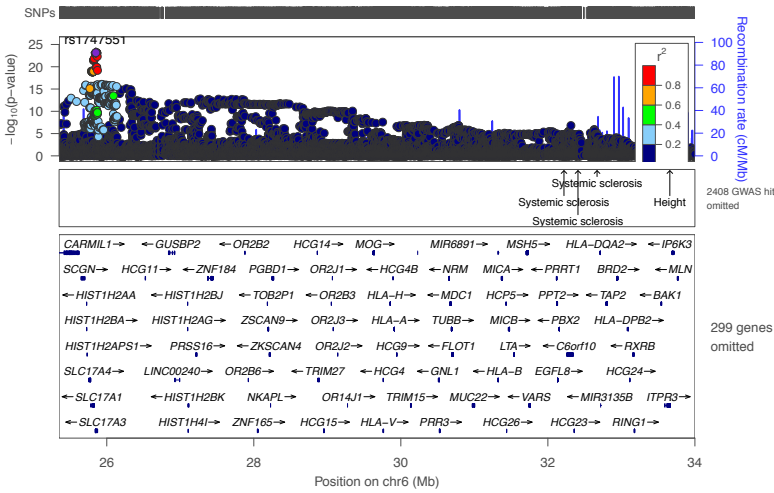
gamma−glutamylhistidine



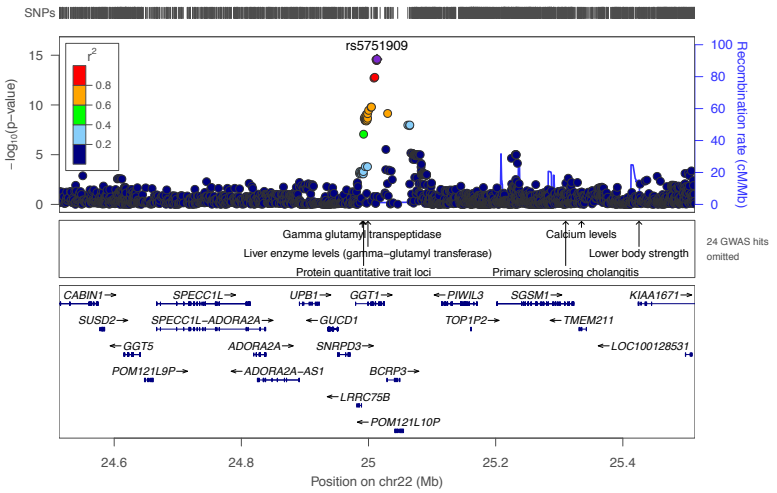
5–aminovalerate



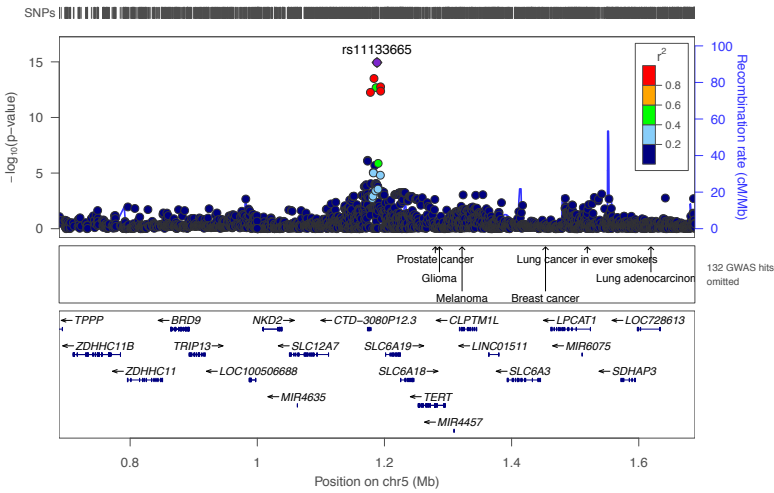
indolelactate



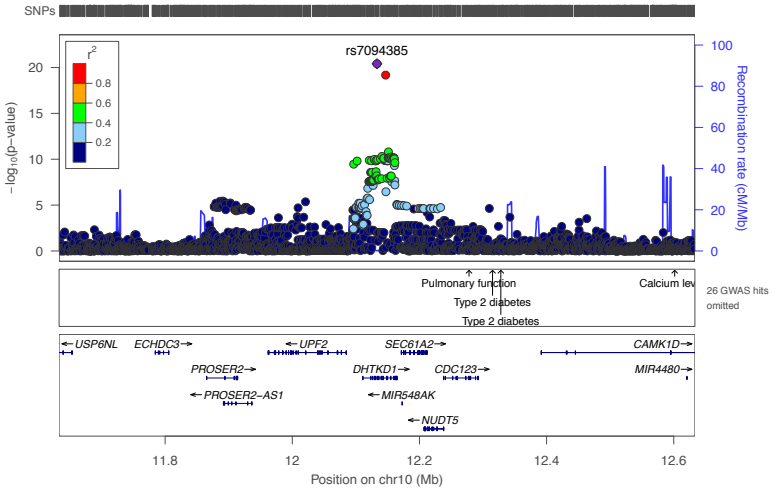
gamma–glutamylleucine



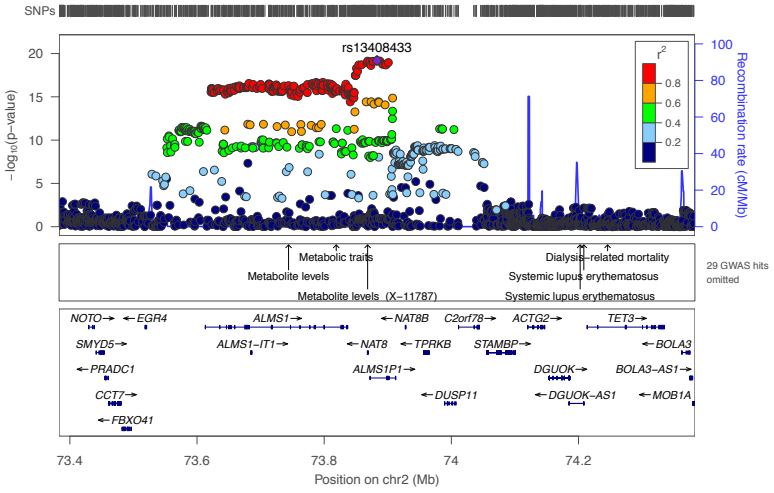
3–hydroxykynurenine



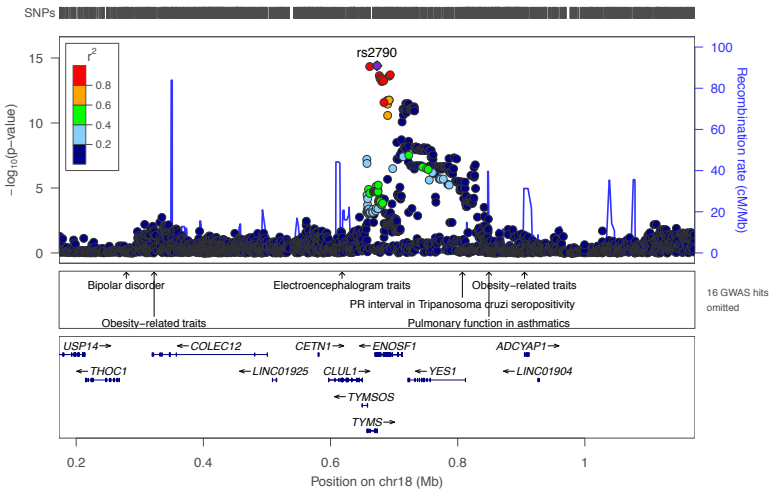
4−ureidobutyrate



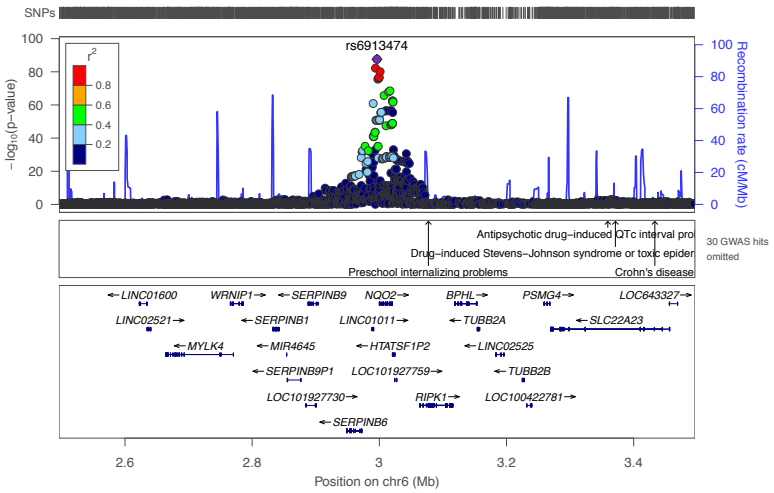
homocitrulline



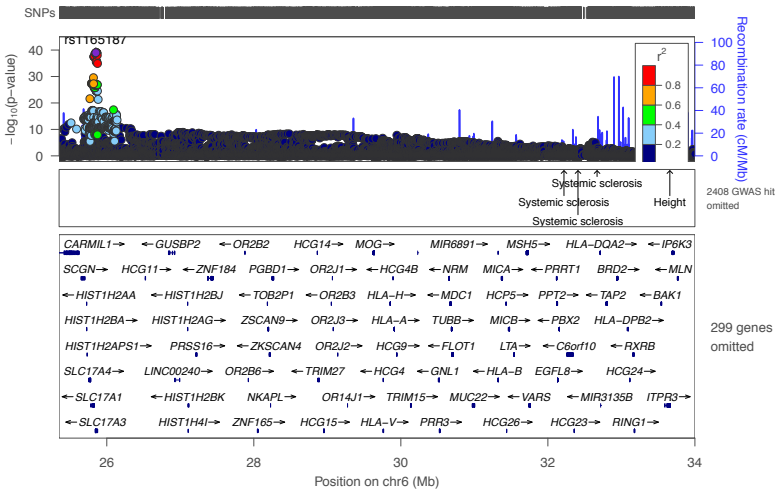
ribonate



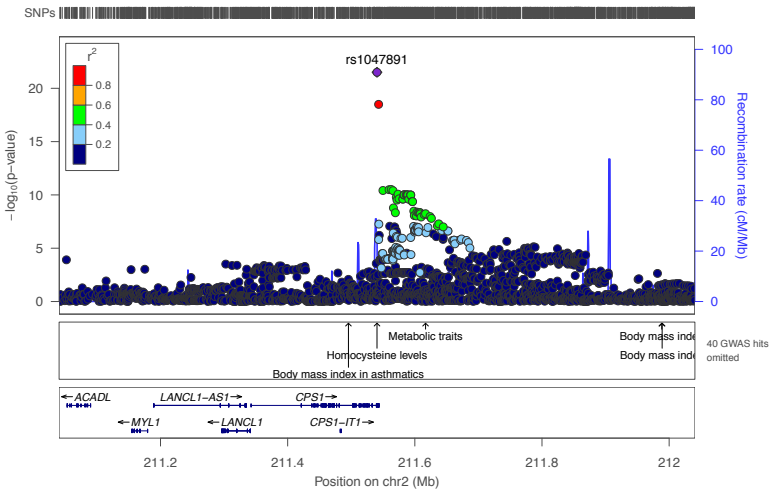
isoxanthopterin



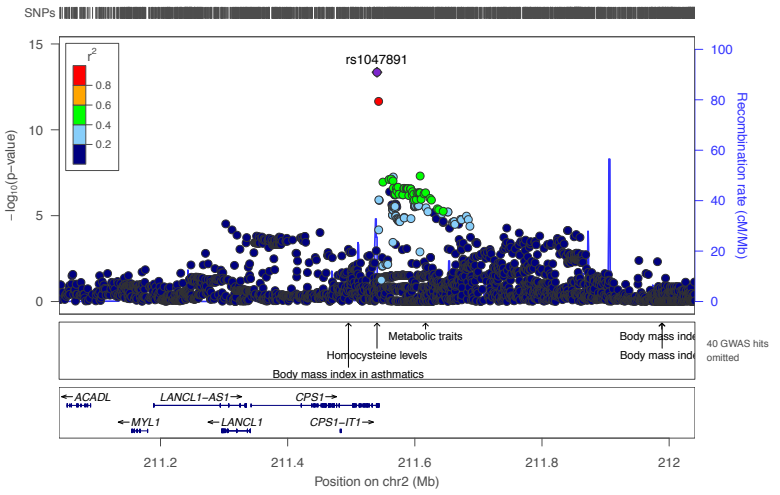
adenosine 3',5'–cyclic monophosphate (cAMP)



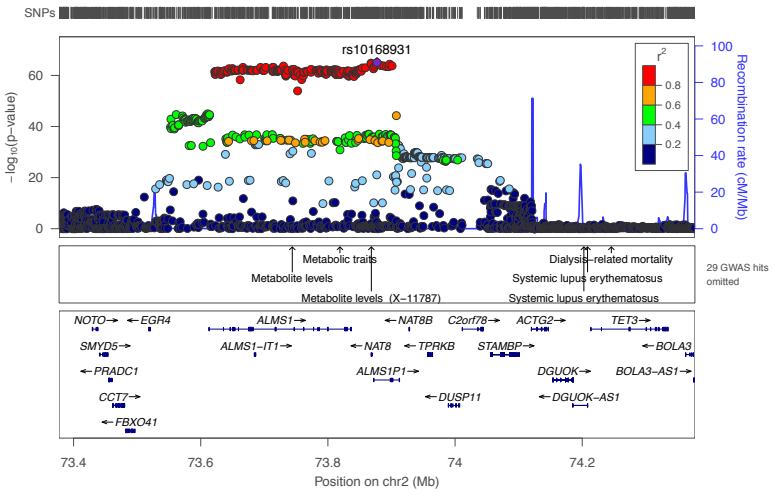
2–methylbutyrylglycine



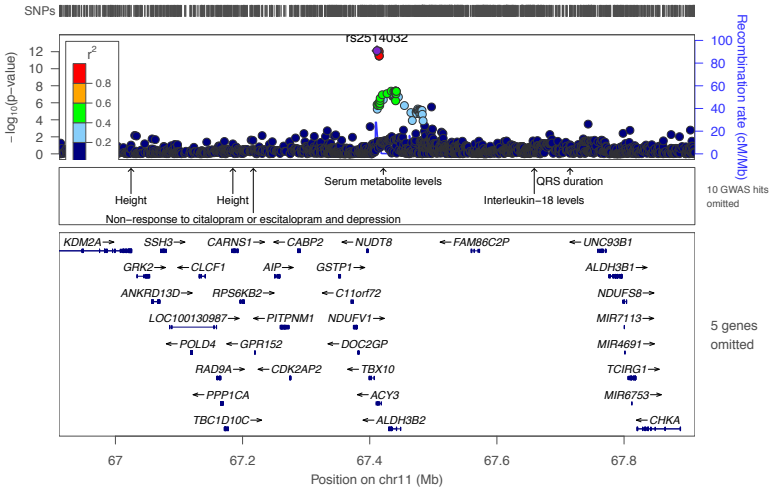
3–methylcrotonylglycine



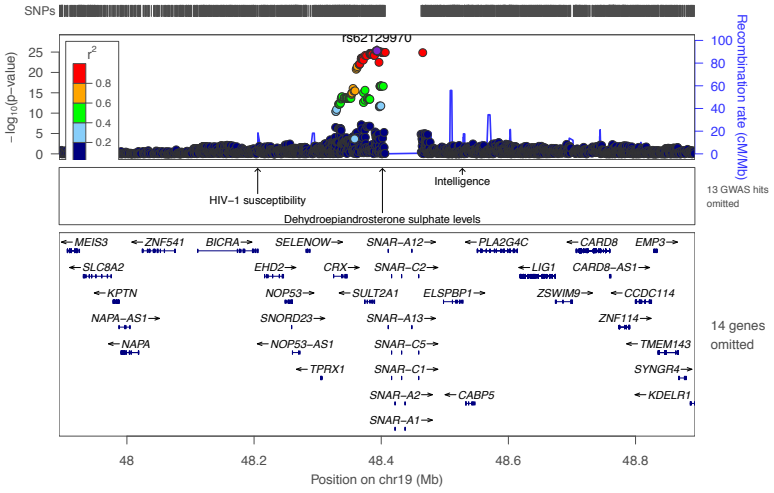
N–acetyltyrosine



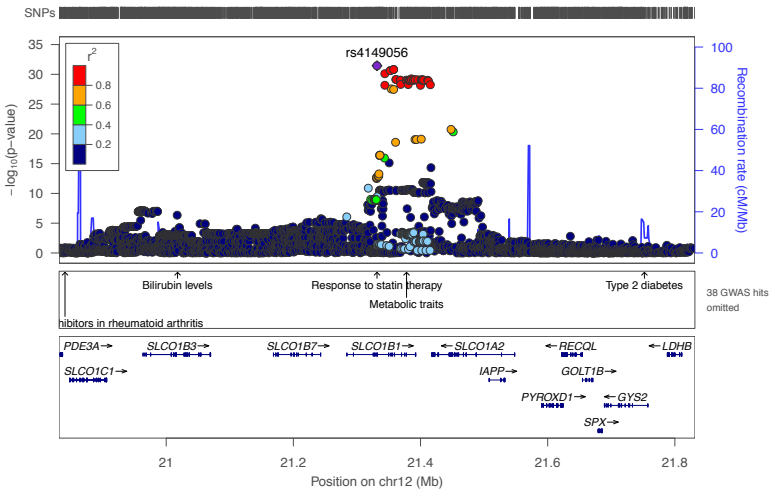
N−acetyltyrosine



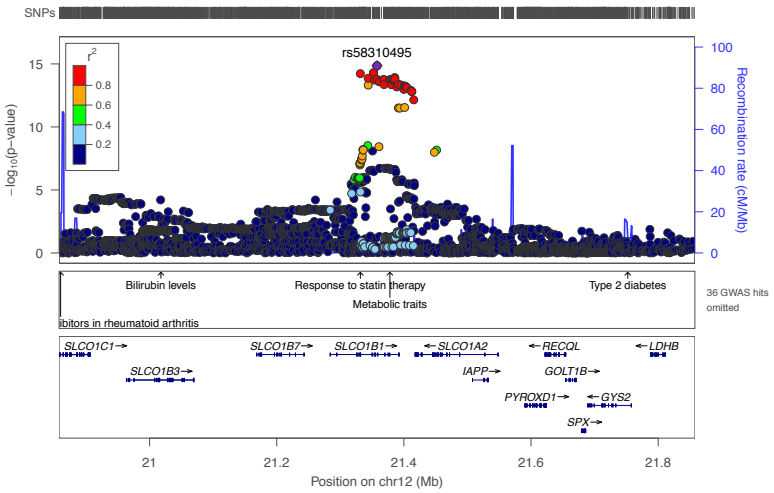
pregnenediol disulfate (C21H34O8S2)*



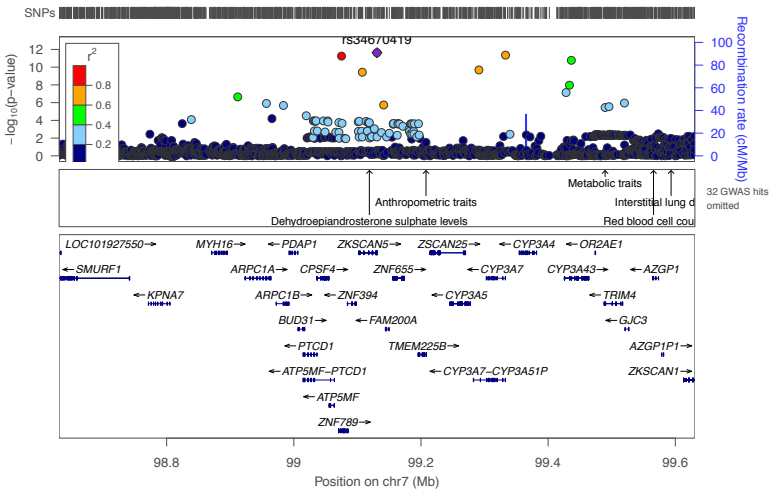
glycocholenate sulfate*



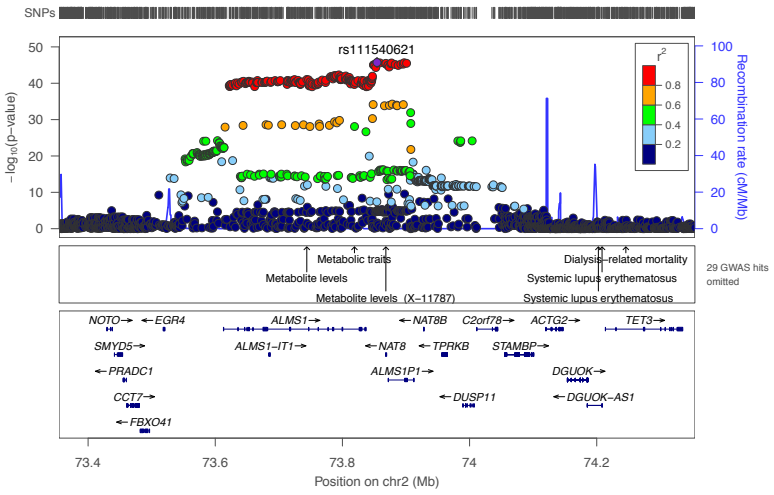
taurocholenate sulfate



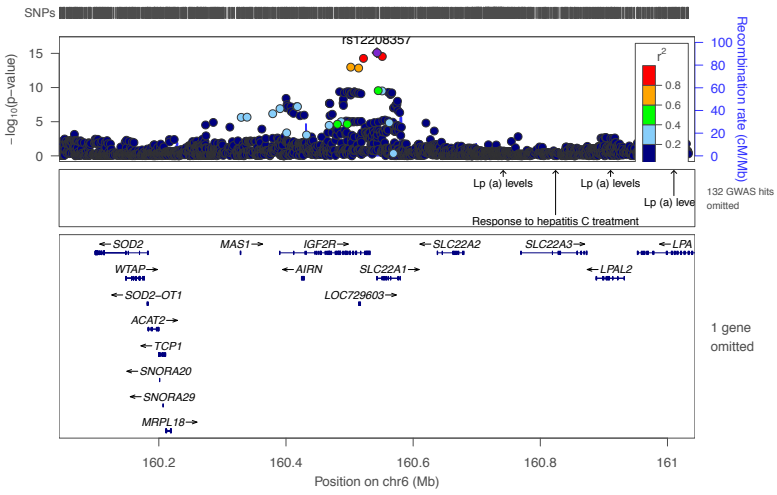
andro steroid monosulfate C19H28O6S (1)*



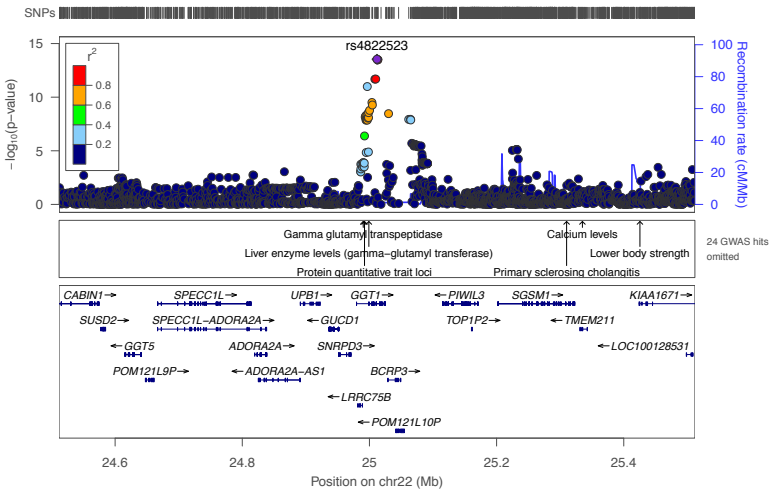
N–alpha–acetylornithine

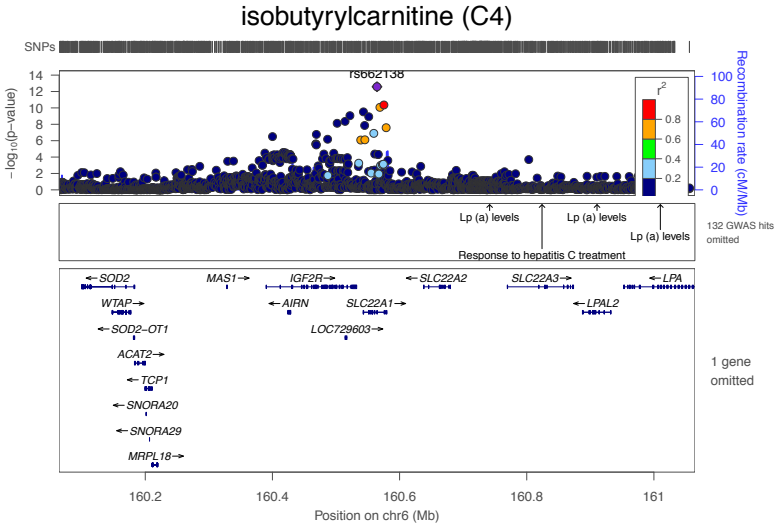


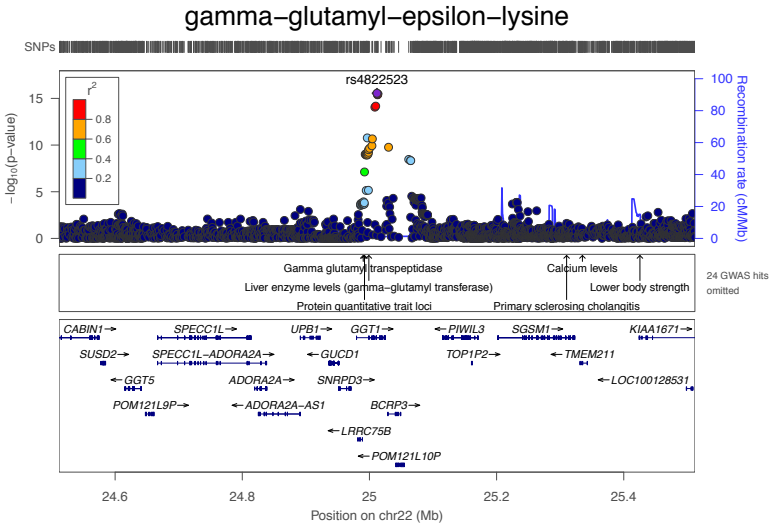
nicotinamide riboside

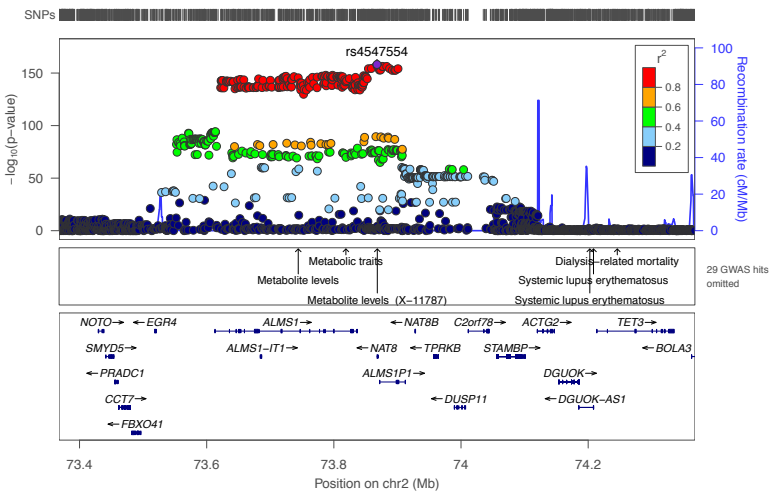


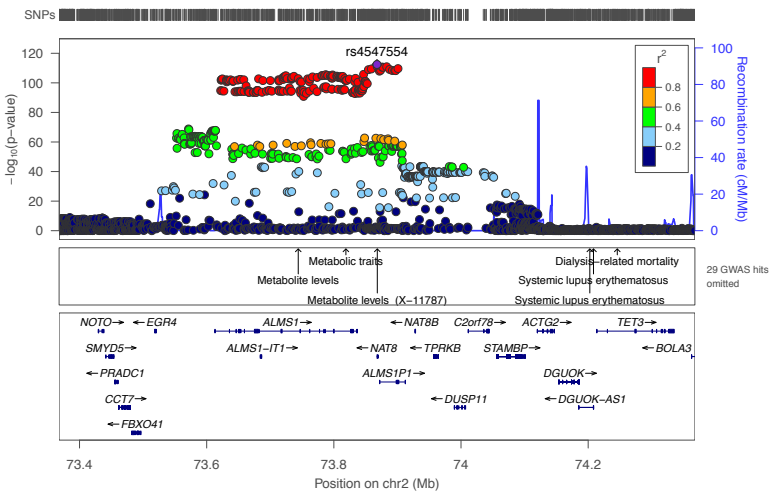
gamma–glutamylphenylalanine



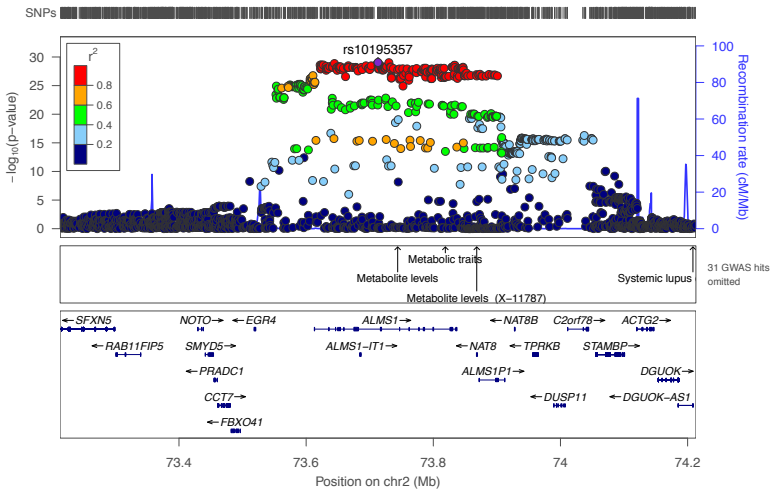

isobutyrylcarnitine (C4)
SNPs
rs662138
−log10(p−value)
Recombination rate (cM/Mb)
r2
0.8
0.6
0.4
0.2
Lp (a) levels
Lp (a) levels
Lp (a) levels
Response to hepatitis C treatment
132 GWAS hits omitted
SOD2
MAS1
IGF2R
SLC22A2
SLC22A3
LPA
WTAP
AIRN
SLC22A1
LPAL2
SOD2−OT1
LOC729603
ACAT2
TCP1
SNORA20
SNORA29
MRPL18
1 gene omitted
160.2
160.4
160.6
160.8
161
Position on chr6 (Mb)


gamma−glutamyl−epsilon−lysine
SNPs
rs4822523
−log10(p−value)
Recombination rate (cM/Mb)
r2
0.8
0.6
0.4
0.2
Gamma glutamyl transpeptidase
Liver enzyme levels (gamma−glutamyl transferase)
Protein quantitative trait loci
Calcium levels
Lower body strength
Primary sclerosing cholangitis
24 GWAS hits omitted
CABIN1
SPECC1L
UPB1
GGT1
PIWIL3
SGSM1
KIAA1671
SUSD2
SPECC1L−ADORA2A
GUCD1
TOP1P2
TMEM211
GGT5
ADORA2A
SNRPD3
LOC100128531
POM121L9P
ADORA2A−AS1
BCRP3
LRRC75B
POM121L10P
24.6
24.8
25
25.2
25.4
Position on chr22 (Mb)

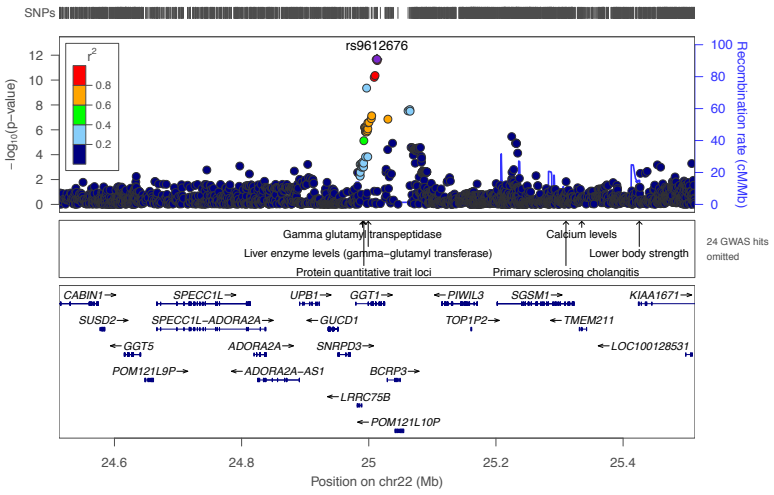

N−acetylasparagine
SNPs
rs4547554
−log10(p−value)
Recombination rate (cM/Mb)
r2
0.8
0.6
0.4
0.2
Metabolic traits
Metabolite levels
Metabolite levels (X−11787)
Dialysis−related mortality
Systemic lupus erythematosus
Systemic lupus erythematosus
29 GWAS hits omitted
NOTO
EGR4
ALMS1
NAT8B
C2orf78
ACTG2
TET3
SMYD5
ALMS1−IT1
NAT8
TPRKB
STAMBP
BOLA3
PRADC1
ALMS1P1
DGUOK
CCT7
DUSP11
DGUOK−AS1
FBXO41
73.4
73.6
73.8
74
74.2
Position on chr2 (Mb)


N−acetylglutamine
SNPs
rs4547554
−log10(p−value)
Recombination rate (cM/Mb)
r2
0.8
0.6
0.4
0.2
Metabolic traits
Metabolite levels
Metabolite levels (X−11787)
Dialysis−related mortality
Systemic lupus erythematosus
Systemic lupus erythematosus
29 GWAS hits omitted
NOTO
EGR4
ALMS1
NAT8B
C2orf78
ACTG2
TET3
SMYD5
ALMS1−IT1
NAT8
TPRKB
STAMBP
BOLA3
PRADC1
ALMS1P1
DGUOK
CCT7
DUSP11
DGUOK−AS1
FBXO41
73.4
73.6
73.8
74
74.2
Position on chr2 (Mb)

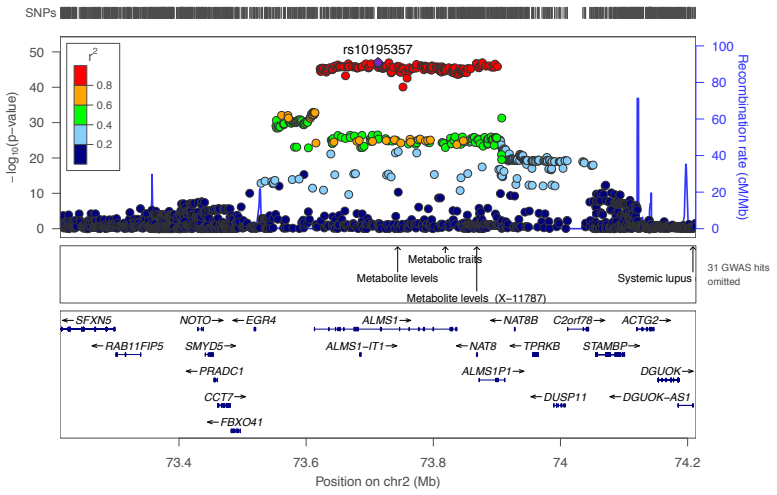
N−acetylhistidine



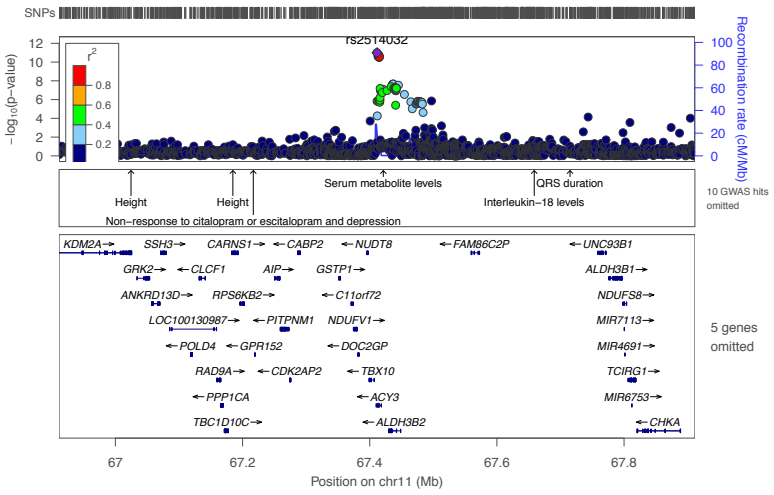
gamma−glutamylglycine



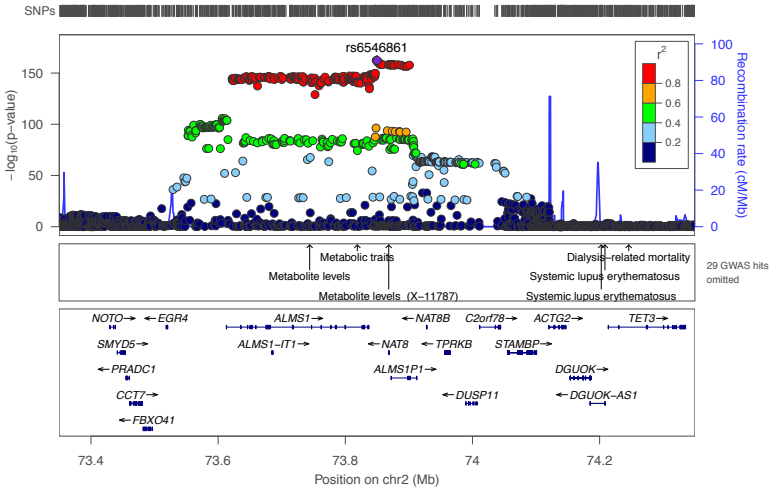
N−acetylphenylalanine



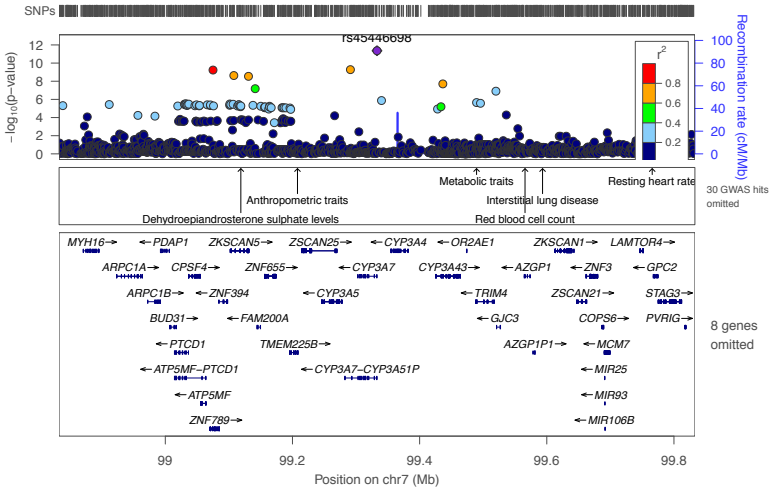
N−acetylphenylalanine



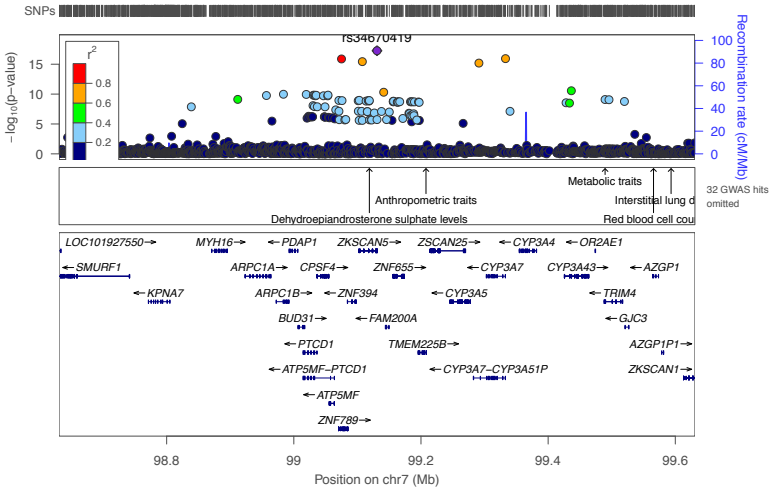
N−acetylarginine



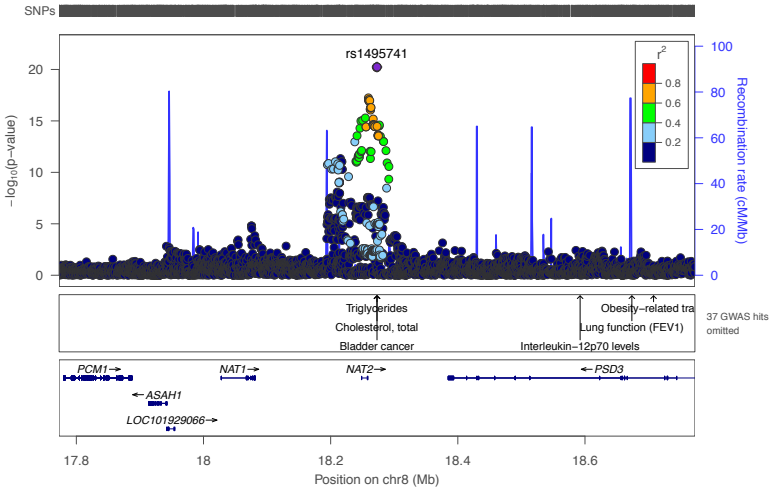
epiandrosterone sulfate



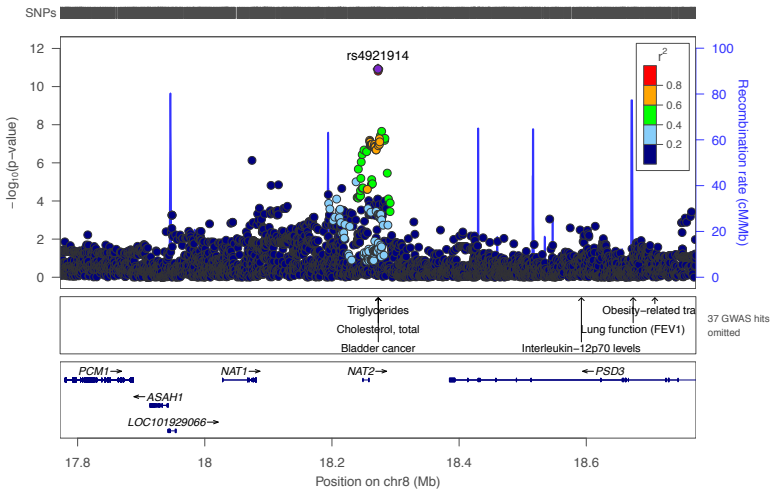
tauro−beta−muricholate



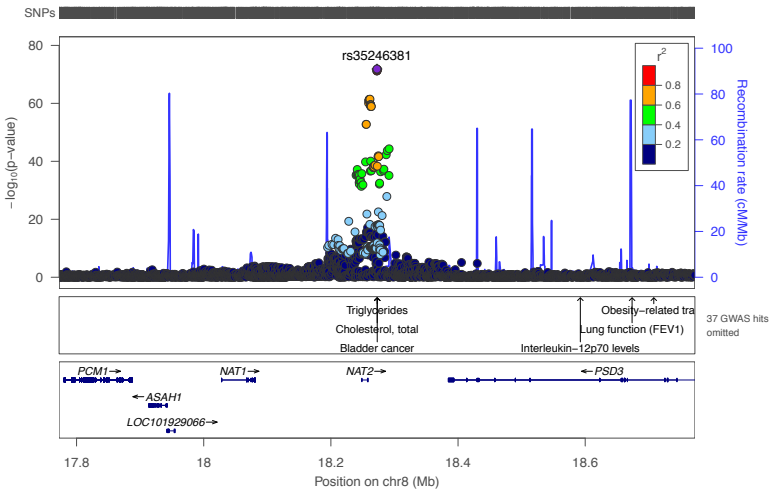
1−methylxanthine



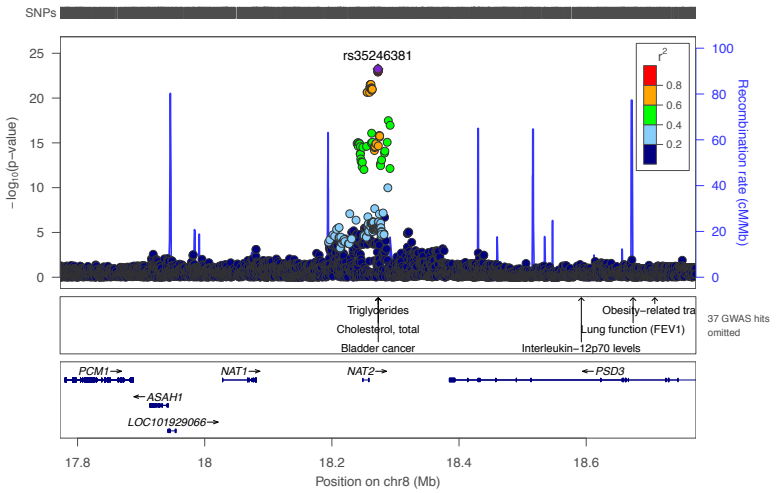
1−methylurate



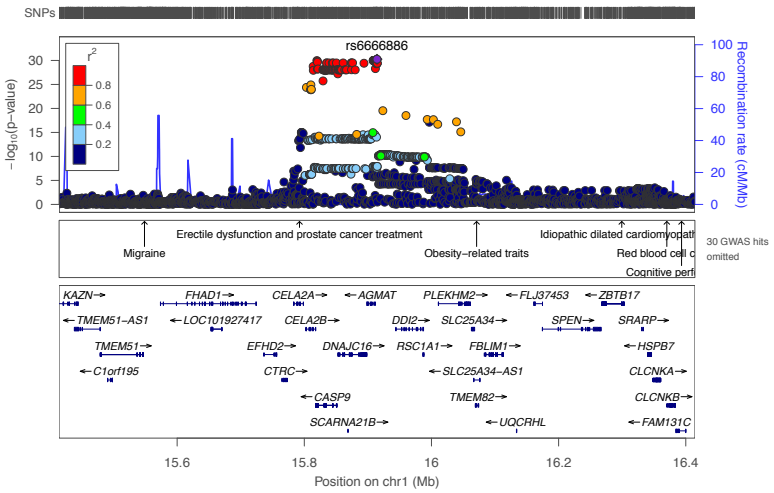
5−acetylamino−6−formylamino−3−methyluracil



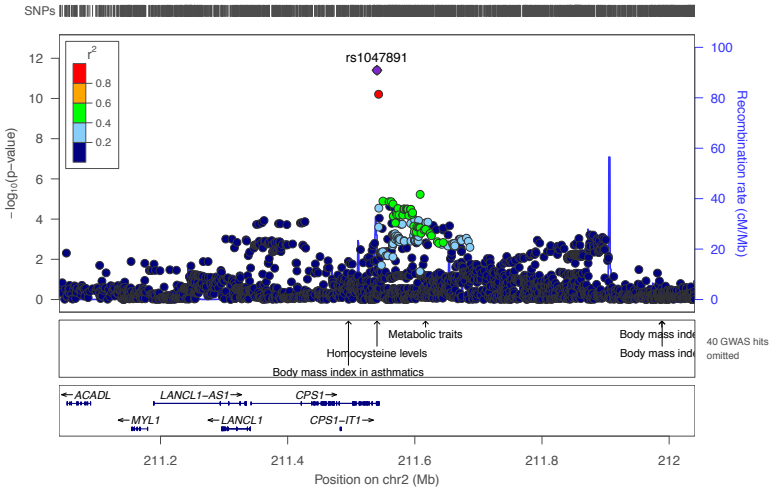
5−acetylamino−6−amino−3−methyluracil



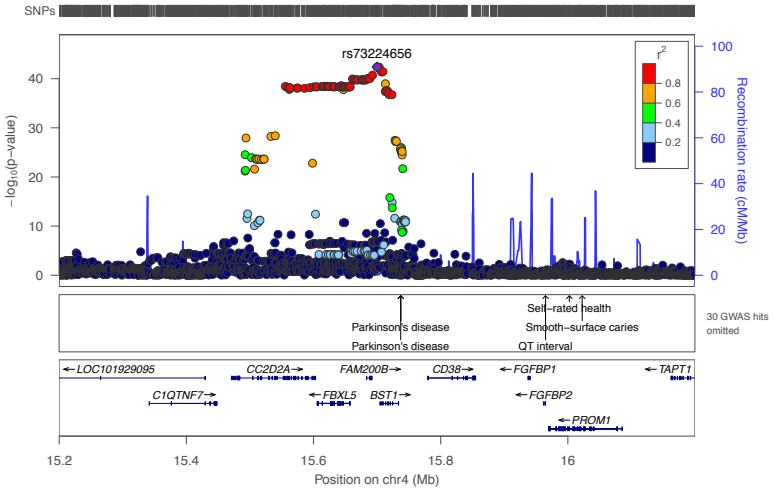
beta−guanidinopropanoate



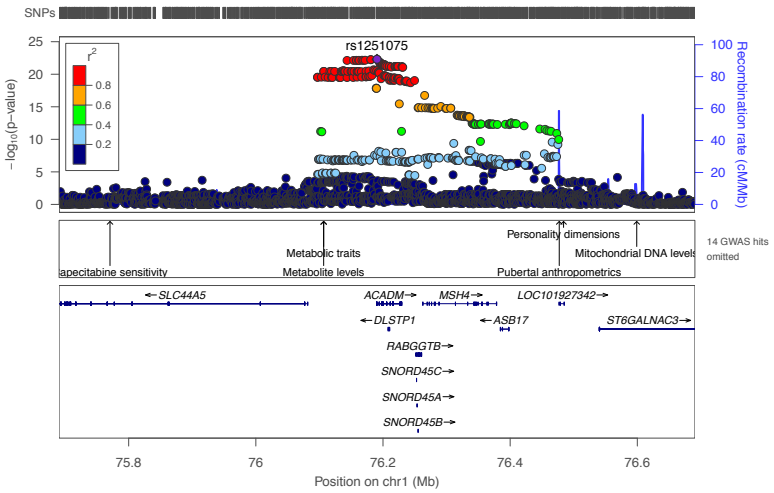
isovalerylglycine



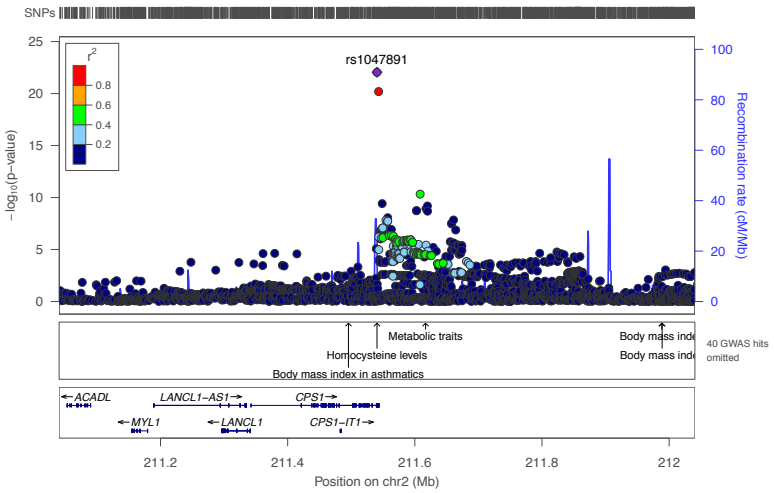
3−methylcytidine



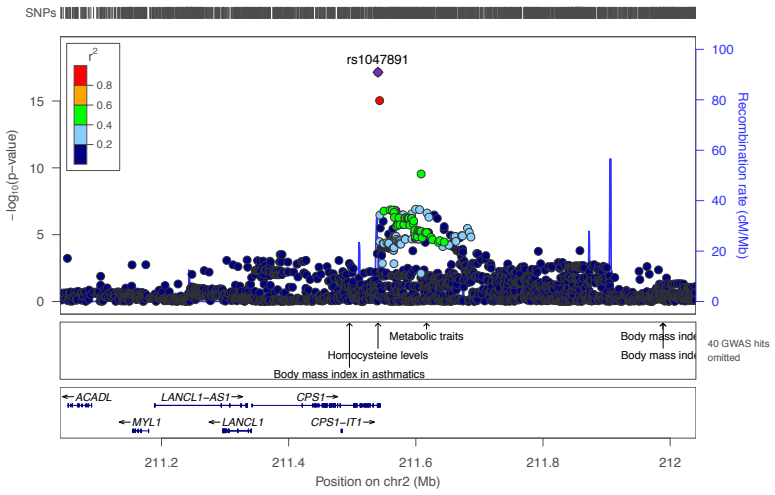
hexanoylglycine



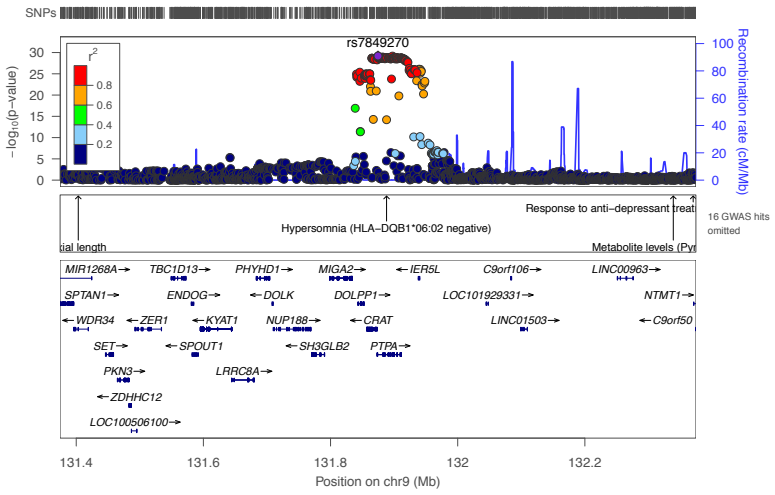
hexanoylglycine



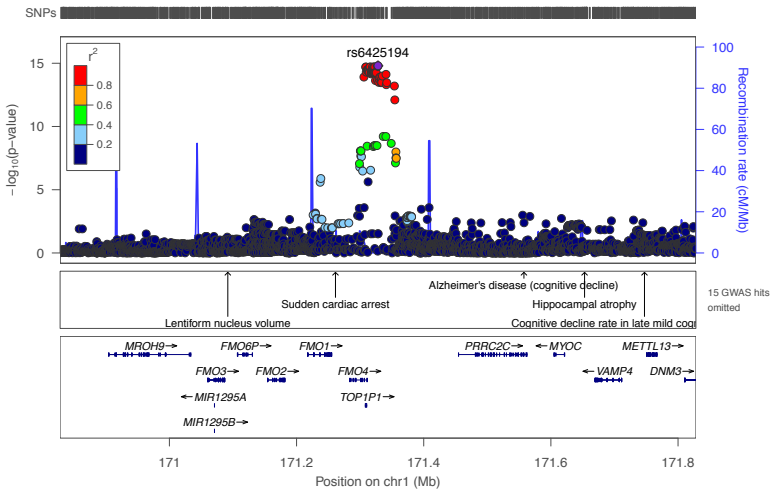
isobutyrylglycine



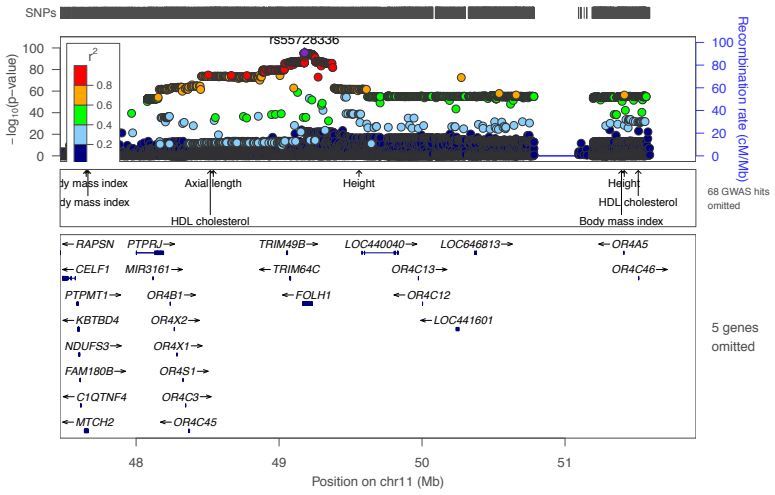
2−methylmalonylcarnitine (C4−DC)

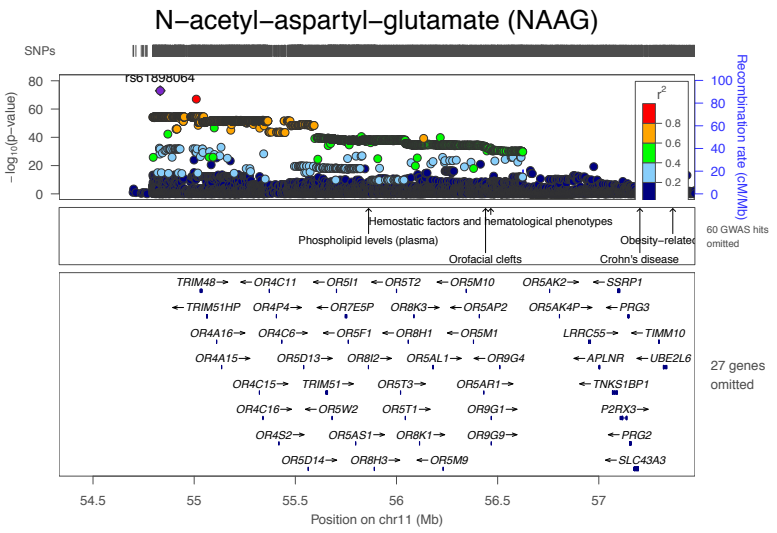


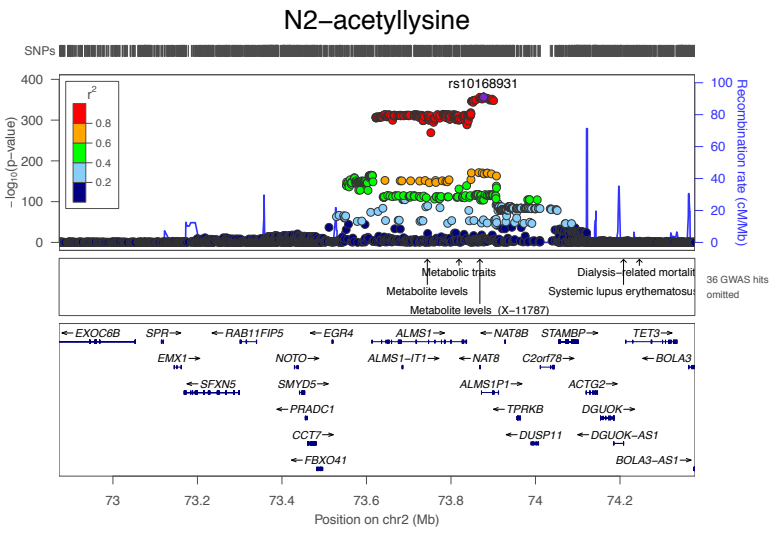
cysteinylglycine

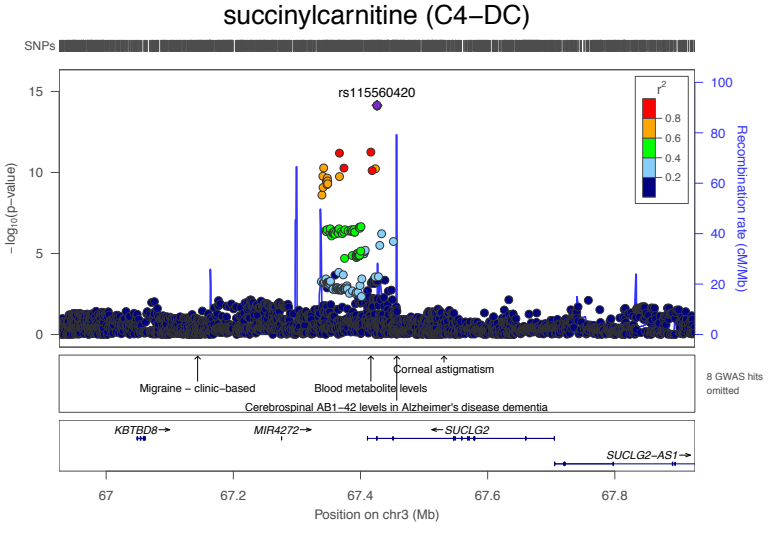


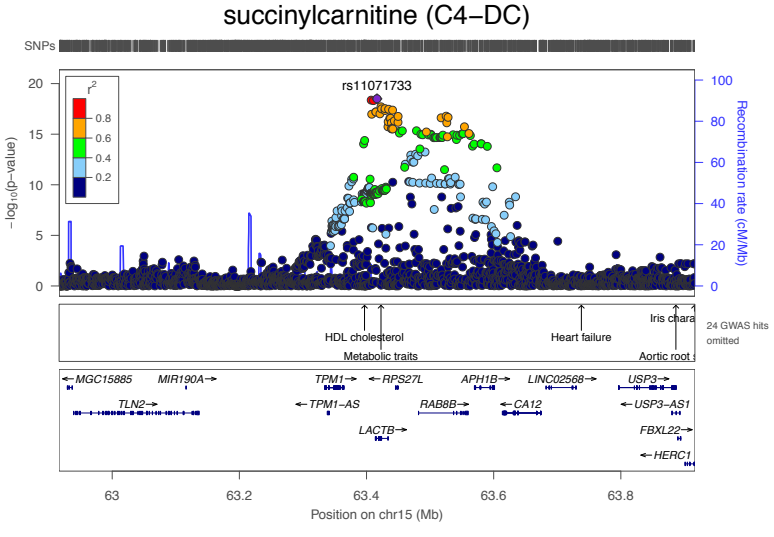
N−acetyl−aspartyl−glutamate (NAAG)



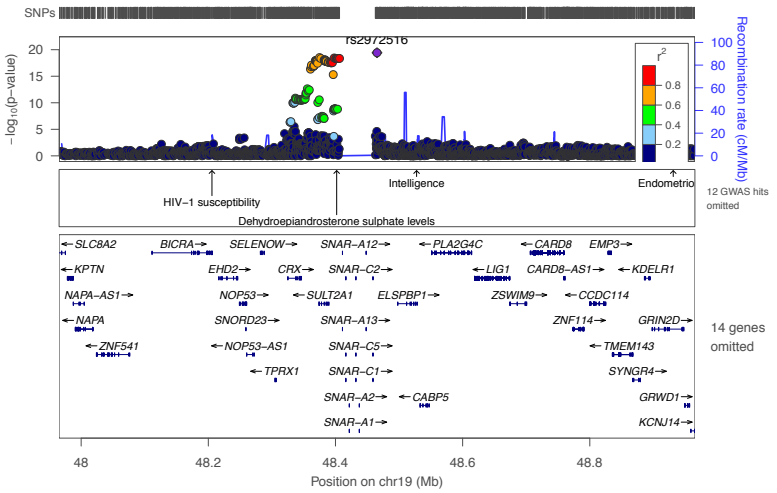

N−acetyl−aspartyl−glutamate (NAAG)
SNPs
rs61898064
−log10(p−value)
Recombination rate (cM/Mb)
r2
Hemostatic factors and hematological phenotypes
Phospholipid levels (plasma)
Orofacial clefts
Obesity−related
Crohn's disease
60 GWAS hits omitted
27 genes omitted
Position on chr11 (Mb)


N2−acetyllysine
SNPs
rs10168931
−log10(p−value)
Recombination rate (cM/Mb)
r2
Metabolic traits
Metabolite levels
Metabolite levels (X−11787)
Dialysis−related mortalit
Systemic lupus erythematosu
36 GWAS hits omitted
Position on chr2 (Mb)

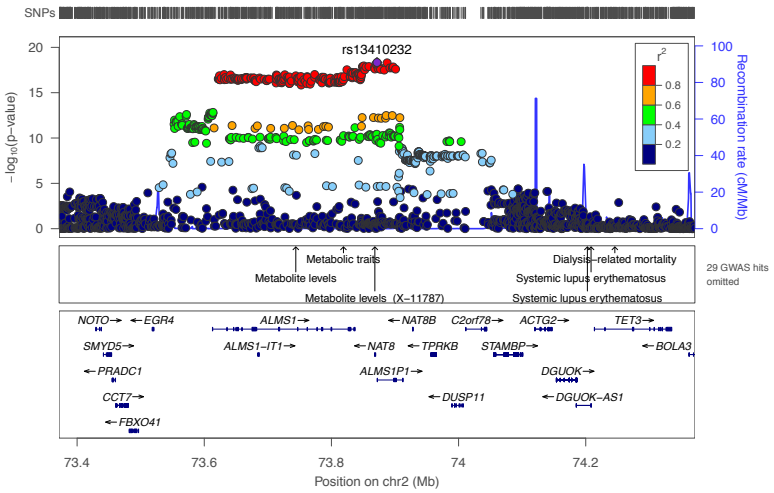

succinylcarnitine (C4−DC)
SNPs
rs115560420
−log10(p−value)
Recombination rate (cM/Mb)
r2
Migraine − clinic−based
Blood metabolite levels
Corneal astigmatism
Cerebrospinal AB1−42 levels in Alzheimer's disease dementia
8 GWAS hits omitted
Position on chr3 (Mb)


succinylcarnitine (C4−DC)
SNPs
rs11071733
−log10(p−value)
Recombination rate (cM/Mb)
r2
HDL cholesterol
Metabolic traits
Heart failure
24 GWAS hits omitted
Position on chr15 (Mb)

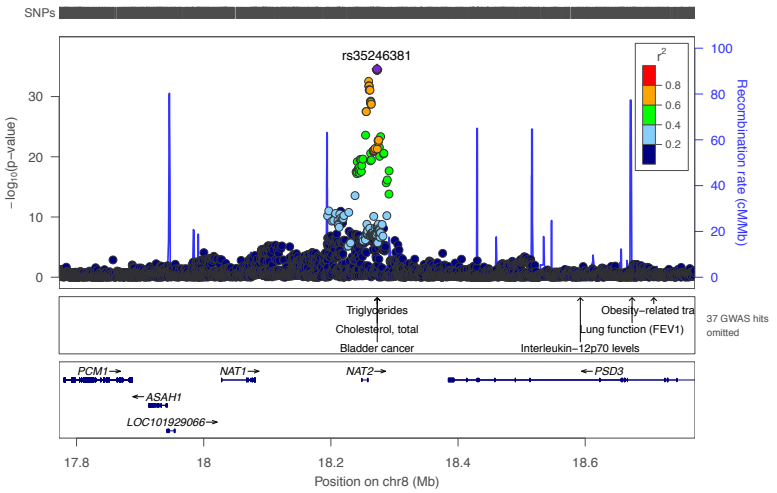
androstenediol (3beta,17beta) disulfate (1)



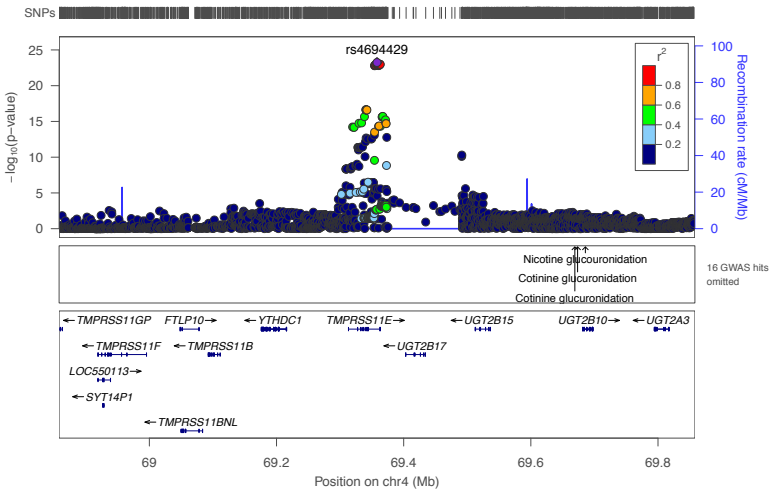
N6−carboxyethyllysine



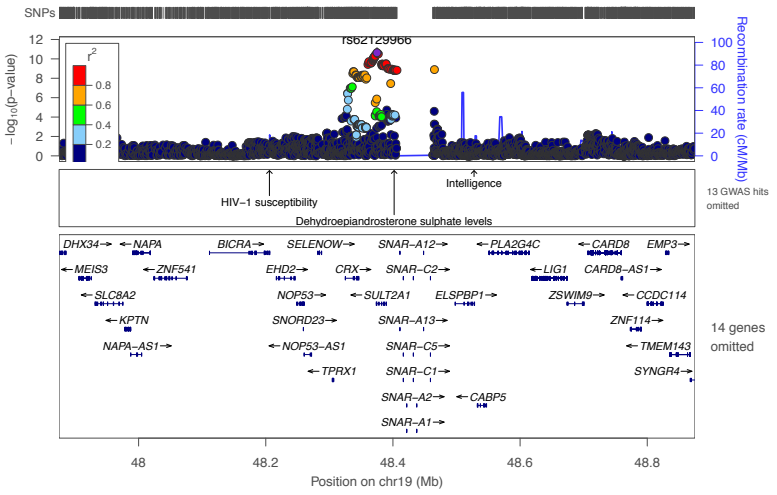
N−acetylputrescine



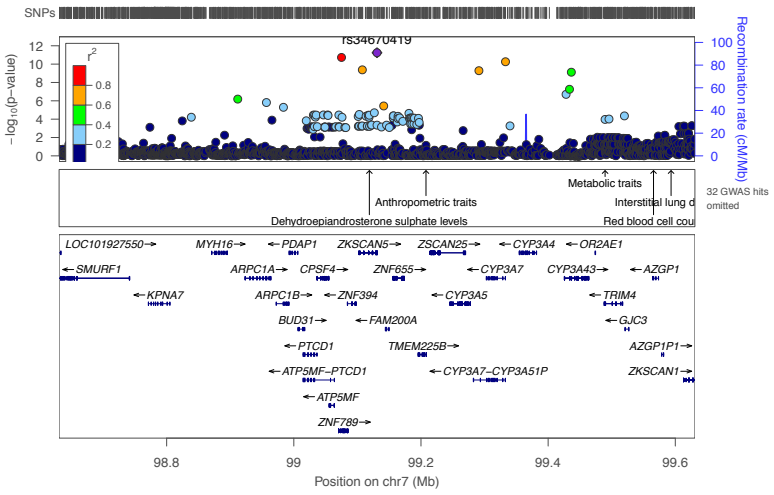
11−ketoetiocholanolone sulfate



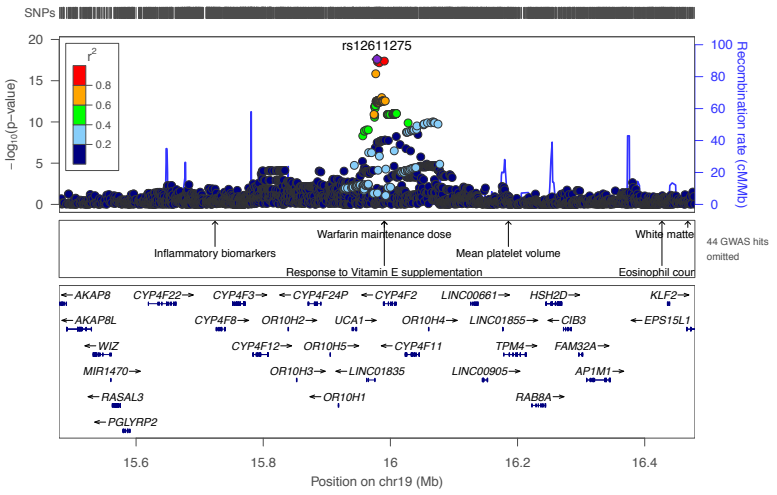
11−ketoetiocholanolone sulfate



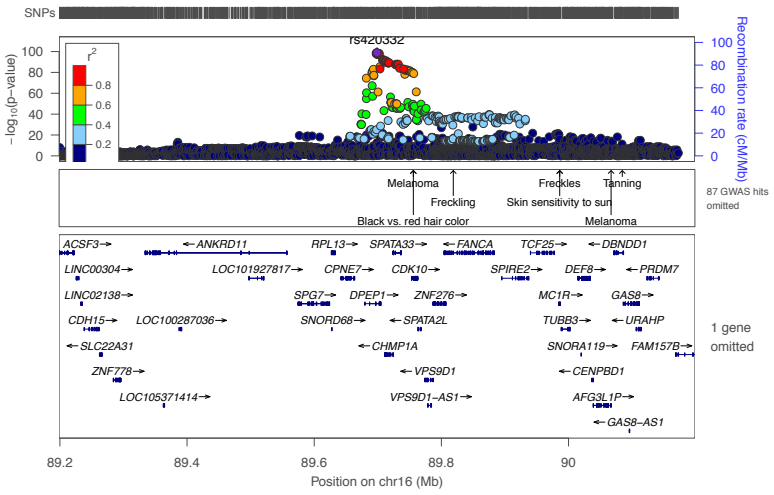
16a−hydroxy DHEA 3−sulfate



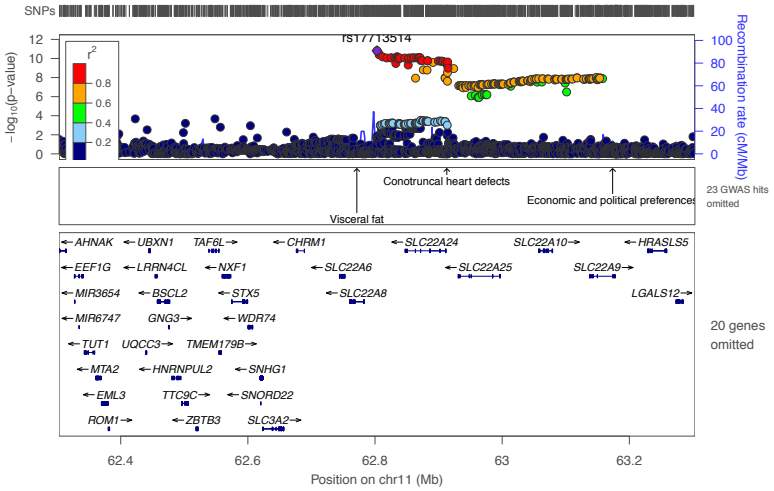
alpha−CEHC glucuronide*



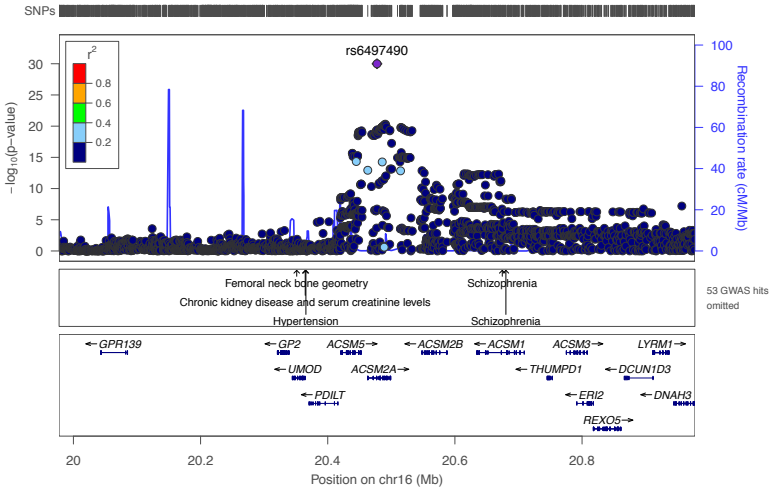
prolylglycine



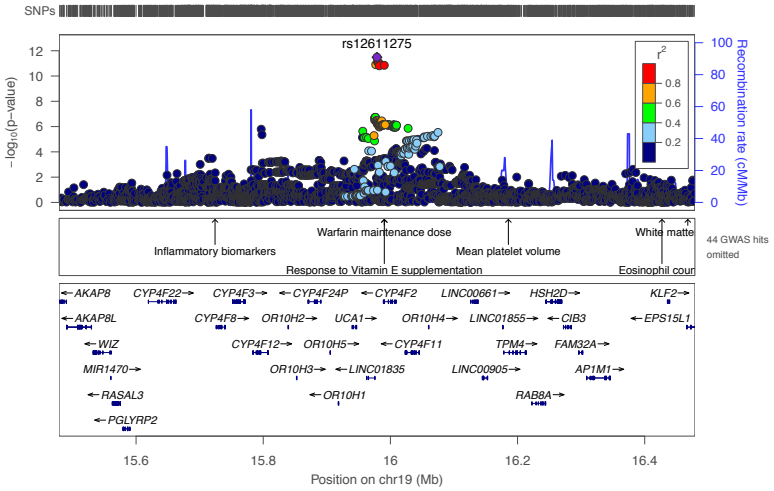
pregnanediol−3−glucuronide



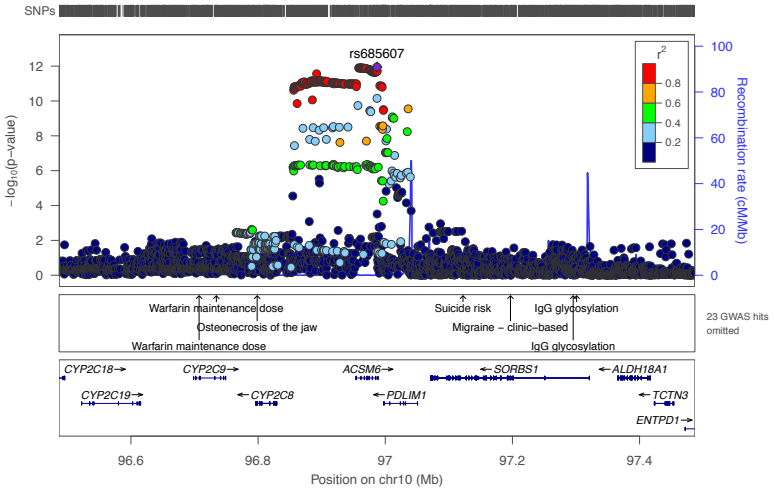
indoleacetylglutamine



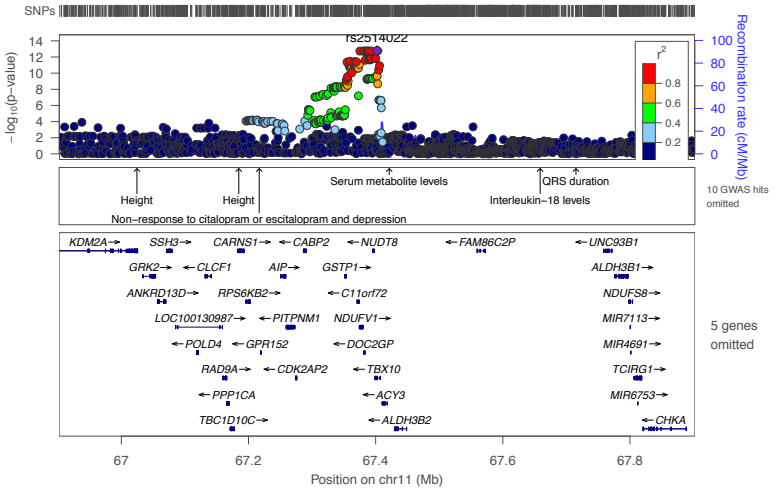
gamma−CEHC glucuronide*



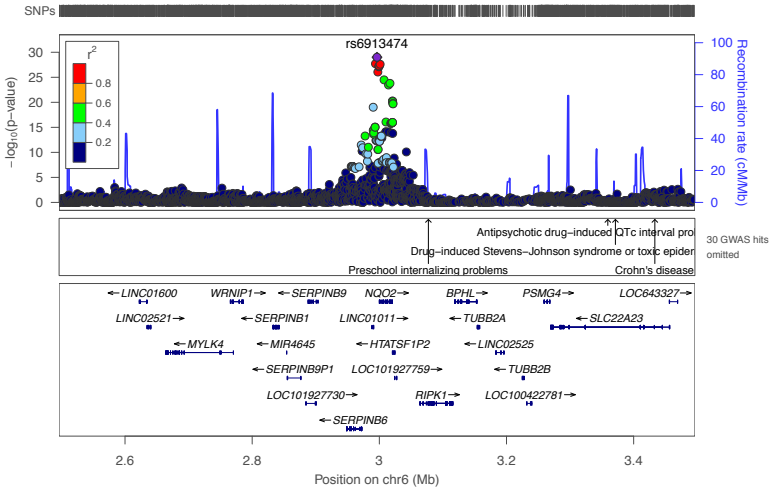
dimethylmalonic acid



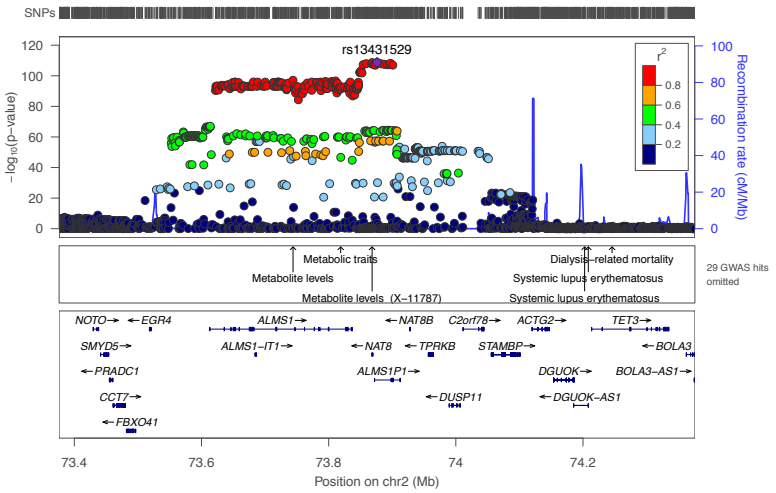
dimethylmalonic acid



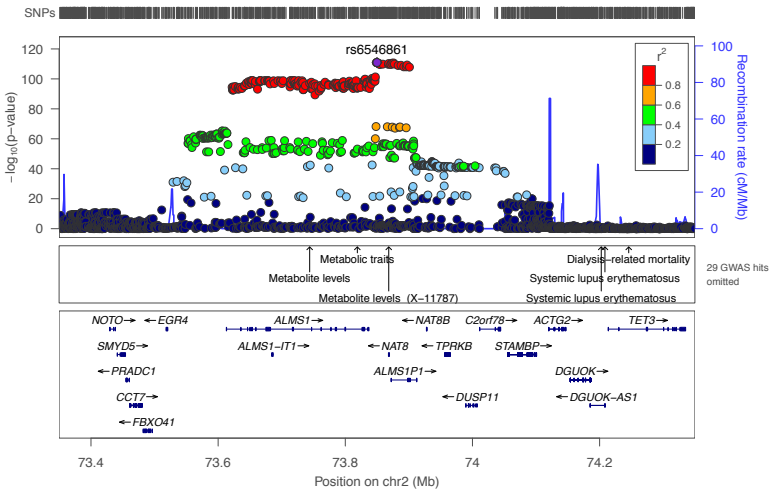
pterin



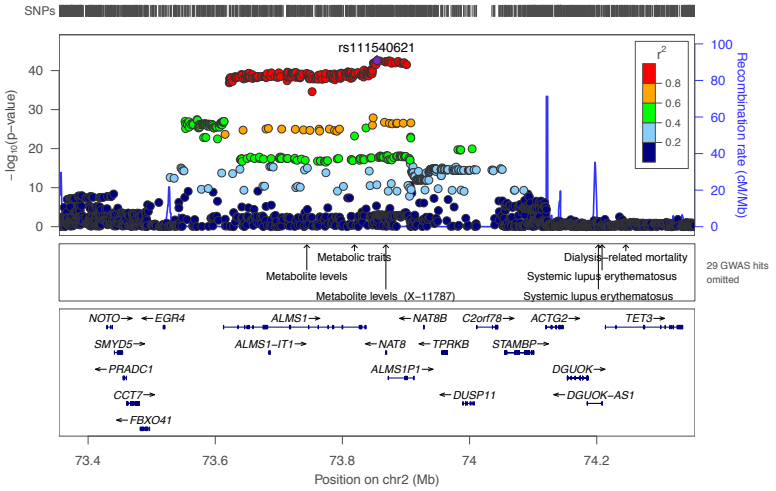
N−delta−acetylornithine



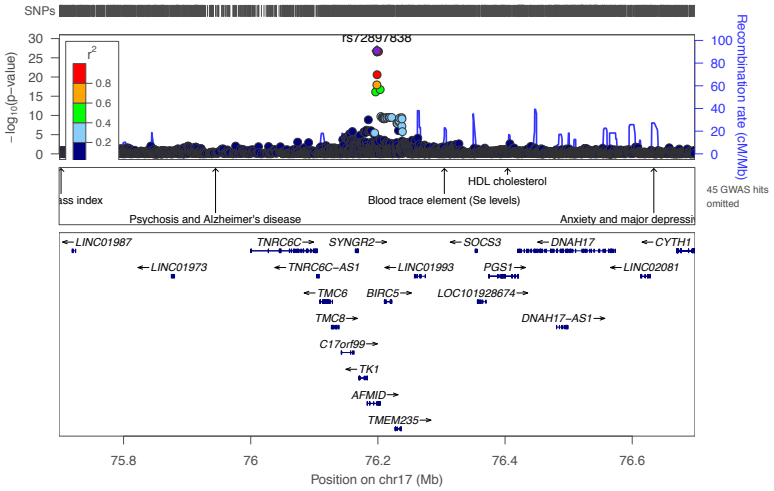
N−acetyl−1−methylhistidine*



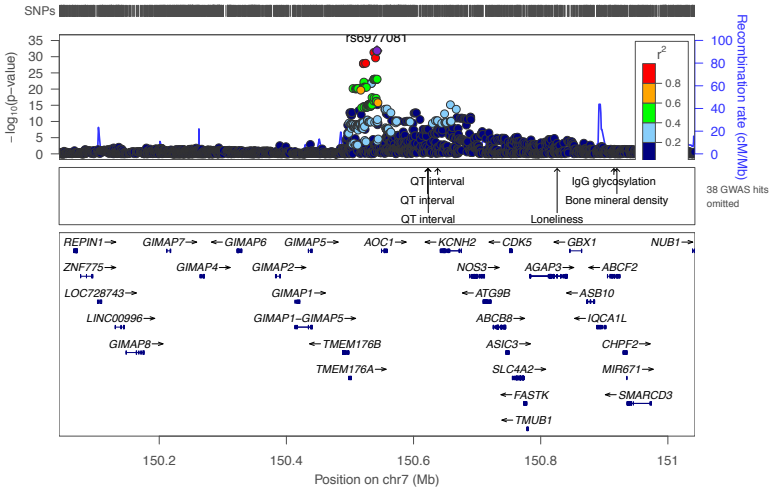
N−acetyl−3−methylhistidine*



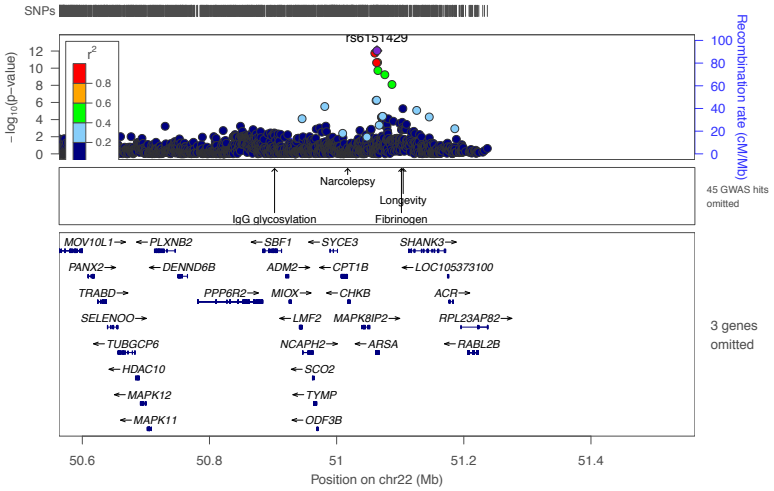
N−formylanthranilic acid



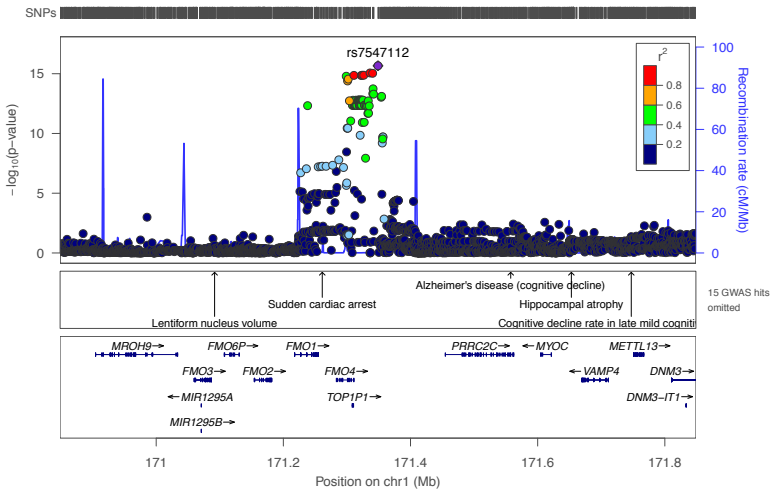
1−methylhistamine



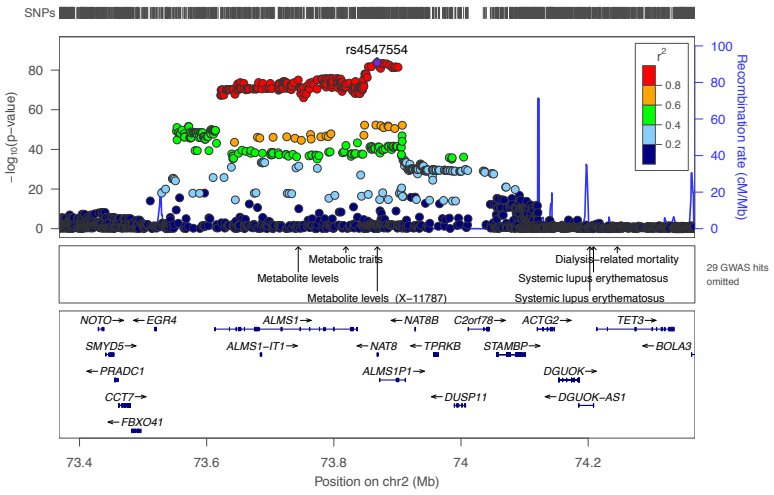
O−sulfo−L−tyrosine



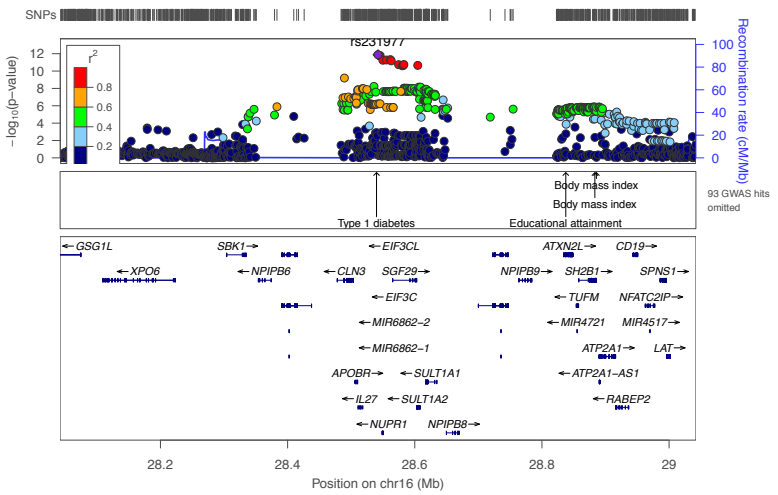
N−acetylmethionine sulfoxide



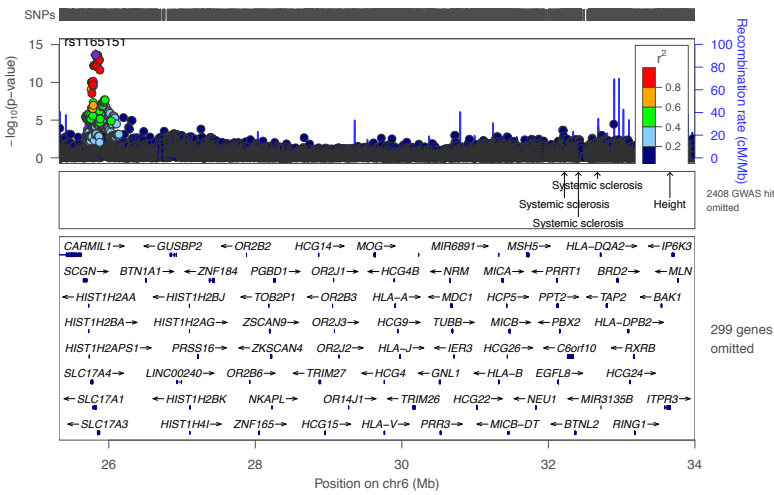
N−acetylmethionine sulfoxide



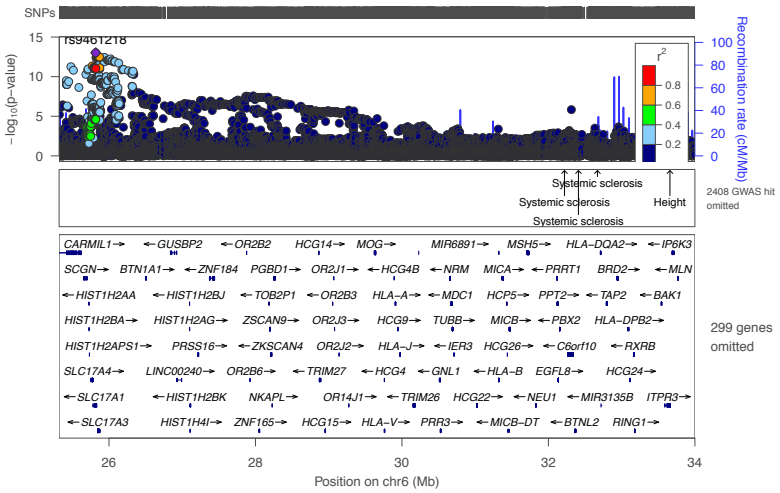
X − 12714



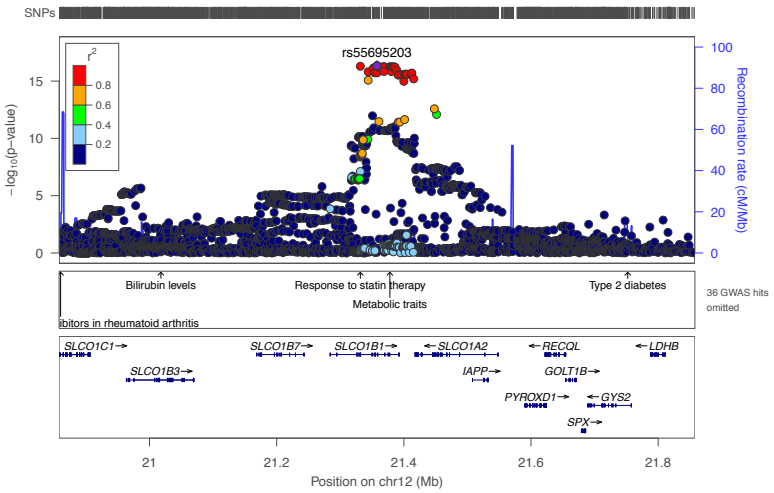
X − 13866



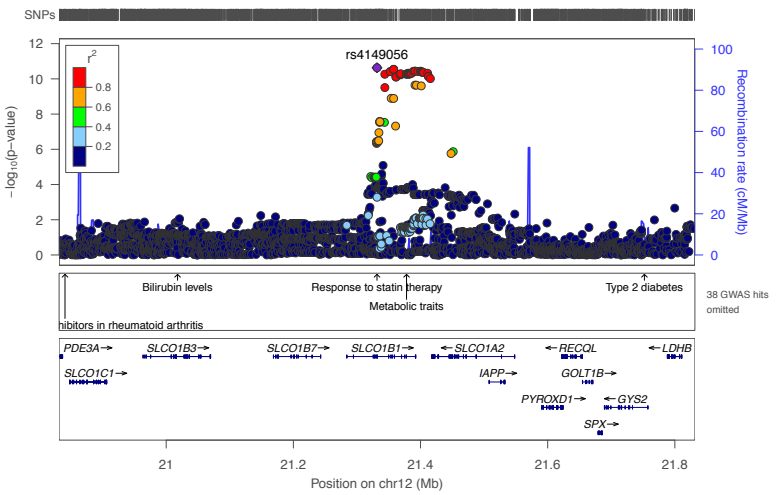
X – 15497 – retired for 1–carboxyethylphenylalanine



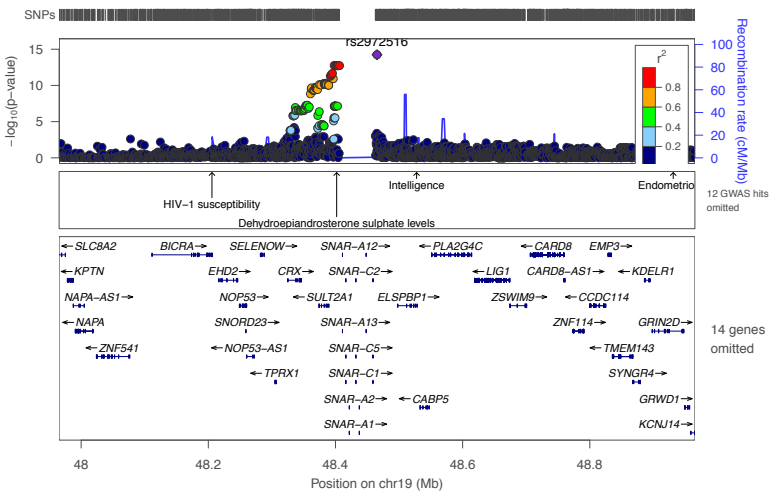
X – 21467



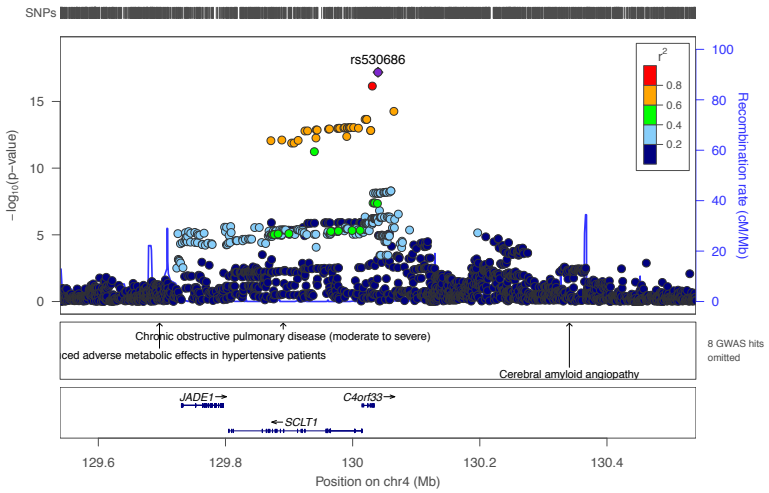
X – 21470



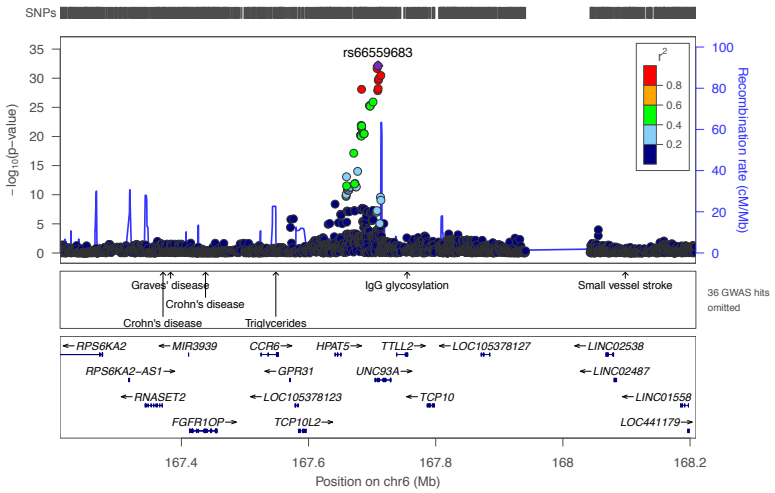
X – 21470



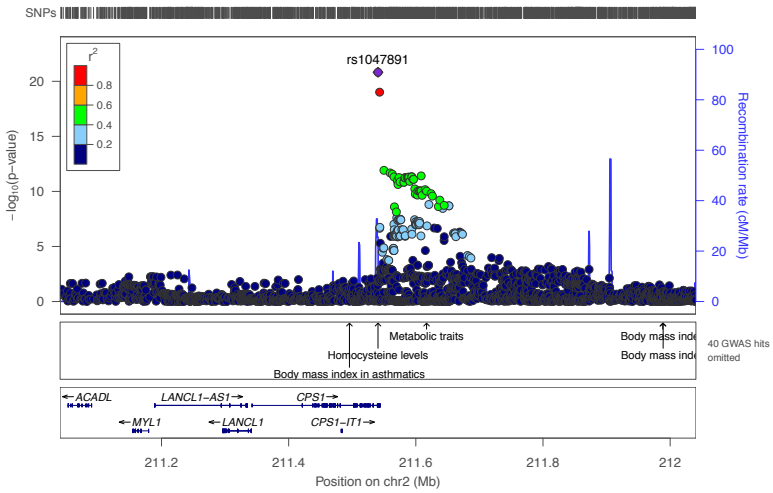
N−acetylglucosamine/N−acetylgalactosamine



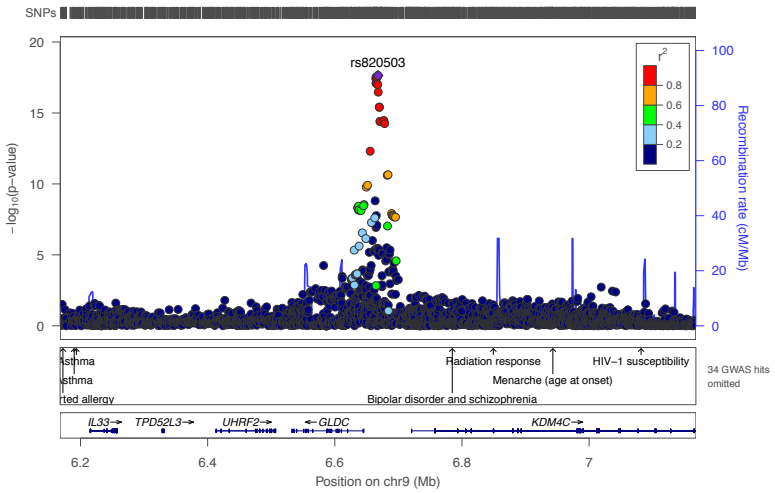
N−acetylglucosamine/N−acetylgalactosamine

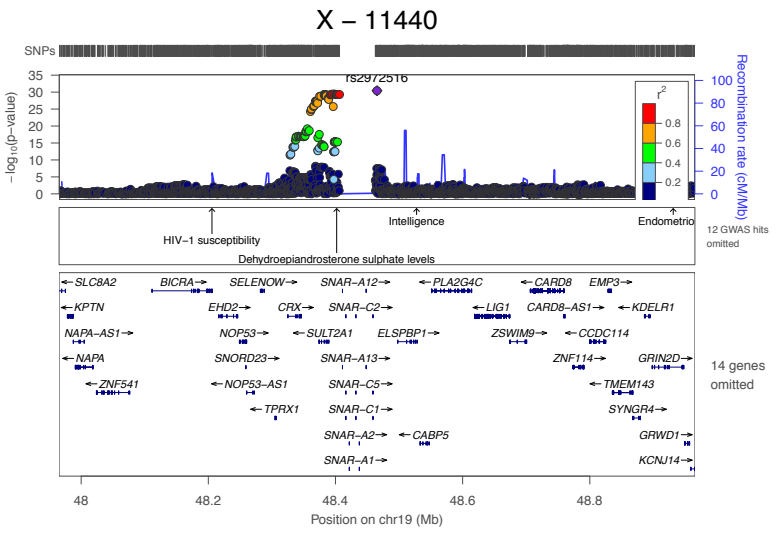


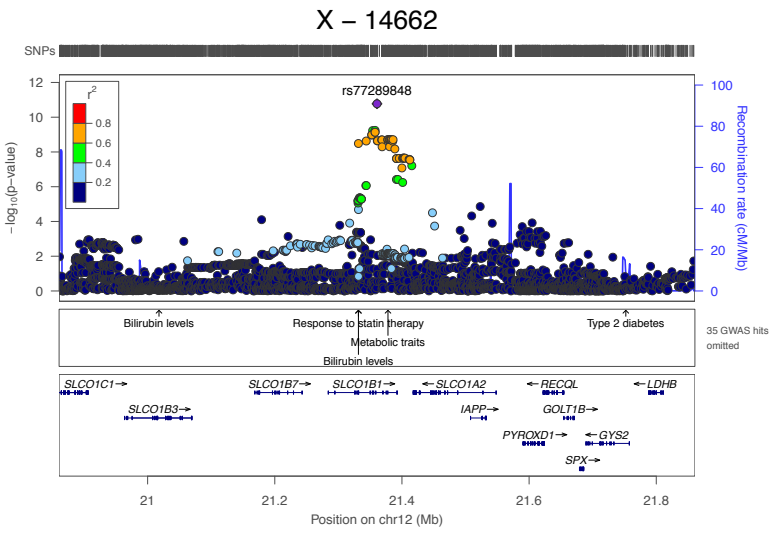
3−methylglutarylcarnitine (2)

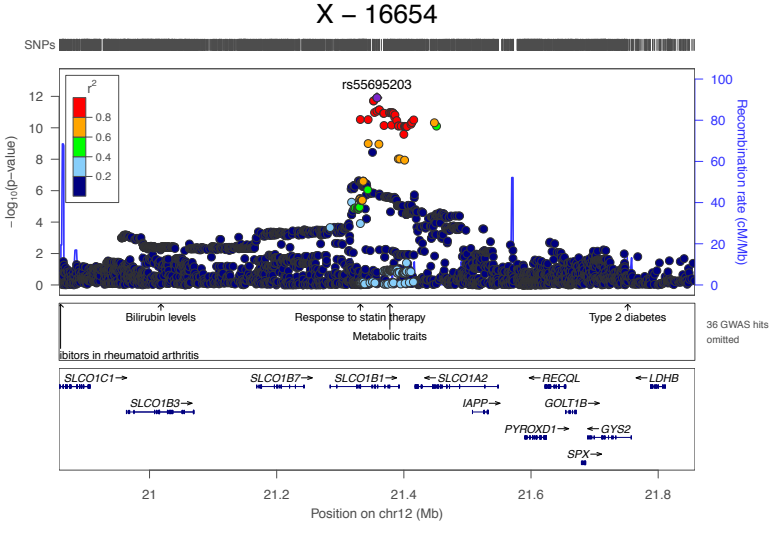


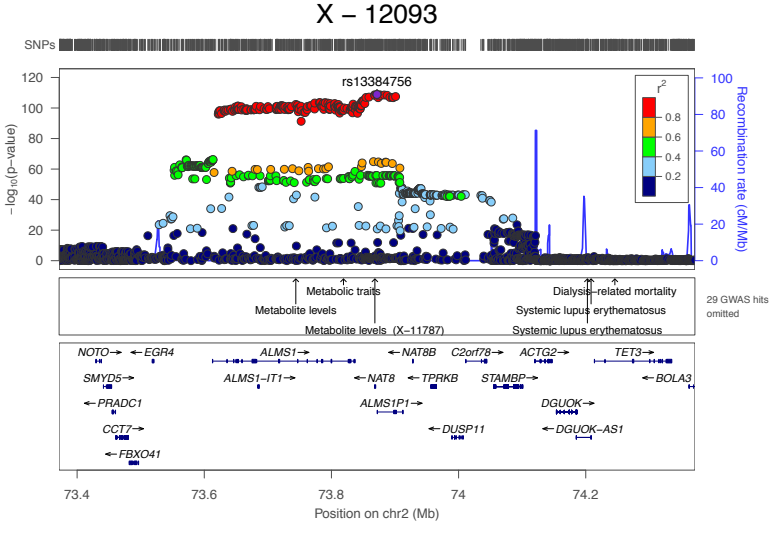
3−methylglutarylcarnitine (2)



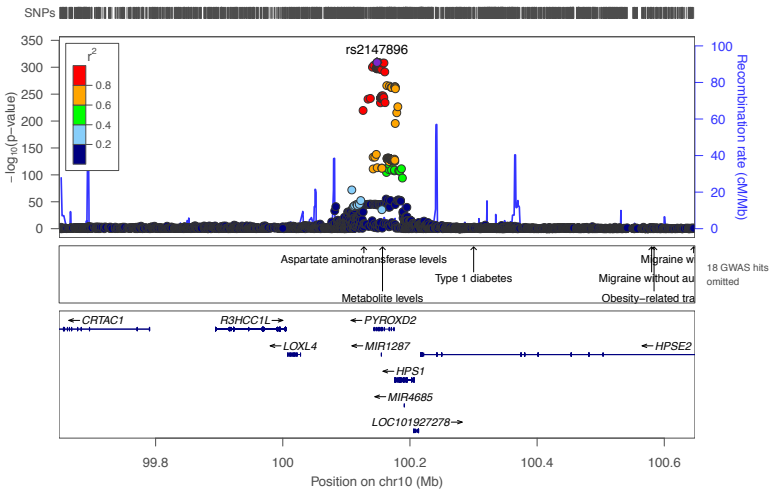

X – 11440
SNPs
rs2972516
HIV−1 susceptibility
Intelligence
Endometrio
Dehydroepiandrosterone sulphate levels
12 GWAS hits omitted
14 genes omitted
Position on chr19 (Mb)
Recombination rate (cM/Mb)


X – 14662
SNPs
rs77289848
Bilirubin levels
Response to statin therapy
Metabolic traits
Type 2 diabetes
35 GWAS hits omitted
Position on chr12 (Mb)
Recombination rate (cM/Mb)

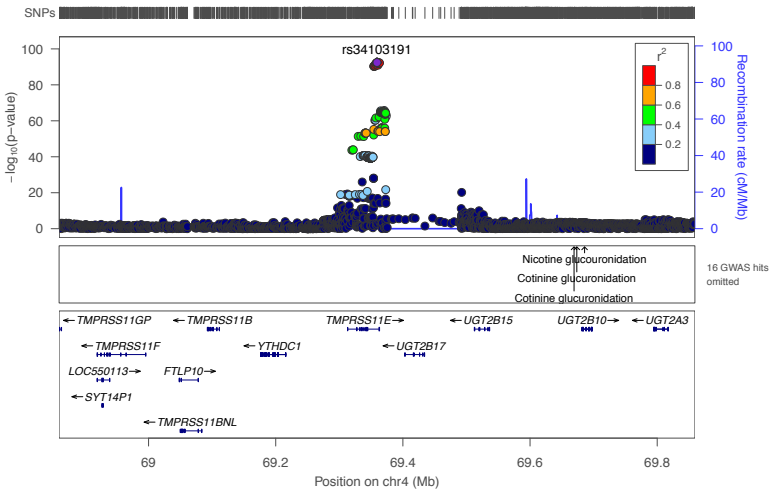

X – 16654
SNPs
rs55695203
Bilirubin levels
Response to statin therapy
Metabolic traits
Type 2 diabetes
36 GWAS hits omitted
Position on chr12 (Mb)
Recombination rate (cM/Mb)


X – 12093
SNPs
rs13384756
Metabolic traits
Metabolite levels
Metabolite levels (X−11787)
Dialysis−related mortality
Systemic lupus erythematosus
Systemic lupus erythematosus
29 GWAS hits omitted
Position on chr2 (Mb)
Recombination rate (cM/Mb)

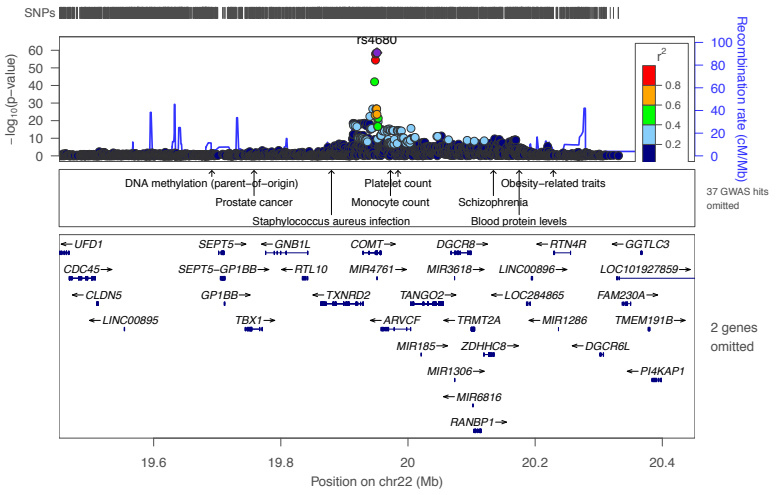
X – 12093



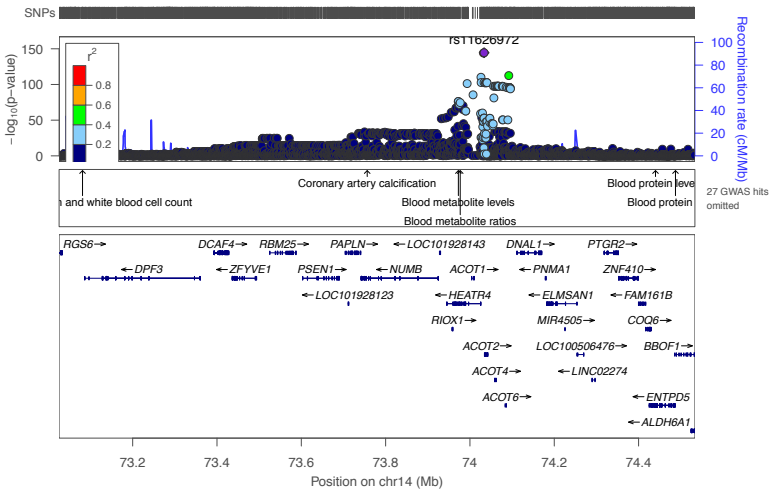
X – 19141



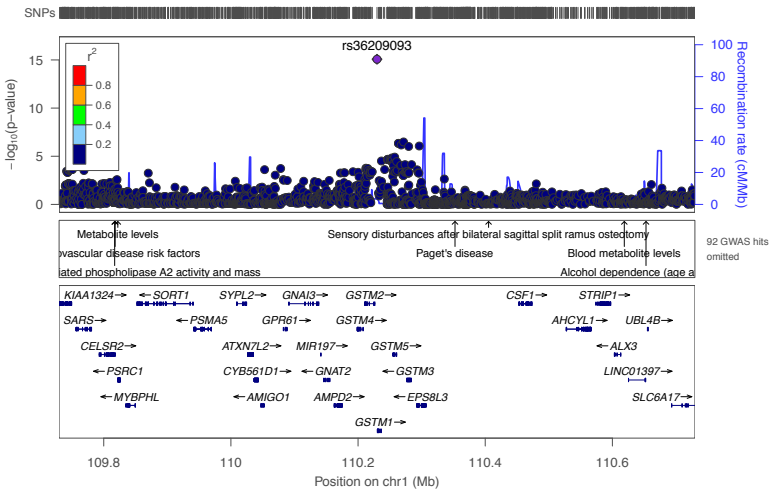
X – 21785



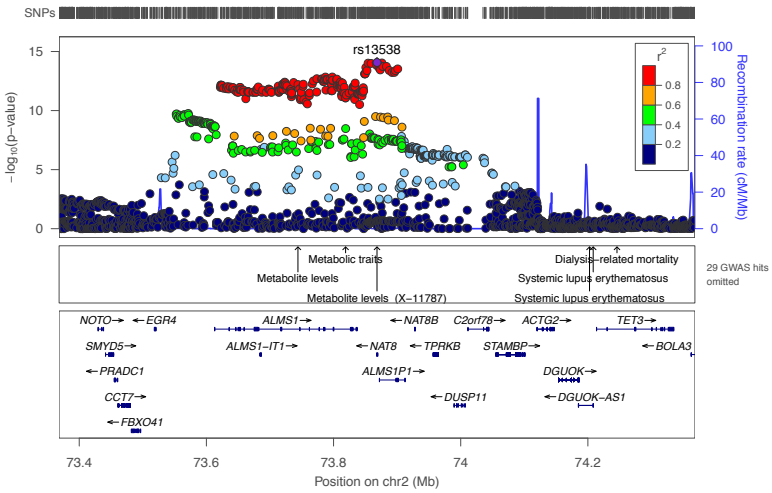
X – 21792



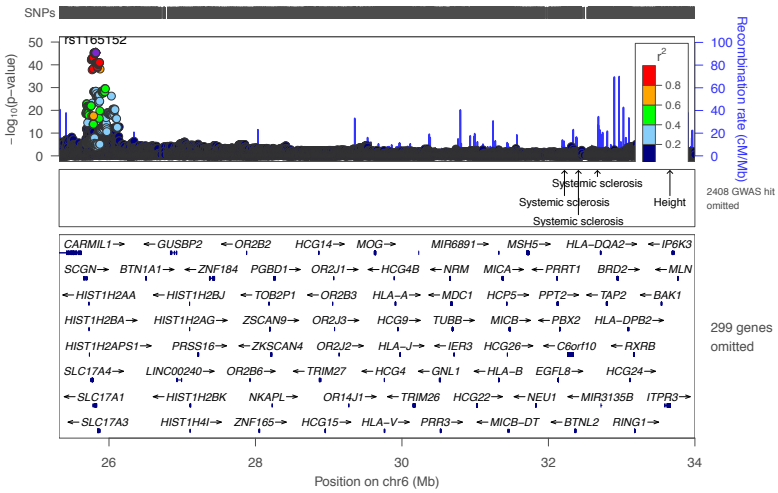
X – 21796



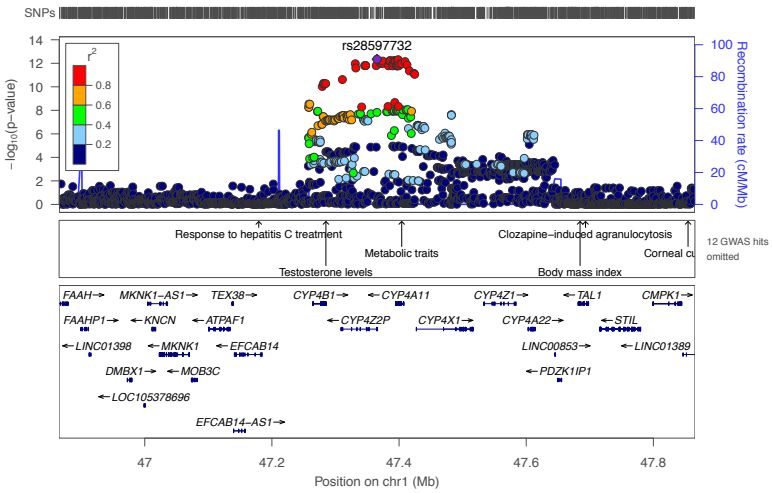
X – 12379



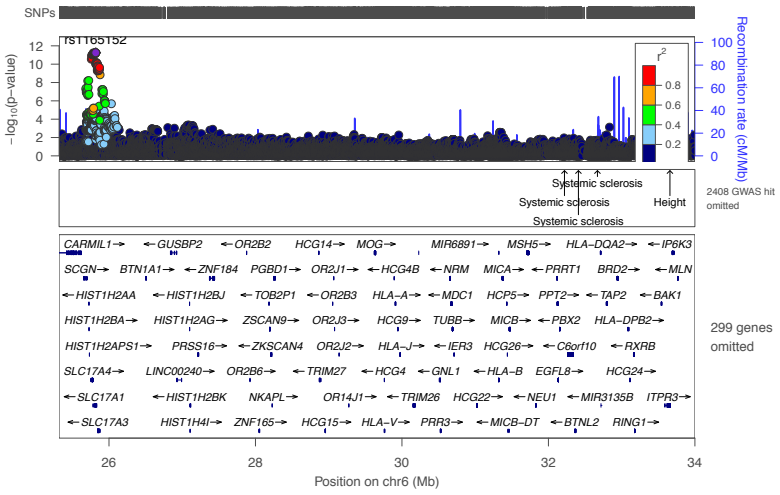
X – 12822



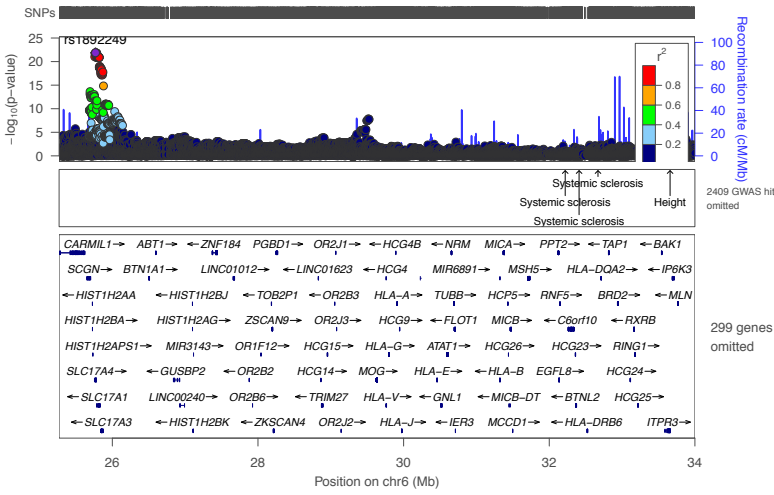
X – 21827



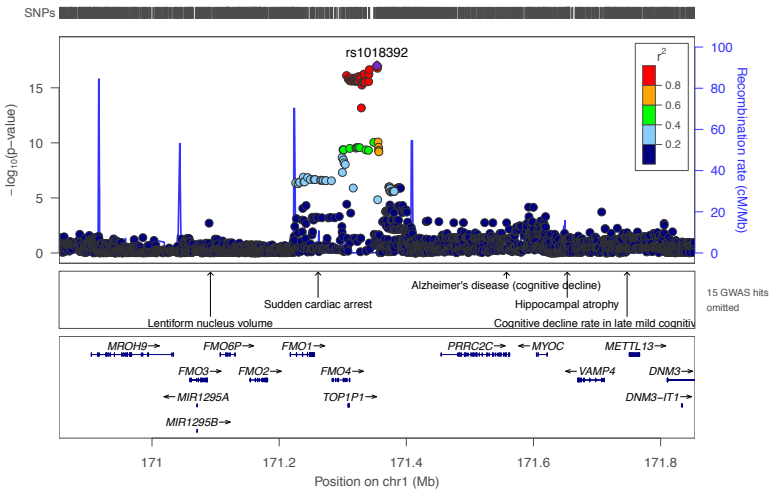
X – 21829



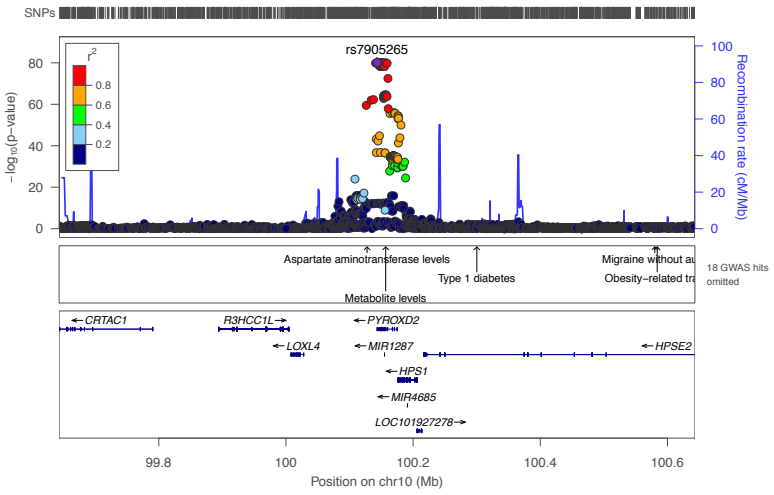
X – 12839



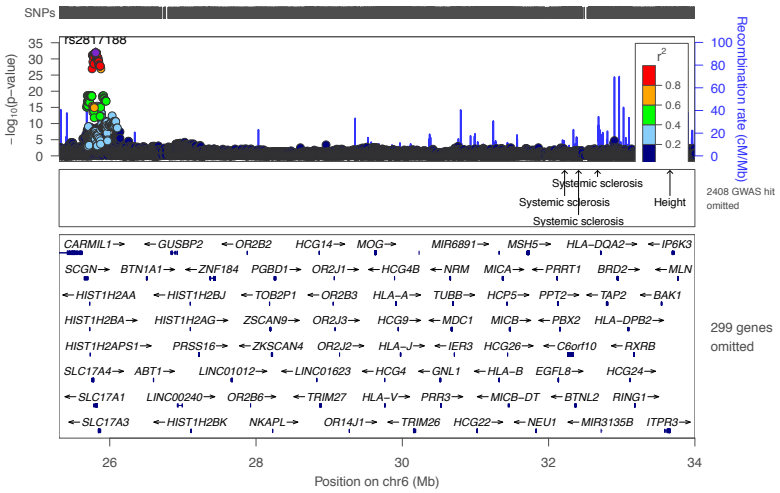
X – 21842



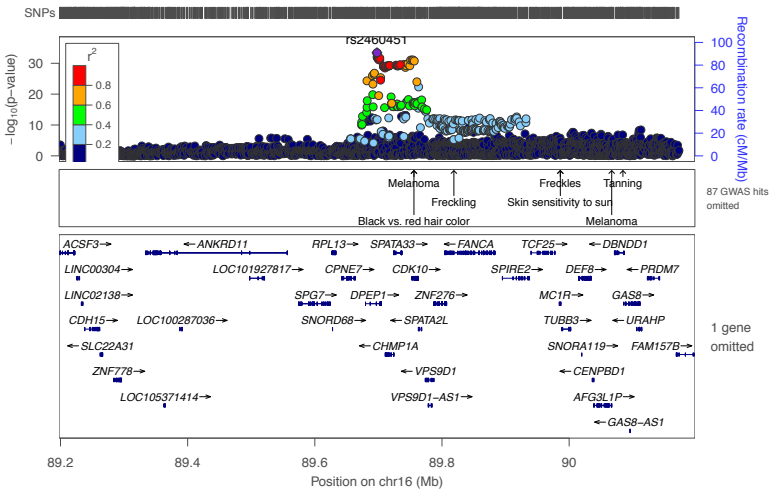
N–methylpipecolate



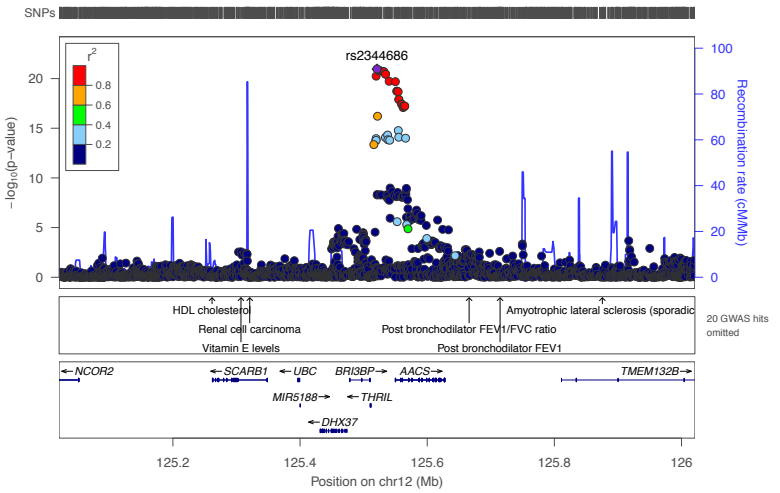
ferulic acid 4−sulfate



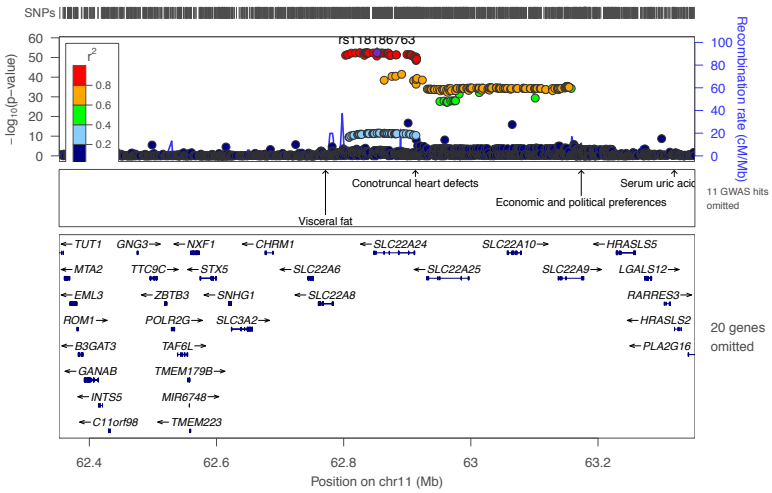
X − 18889 − retired for 1−carboxyethylleucine



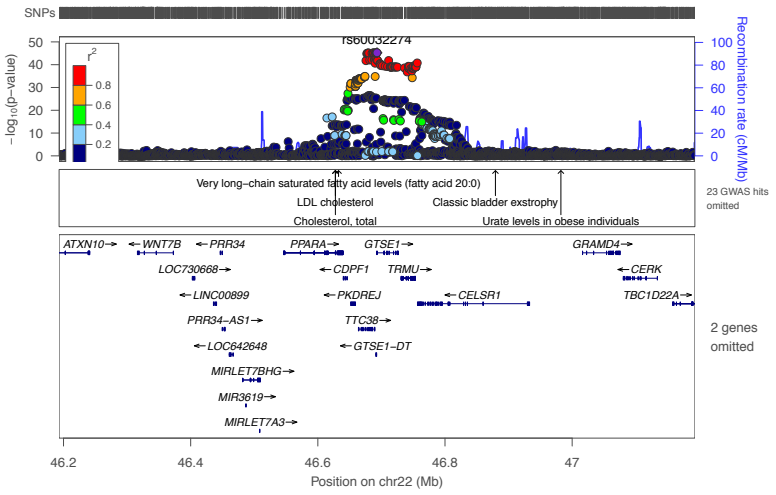
X − 19913



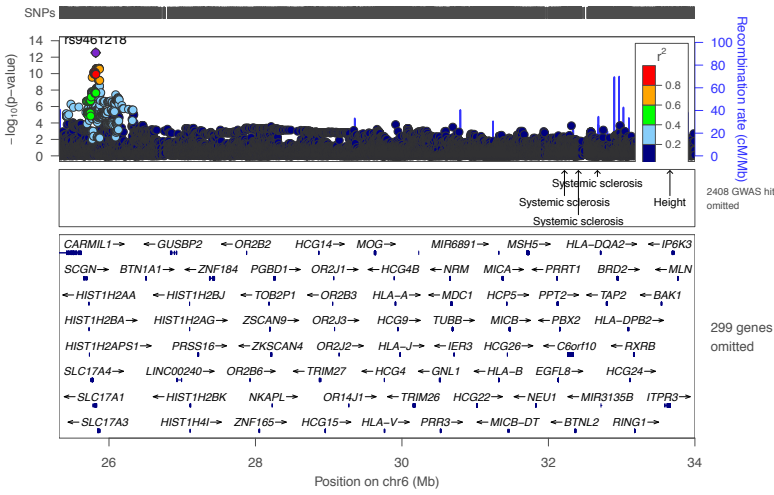
17alpha−hydroxypregnanolone glucuronide



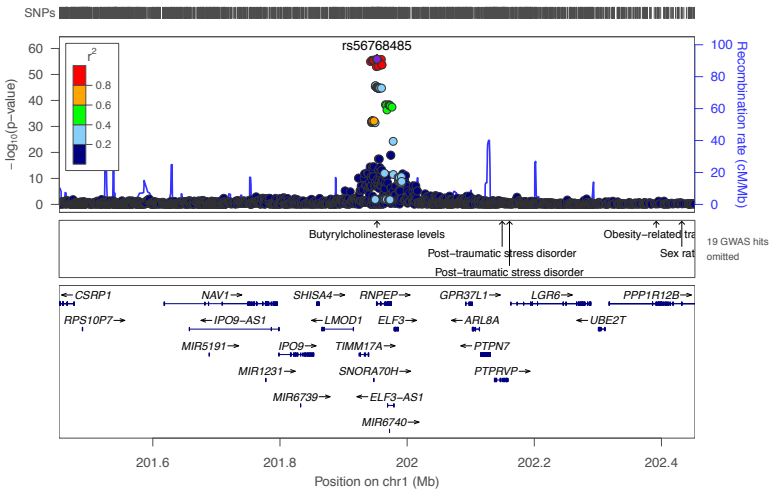
X – 22162



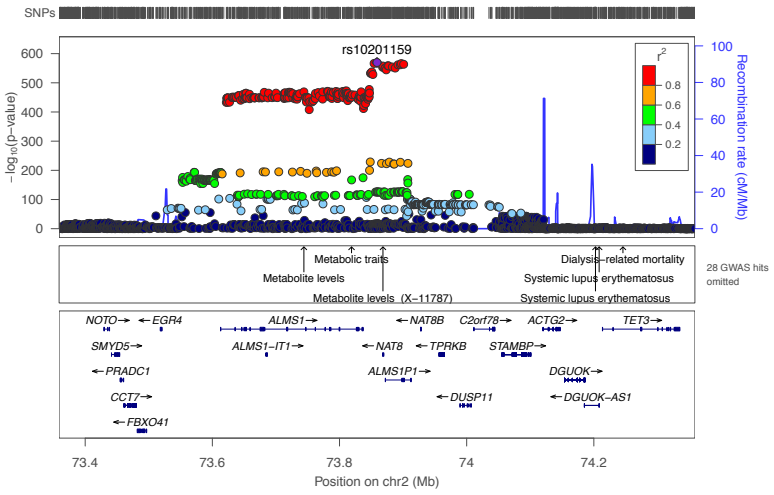
X – 19561 – retired for 1−carboxyethyltyrosine



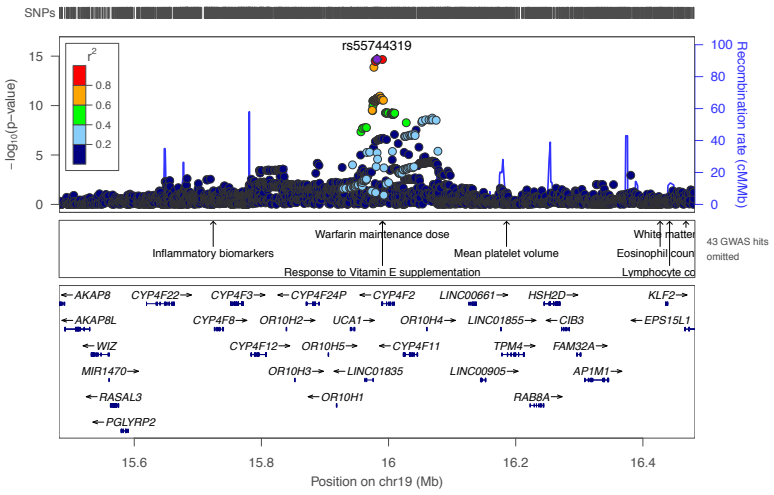
X – 10457



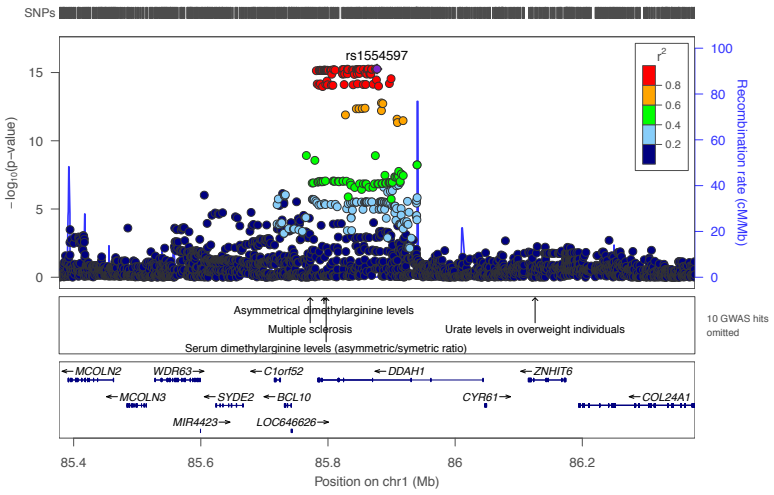
X – 12125



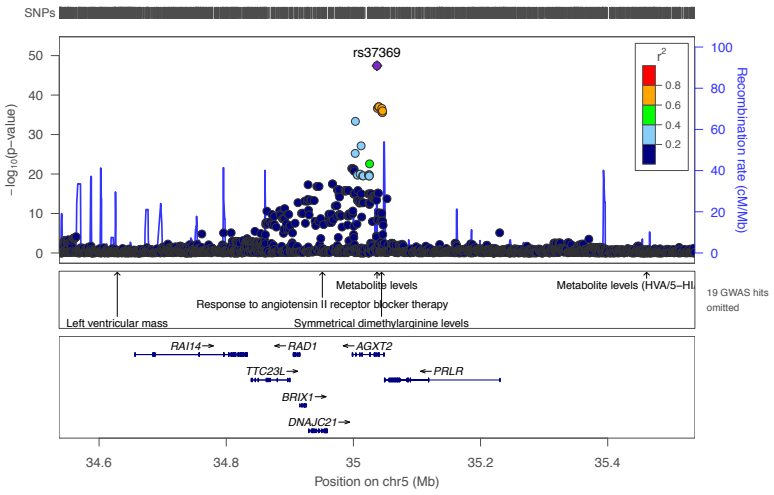
alpha−CEHC sulfate



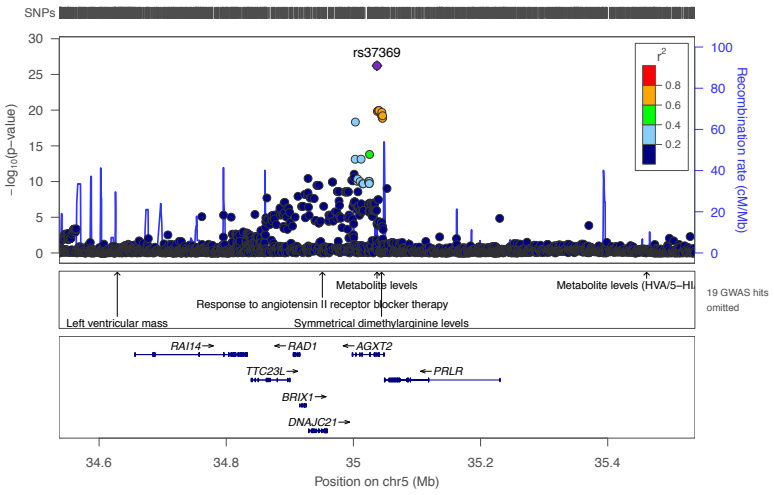
X – 12097



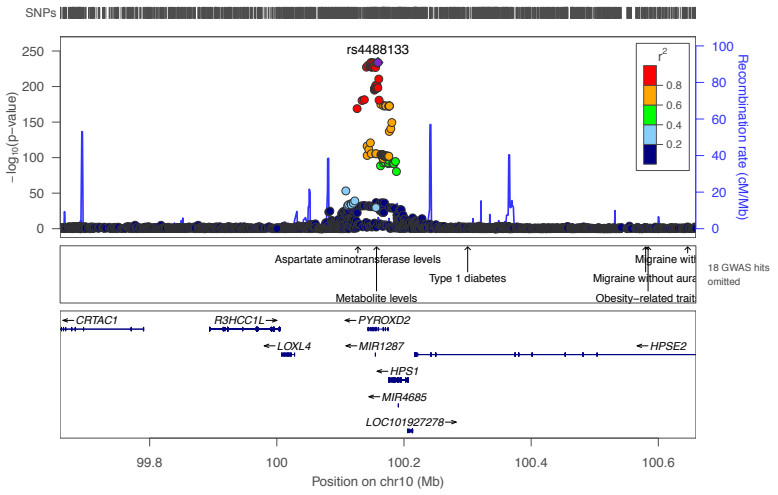
X – 12097



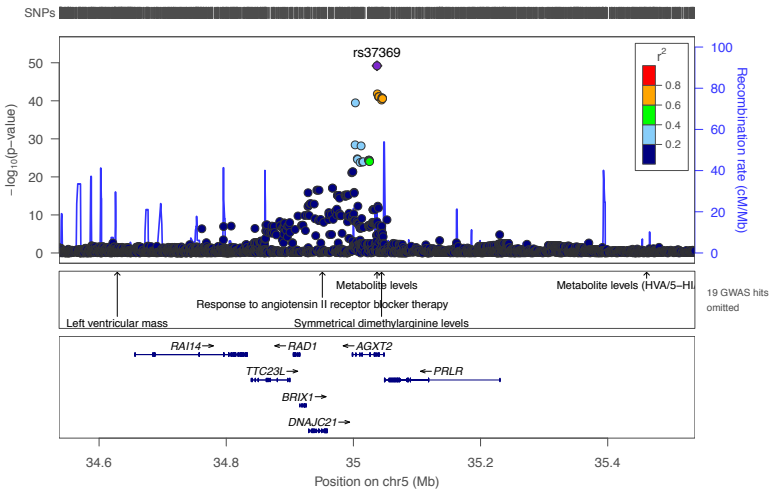
X – 12096



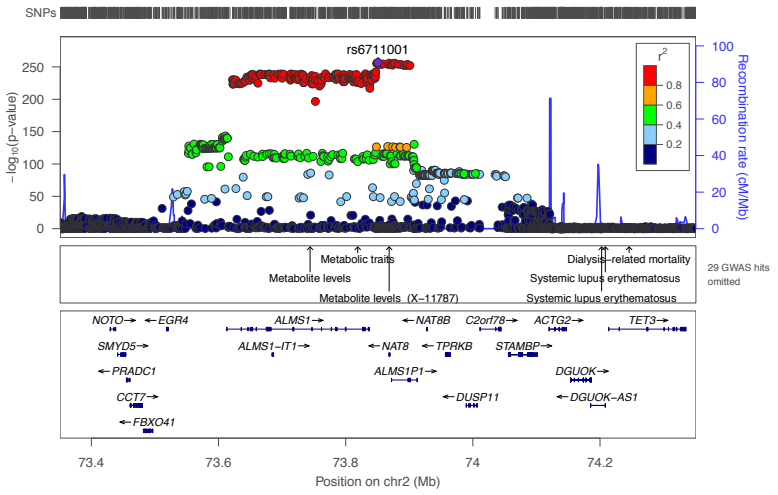
X – 12112



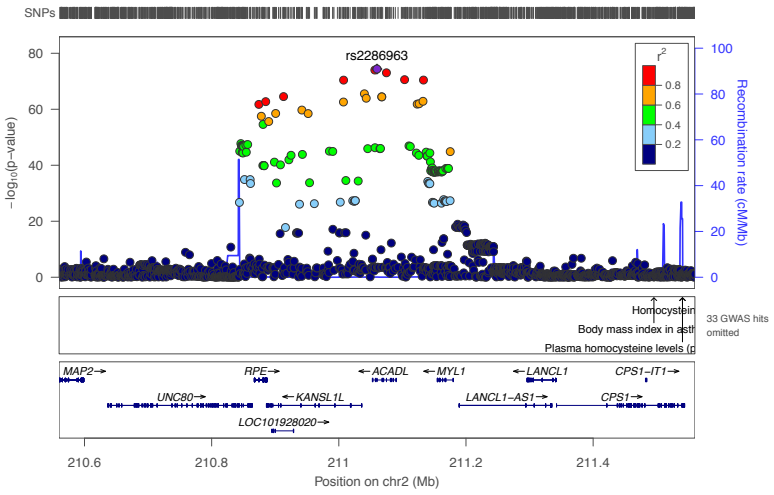
X – 12117



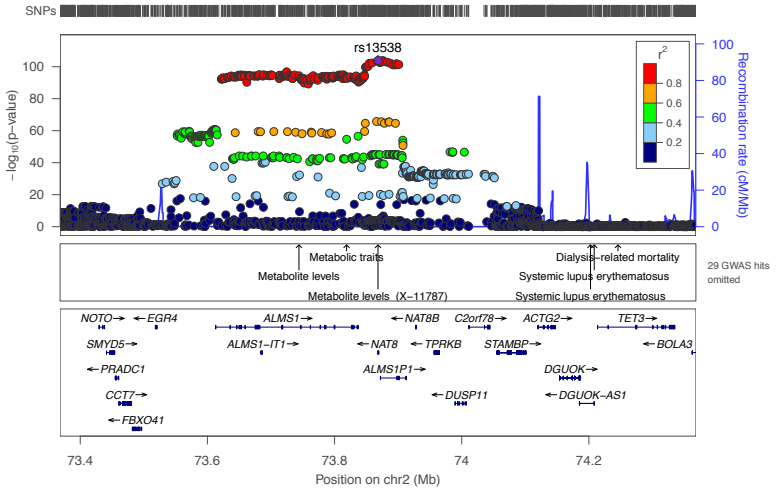
X – 12124



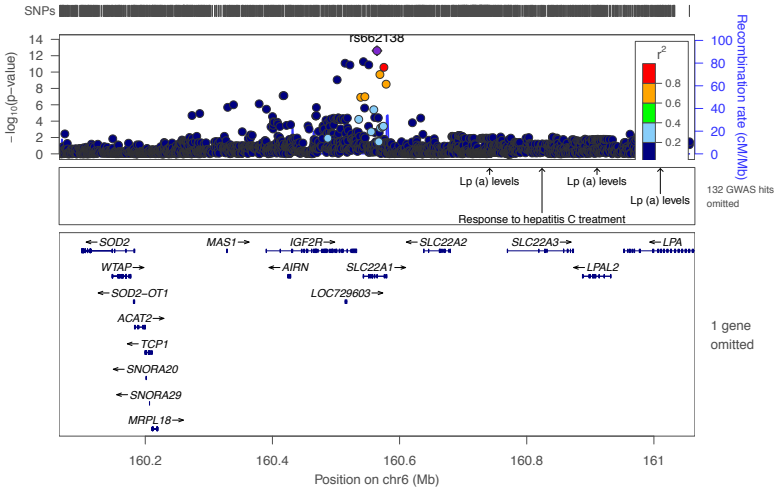
X – 13431



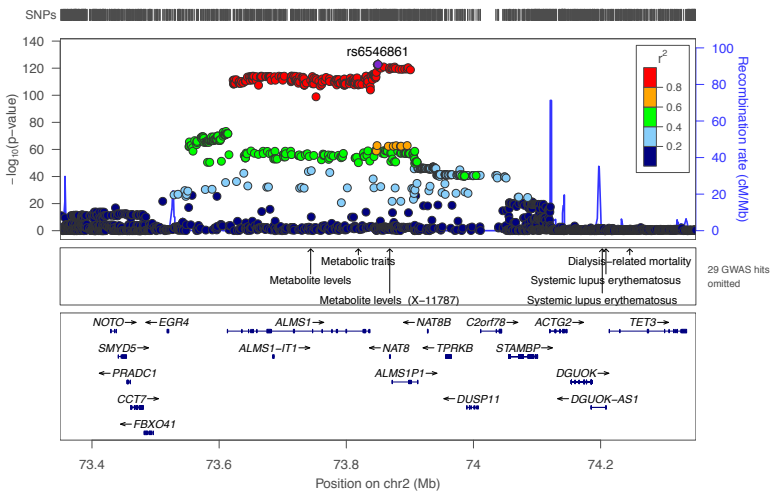
X – 15666



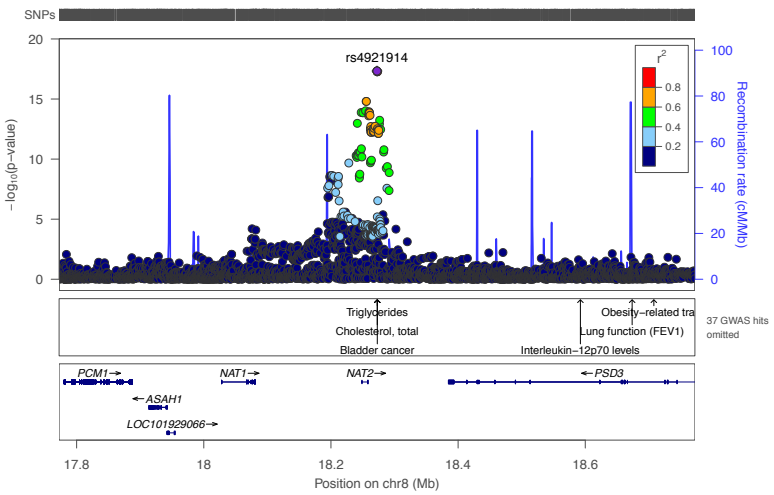
X – 17323



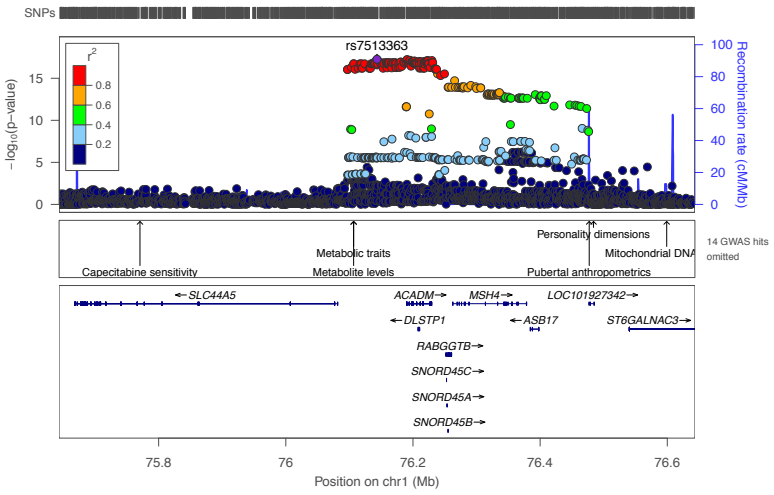
X – 11357



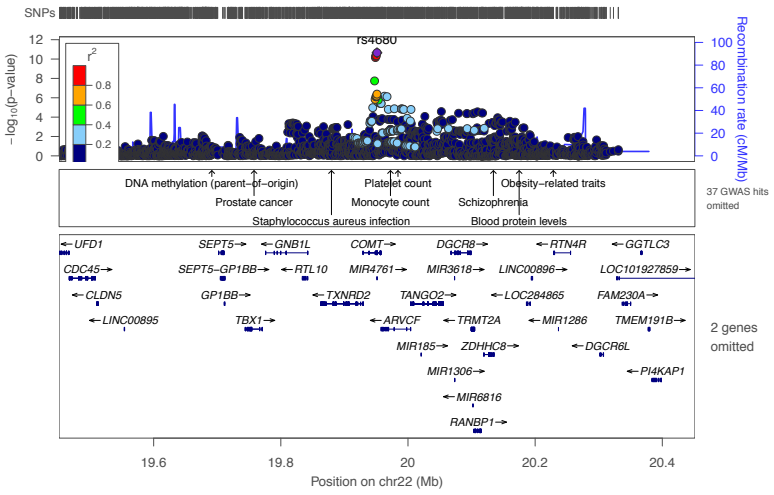
X – 12410



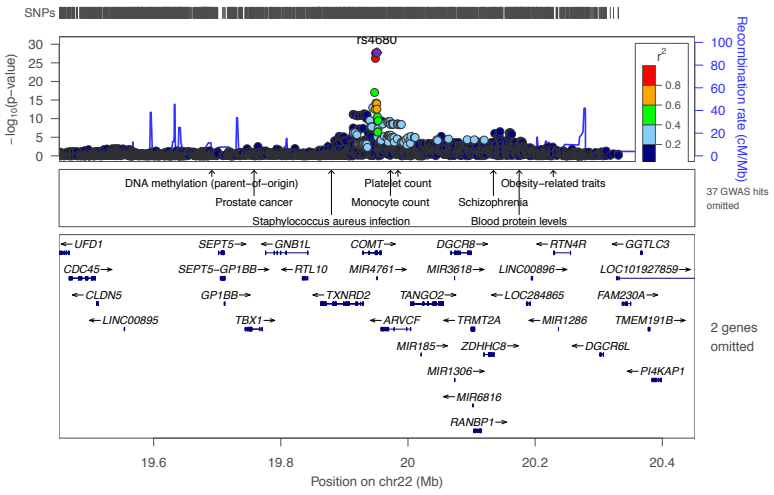
X – 12636



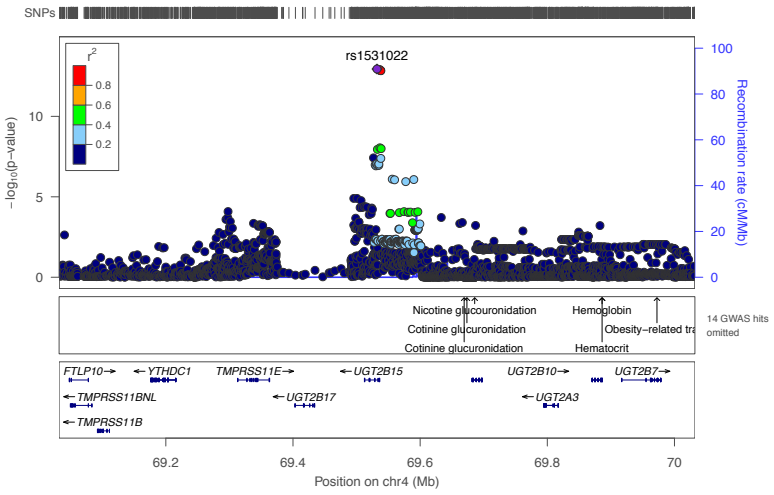
X – 12707



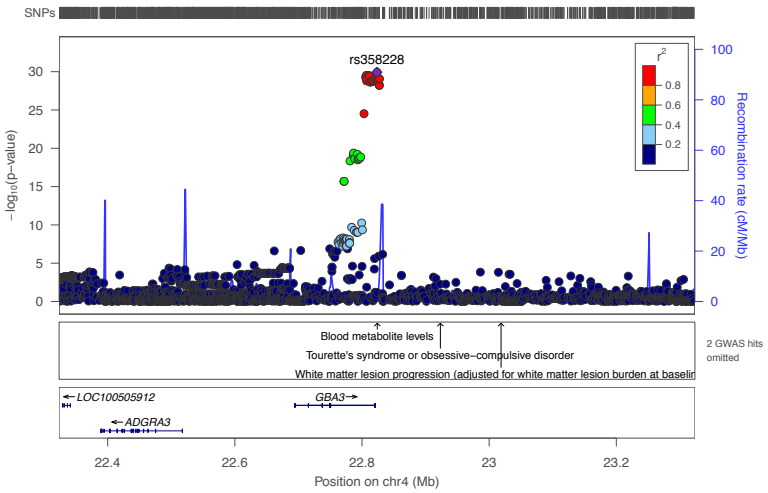
X – 12708



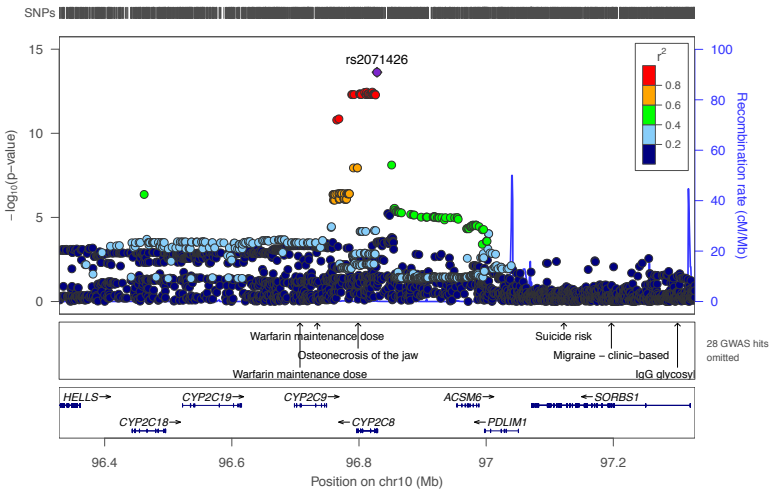
X – 13688



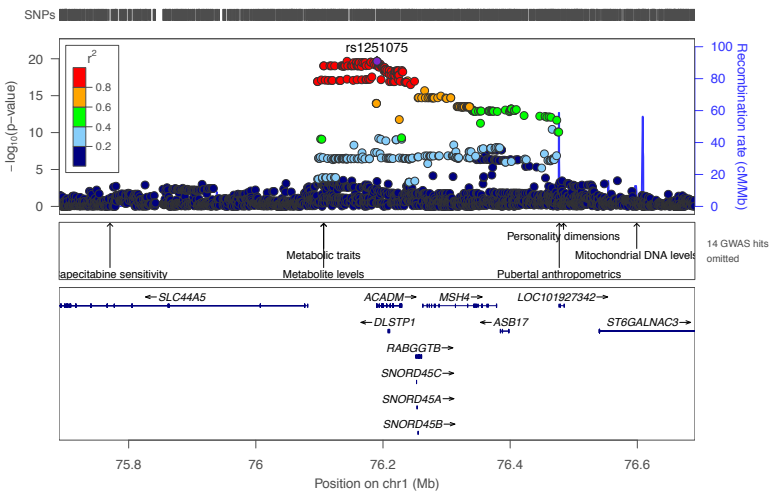
X – 13847



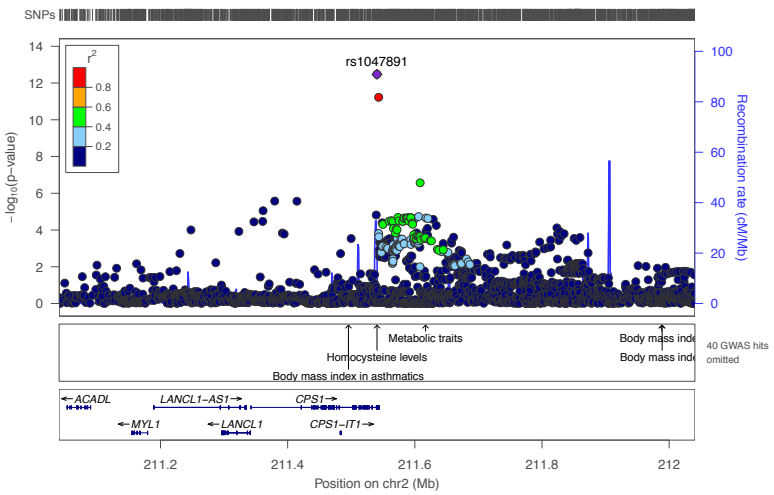
X – 16087

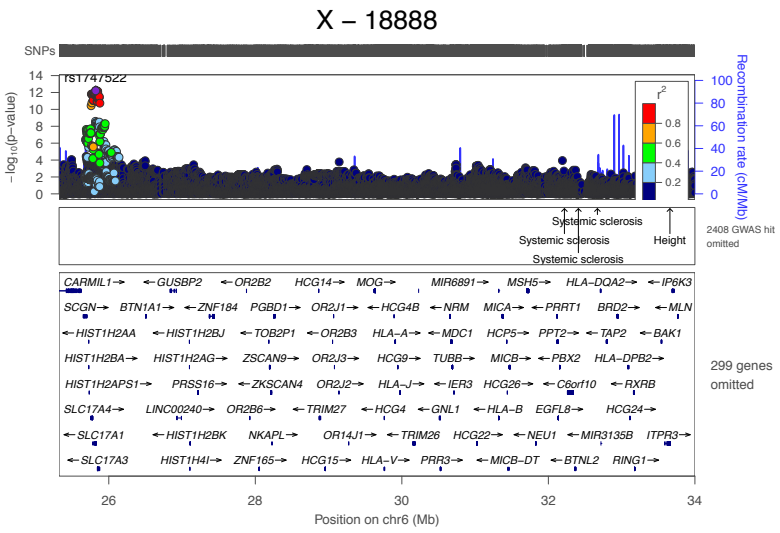


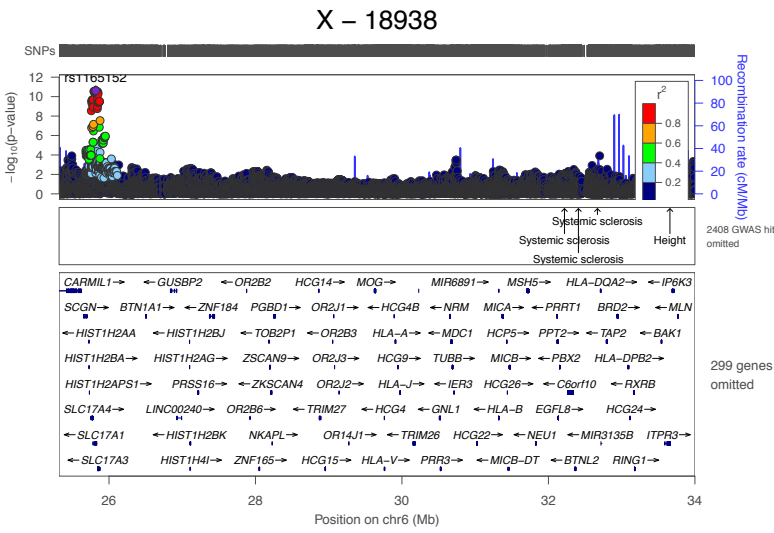
X – 16567

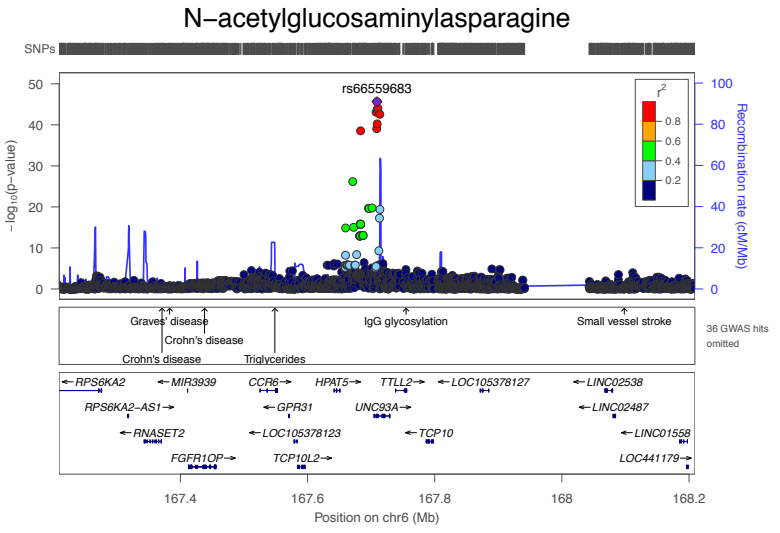


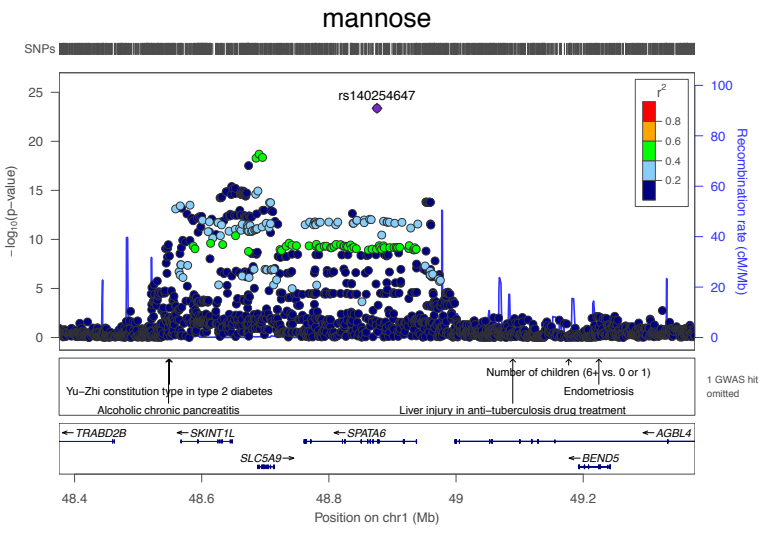
X – 16567



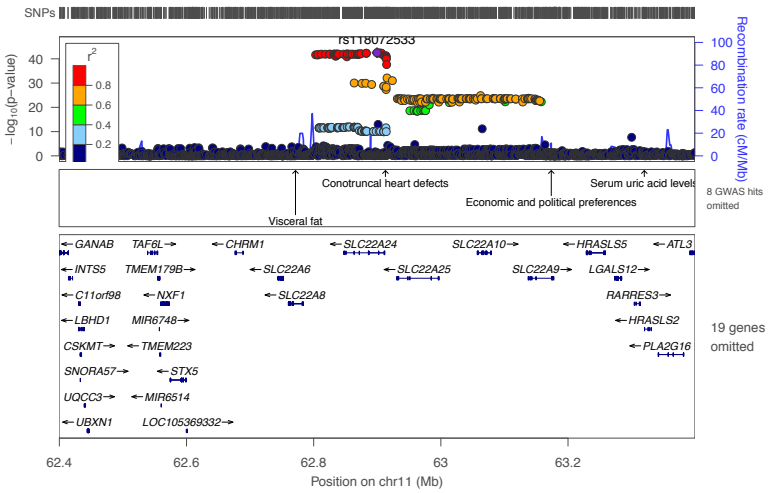

X – 18888
SNPs
rs1747522
−log10(p−value)
Recombination rate (cM/Mb)
r2
Systemic sclerosis
Systemic sclerosis
Systemic sclerosis
Height
2408 GWAS hit omitted
299 genes omitted
Position on chr6 (Mb)


X – 18938
SNPs
rs1165152
−log10(p−value)
Recombination rate (cM/Mb)
r2
Systemic sclerosis
Systemic sclerosis
Systemic sclerosis
Height
2408 GWAS hit omitted
299 genes omitted
Position on chr6 (Mb)

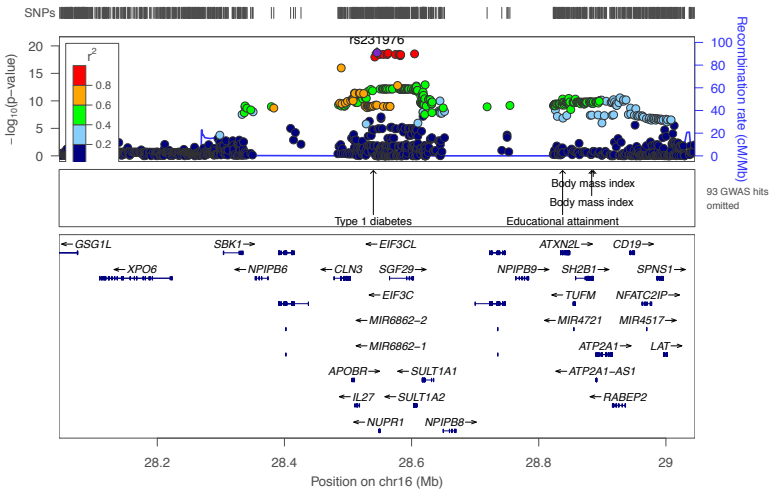

N−acetylglucosaminylasparagine
SNPs
rs66559683
−log10(p−value)
Recombination rate (cM/Mb)
r2
Graves' disease
Crohn's disease
Crohn's disease
Triglycerides
IgG glycosylation
Small vessel stroke
36 GWAS hits omitted
Position on chr6 (Mb)


mannose
SNPs
rs140254647
−log10(p−value)
Recombination rate (cM/Mb)
r2
Yu−Zhi constitution type in type 2 diabetes
Alcoholic chronic pancreatitis
Number of children (6+ vs. 0 or 1)
Endometriosis
Liver injury in anti−tuberculosis drug treatment
1 GWAS hit omitted
Position on chr1 (Mb)

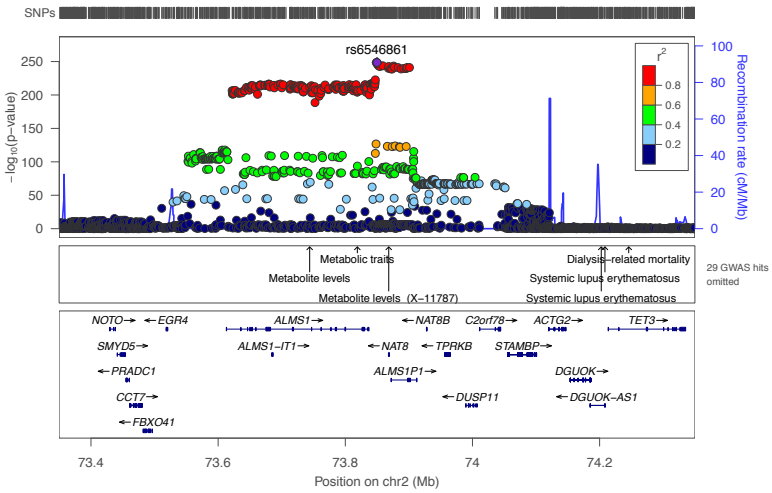
X – 22834



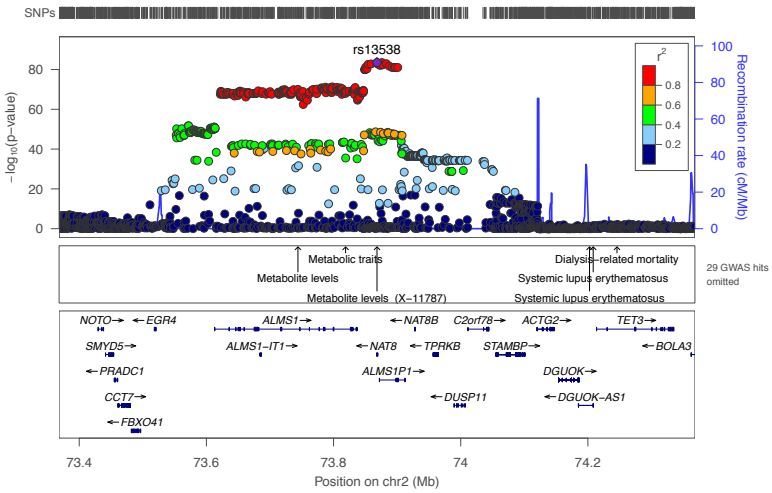
3–hydroxyindolin–2–one sulfate



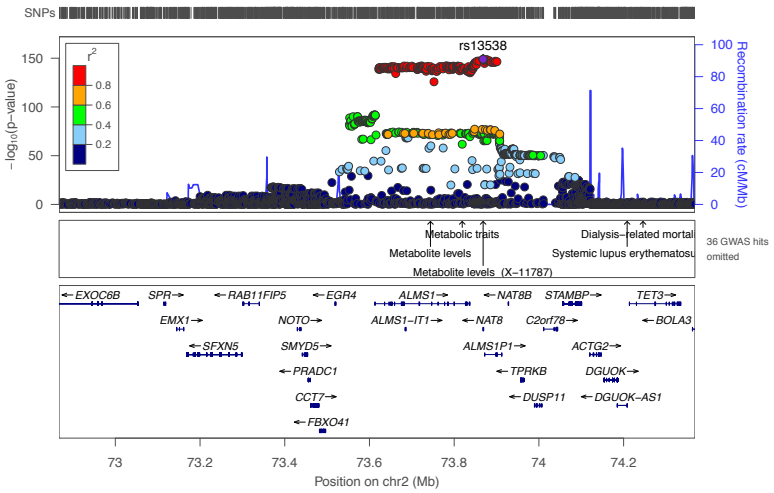
N–acetylcitrulline



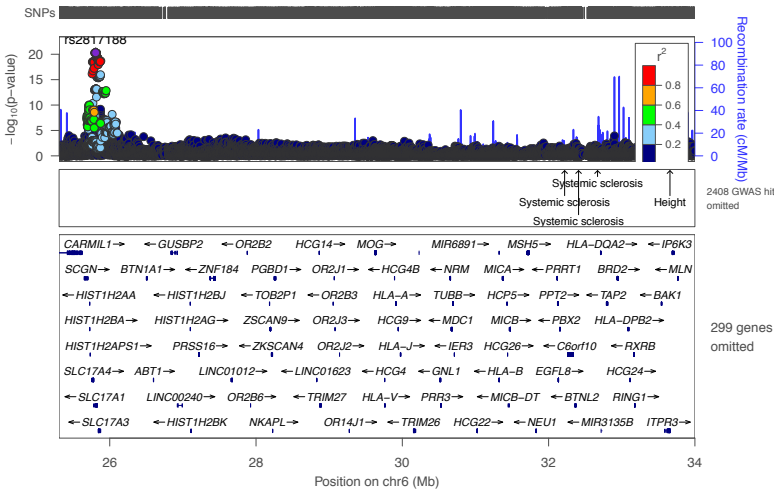
N–acetylpyrraline



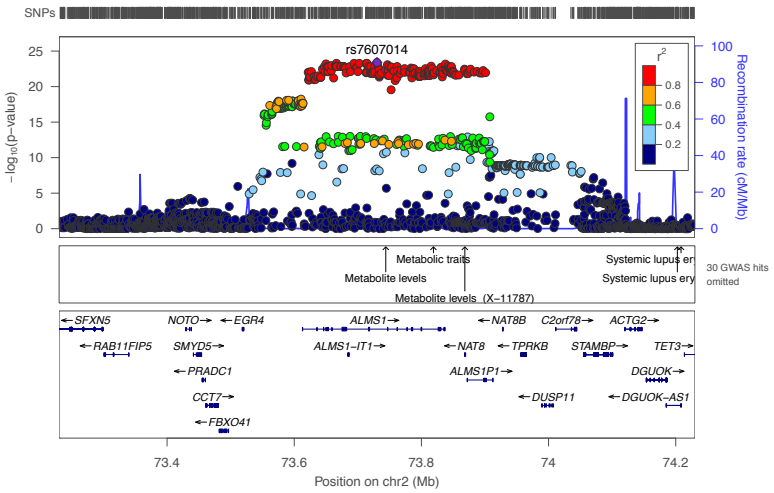
N2,N6−diacetyllysine



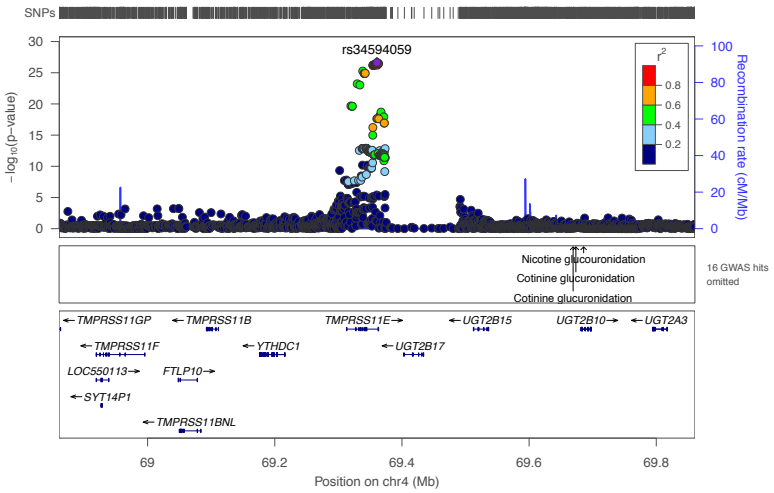
X − 23518



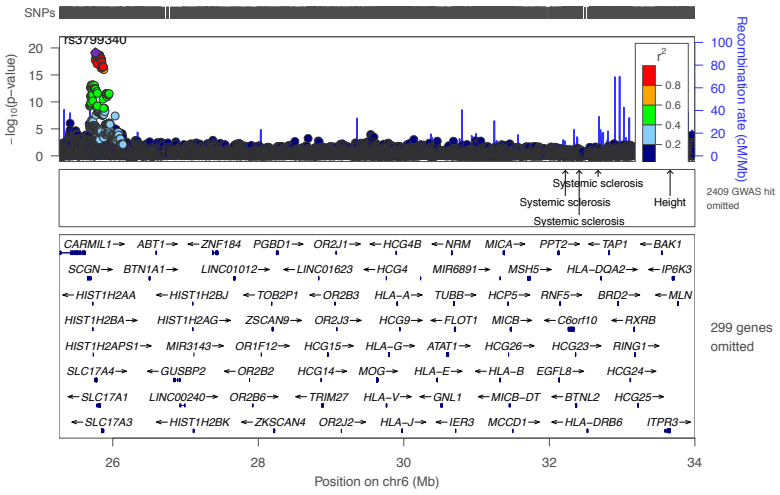
N−acetylkynurenine (2)



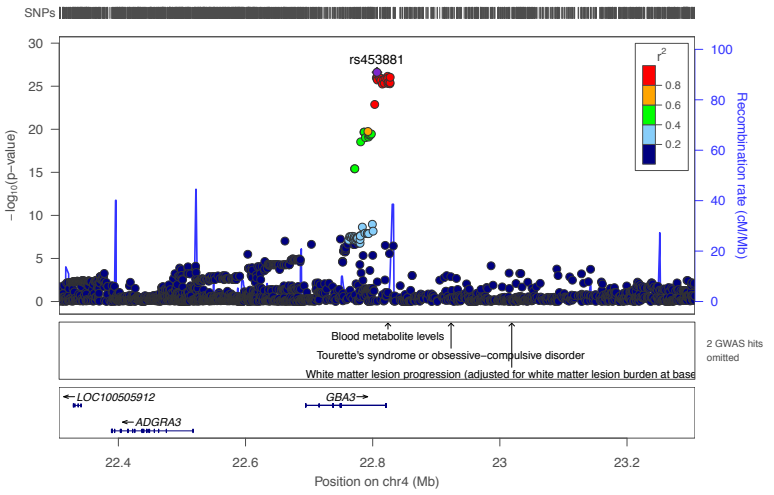
X − 19434 retired for cholic acid glucuronide



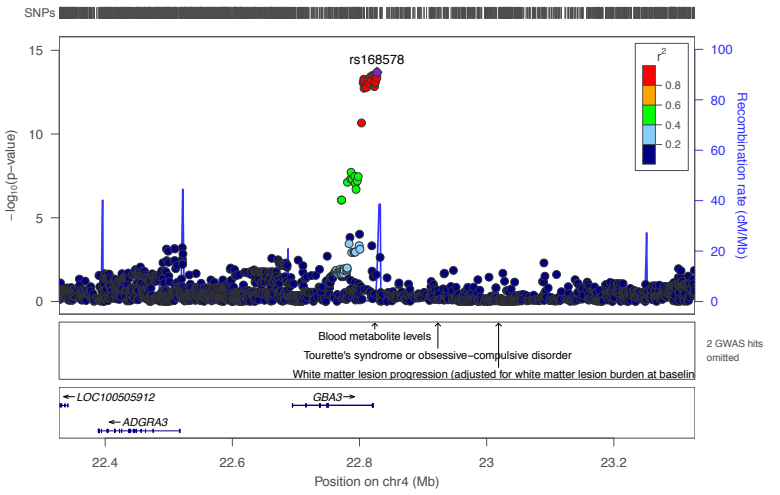
X – 23196



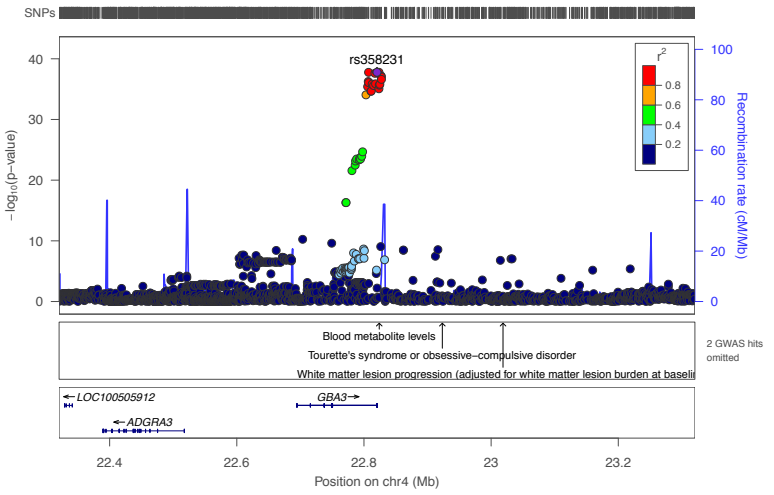
X – 23299



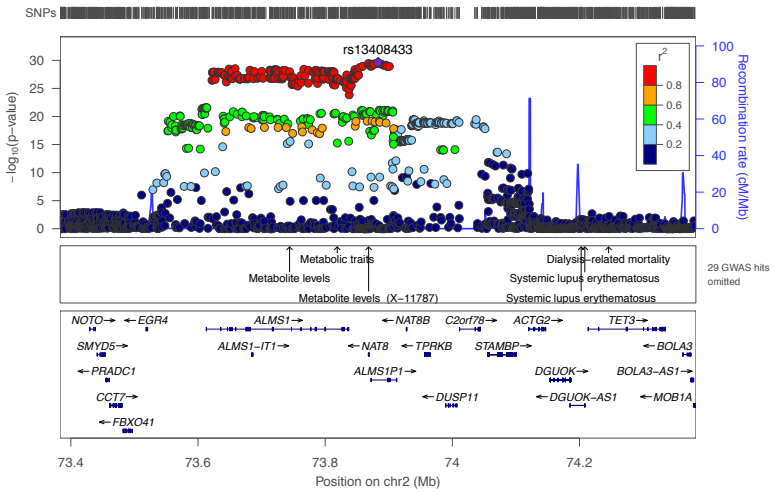
X – 23311



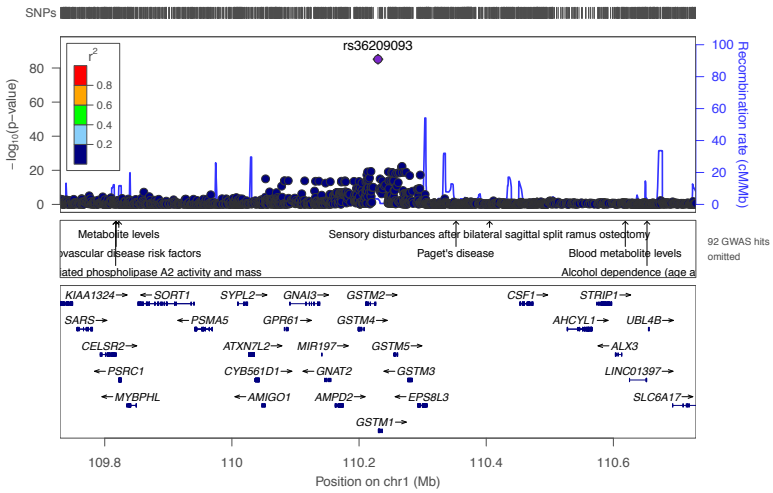
X – 23314



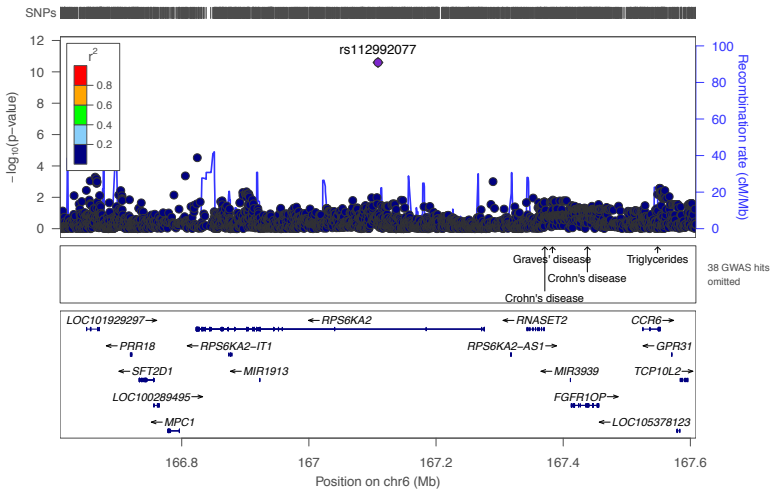
X – 23423



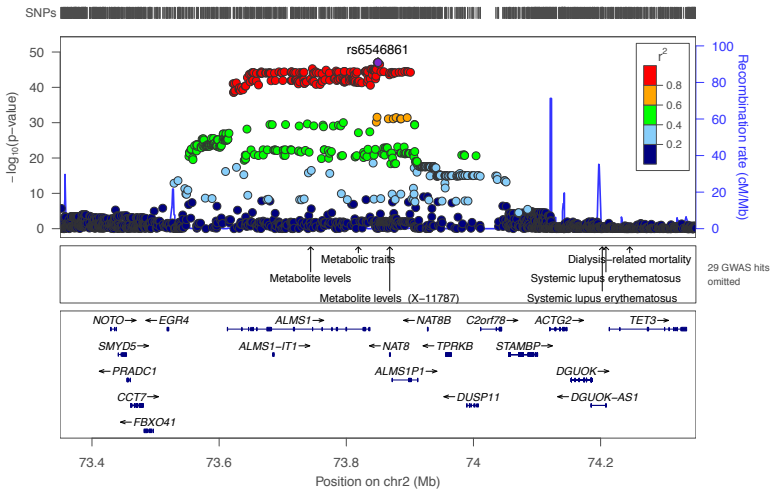
X – 23438



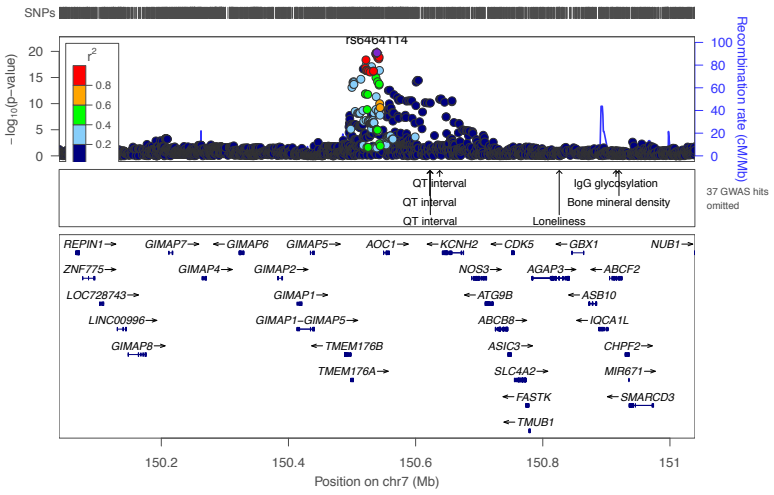
X – 23581



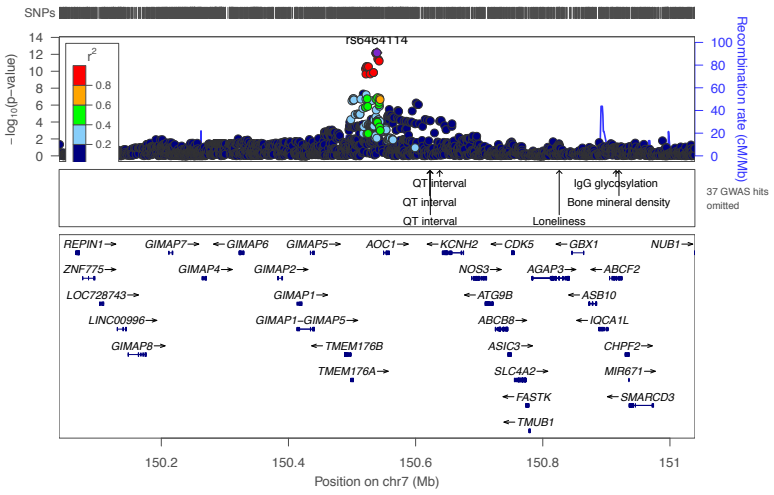
X – 23590



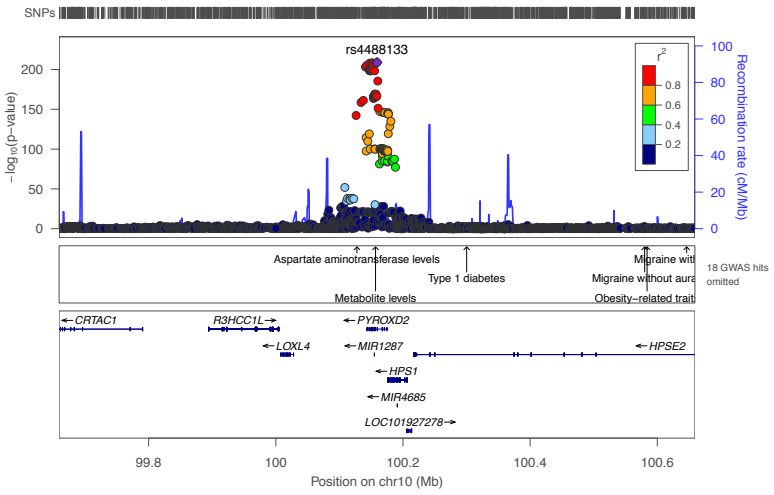
X – 23656



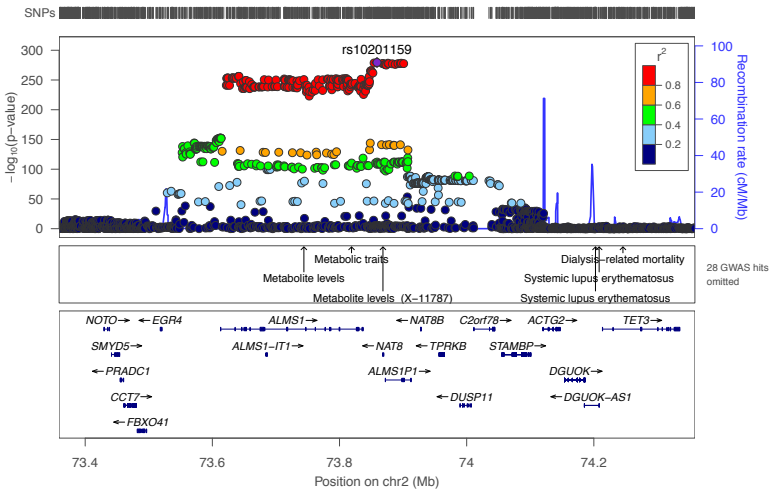
X – 23668



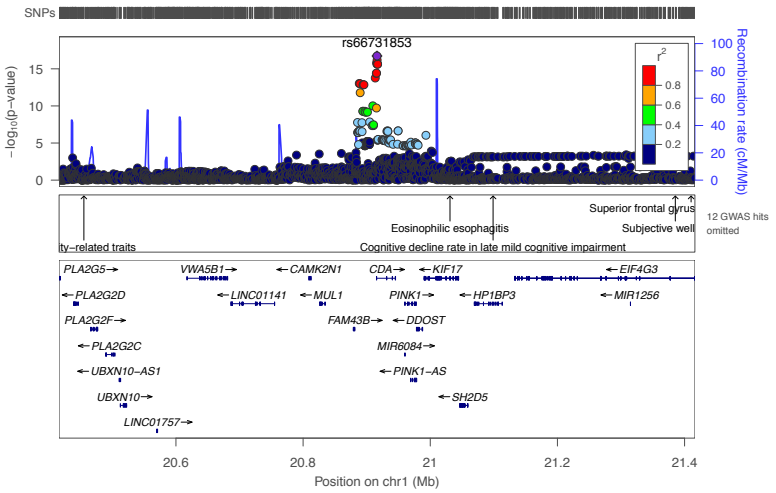
X – 23776



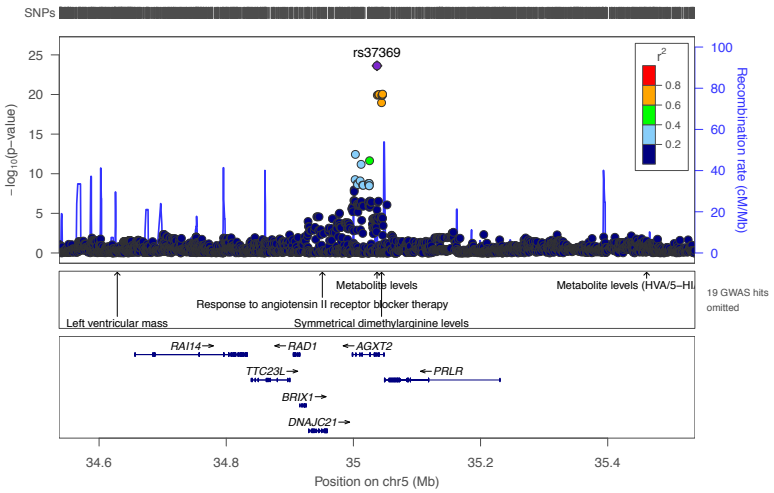
X – 13698



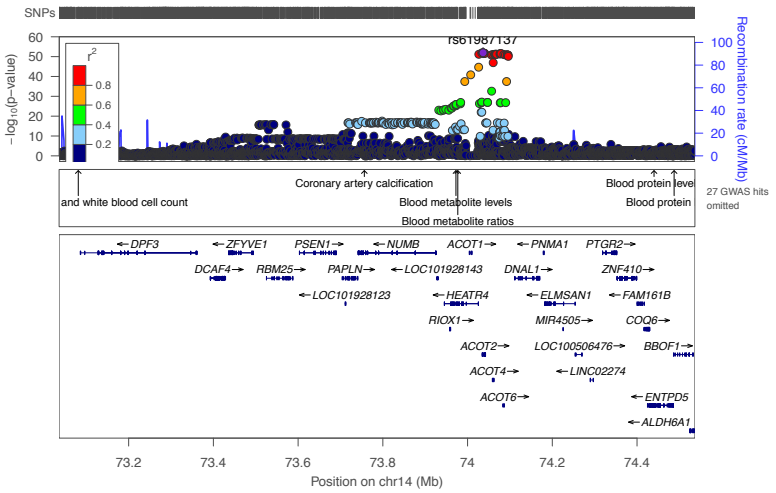
cytidine



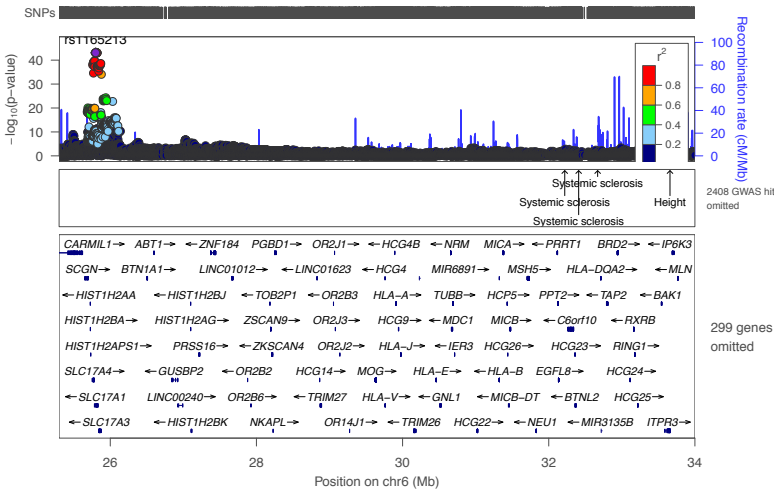
X – 24387



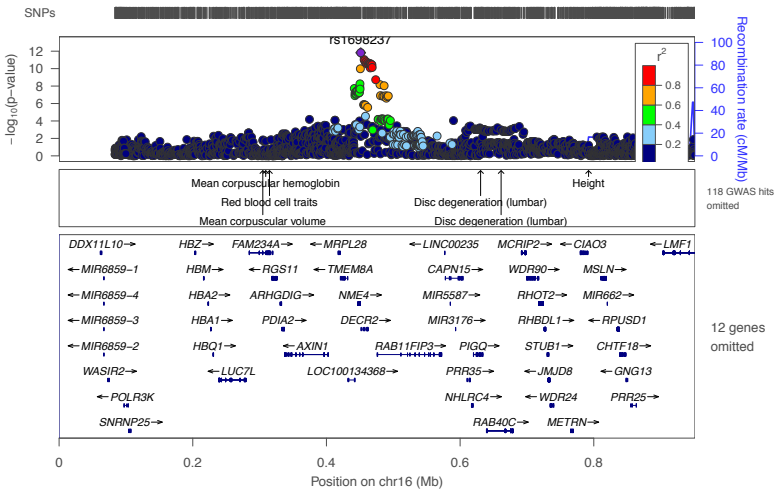
X – 24330



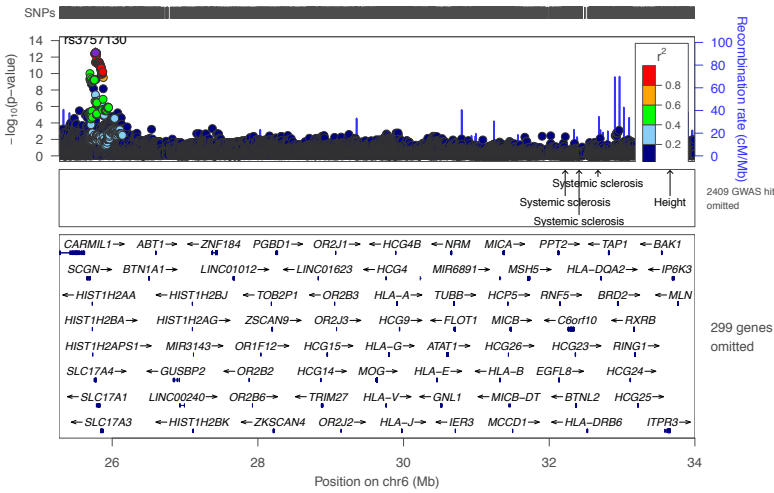
X – 24339



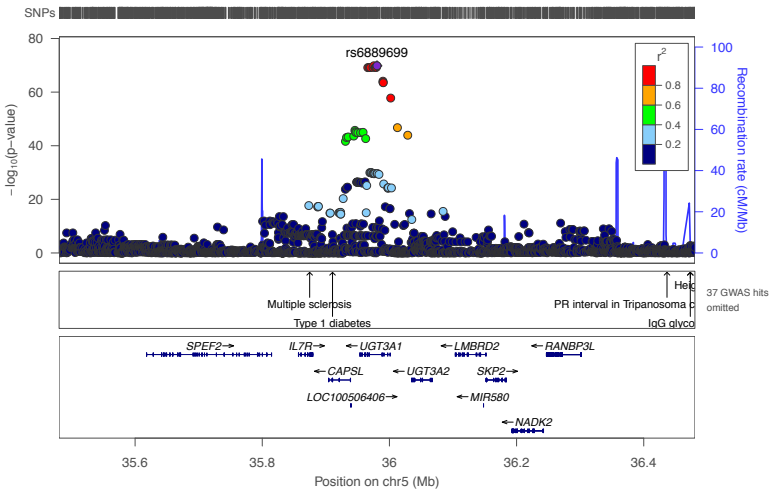
X – 24339



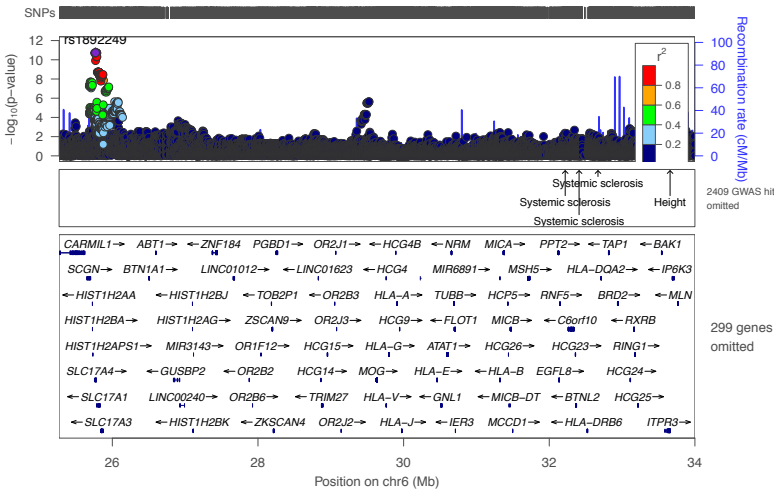
X – 24347



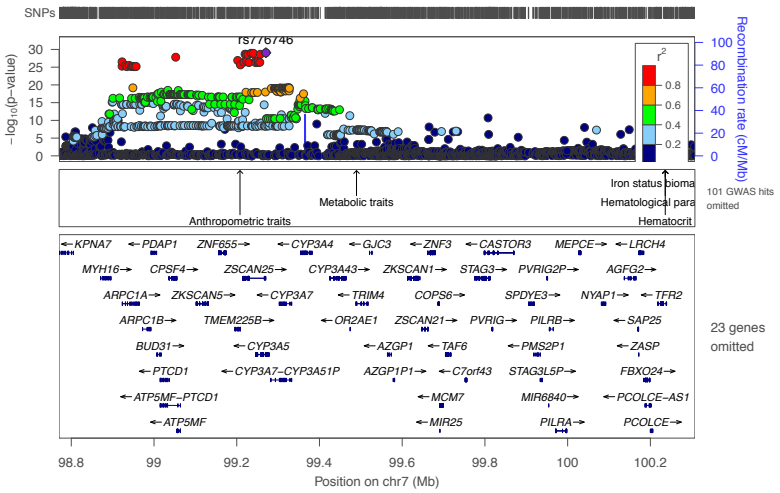
X – 24348



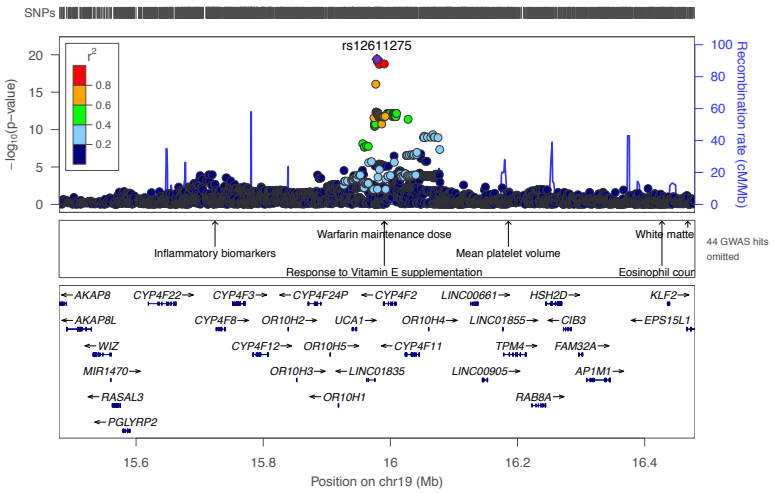
X – 24350



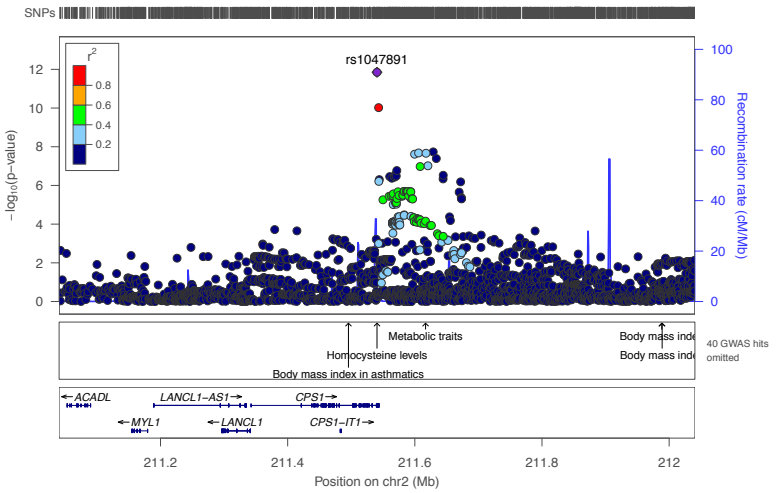
X – 24353



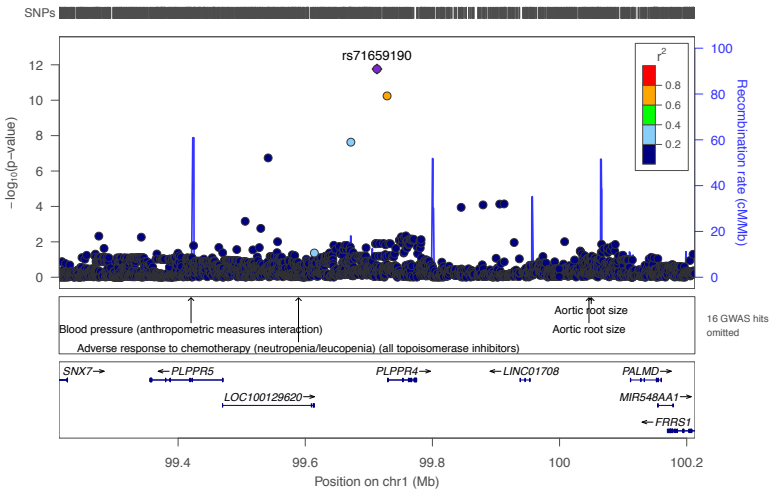
X – 24359



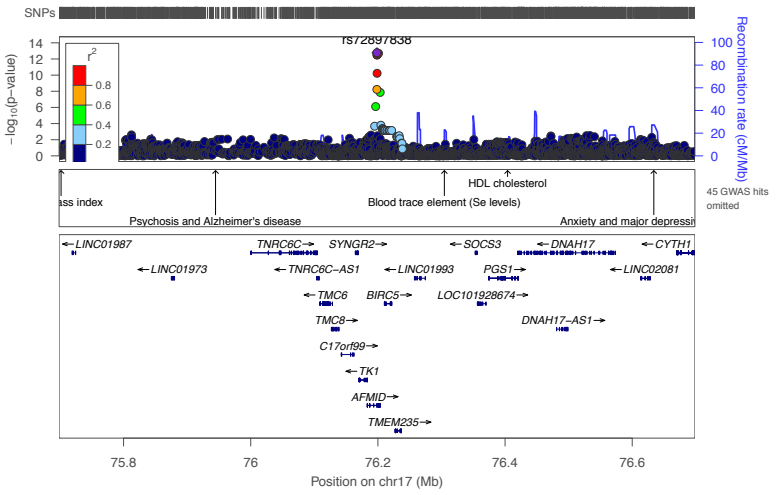
X – 24402



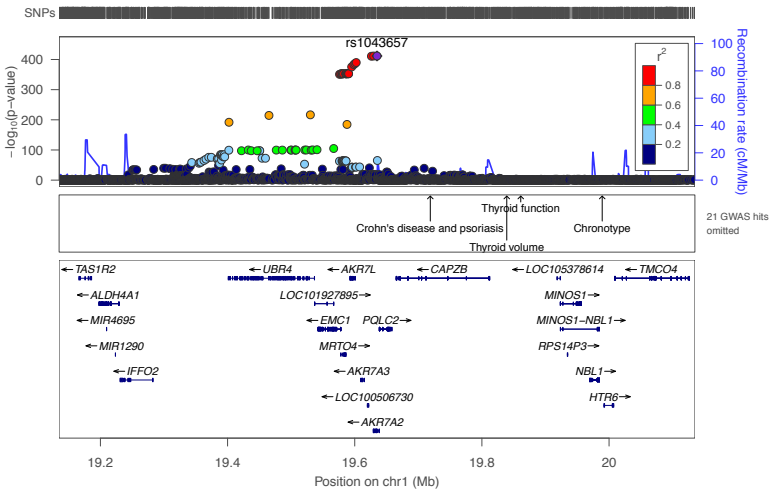
X – 24403



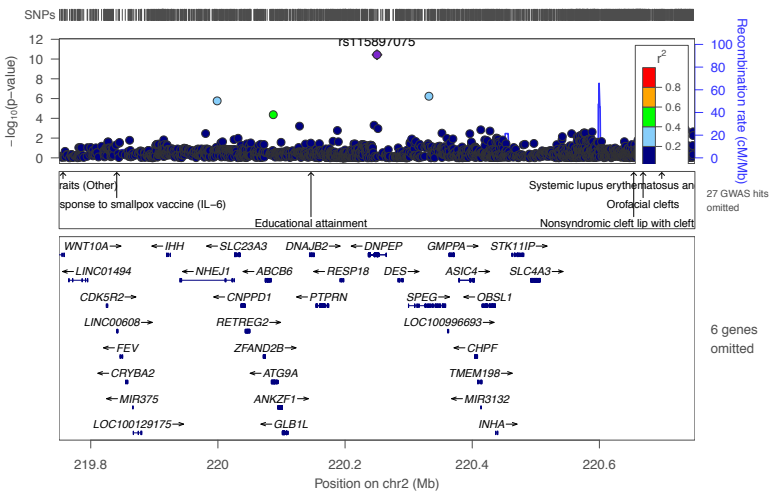
X – 24455



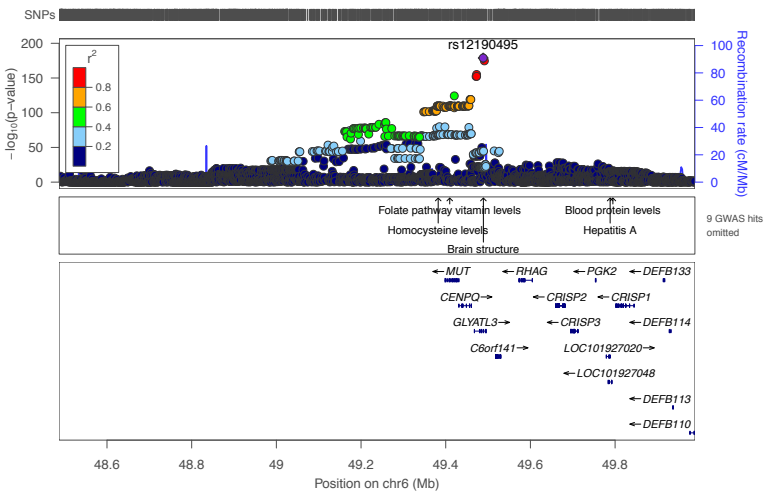
X – 24462

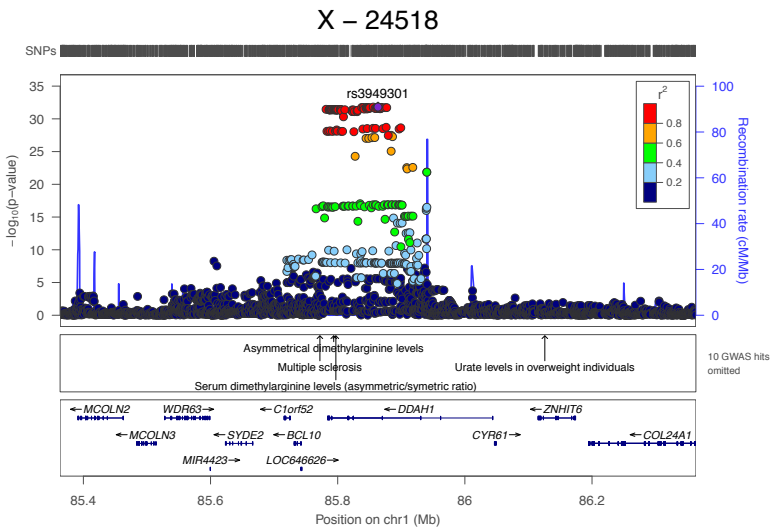


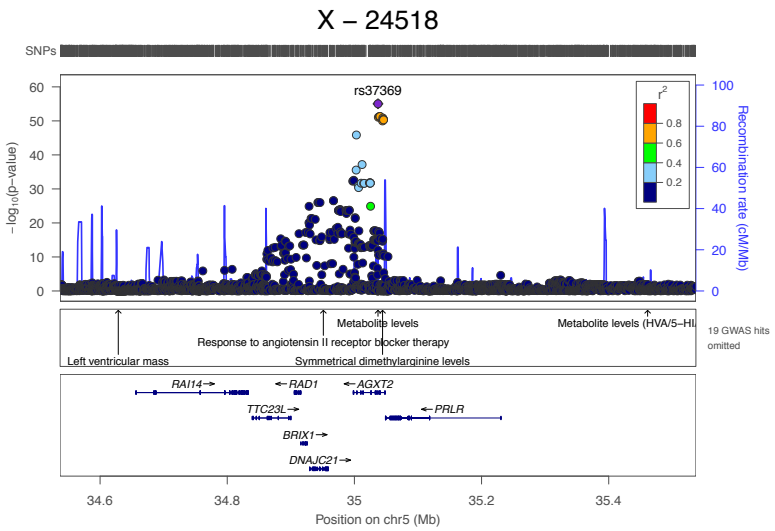
X – 24462

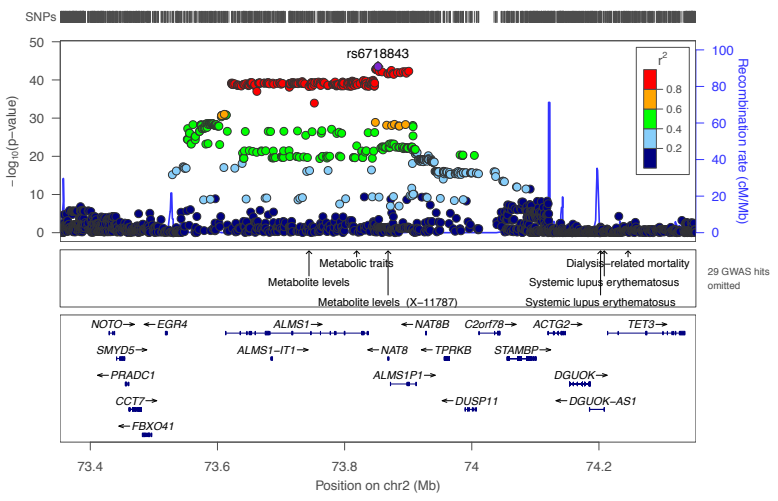


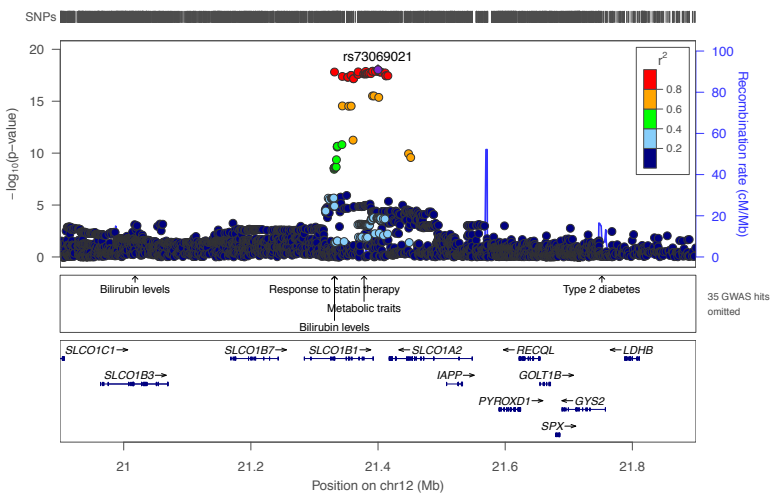
X – 24515



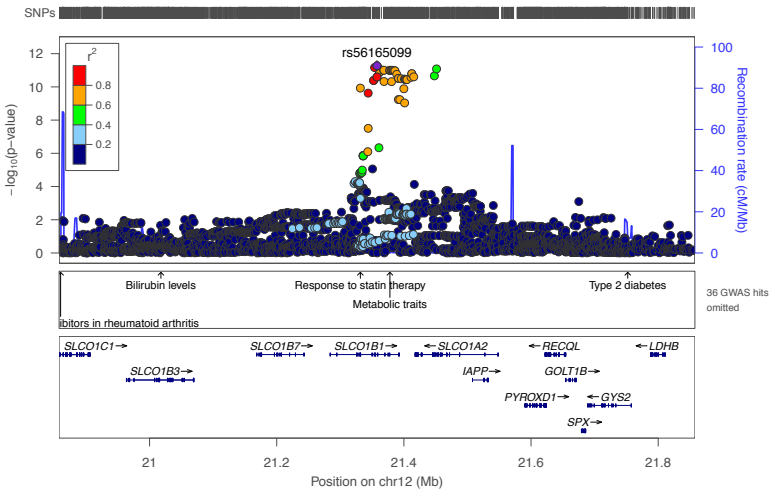

X – 24518
SNPs
rs3949301
r^2
0.8
0.6
0.4
0.2
$-\log_{10}$(p-value)
Recombination rate (cM/Mb)
Asymmetrical dimethylarginine levels
Multiple sclerosis
Serum dimethylarginine levels (asymmetric/symetric ratio)
Urate levels in overweight individuals
10 GWAS hits omitted
MCOLN2
WDR63
C1orf52
DDAH1
ZNHIT6
MCOLN3
SYDE2
BCL10
CYR61
COL24A1
MIR4423
LOC646626
Position on chr1 (Mb)


X – 24518
SNPs
rs37369
r^2
0.8
0.6
0.4
0.2
$-\log_{10}$(p-value)
Recombination rate (cM/Mb)
Left ventricular mass
Metabolite levels
Response to angiotensin II receptor blocker therapy
Symmetrical dimethylarginine levels
Metabolite levels (HVA/5–HI
19 GWAS hits omitted
RAI14
RAD1
AGXT2
TTC23L
PRLR
BRIX1
DNAJC21
Position on chr5 (Mb)

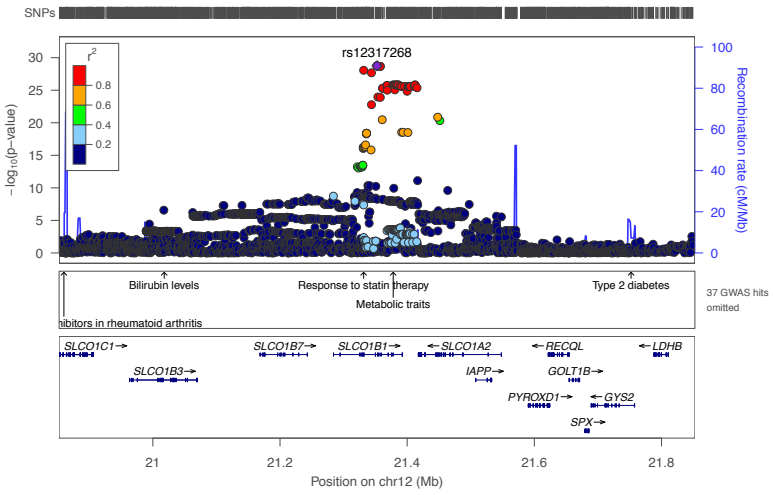

X – 24519
SNPs
rs6718843
r^2
0.8
0.6
0.4
0.2
$-\log_{10}$(p-value)
Recombination rate (cM/Mb)
Metabolic traits
Metabolite levels
Metabolite levels (X–11787)
Dialysis–related mortality
Systemic lupus erythematosus
Systemic lupus erythematosus
29 GWAS hits omitted
NOTO
EGR4
ALMS1
NAT8B
C2orf78
ACTG2
TET3
SMYD5
ALMS1–IT1
NAT8
TPRKB
STAMBP
PRADC1
ALMS1P1
DGUOK
CCT7
DUSP11
DGUOK–AS1
FBXO41
Position on chr2 (Mb)


X – 24546
SNPs
rs73069021
r^2
0.8
0.6
0.4
0.2
$-\log_{10}$(p-value)
Recombination rate (cM/Mb)
Bilirubin levels
Response to statin therapy
Metabolic traits
Bilirubin levels
Type 2 diabetes
35 GWAS hits omitted
SLCO1C1
SLCO1B7
SLCO1B1
SLCO1A2
RECQL
LDHB
SLCO1B3
IAPP
GOLT1B
PYROXD1
GYS2
SPX
Position on chr12 (Mb)

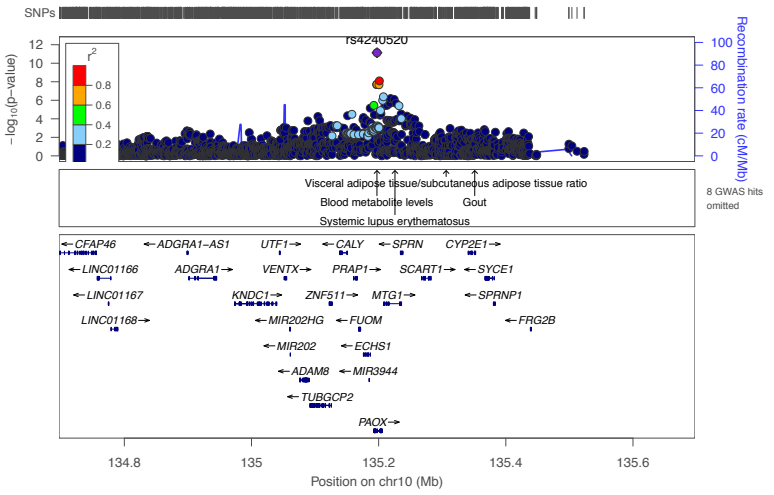
glycodeoxycholate sulfate



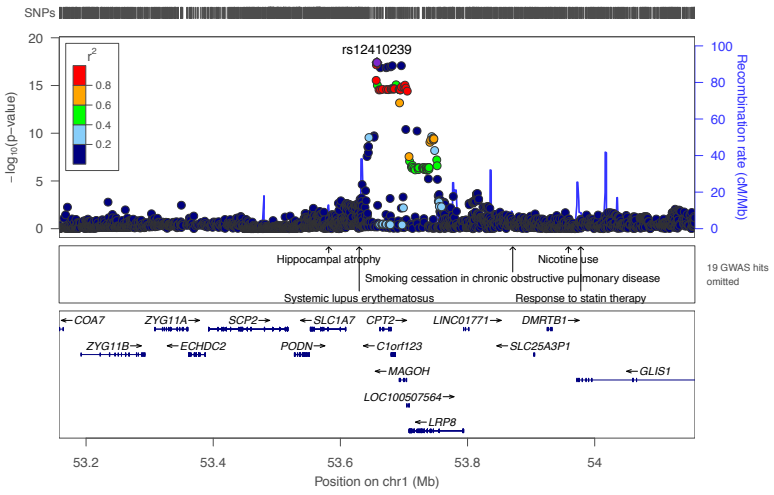
glycochenodeoxycholate glucuronide (1)



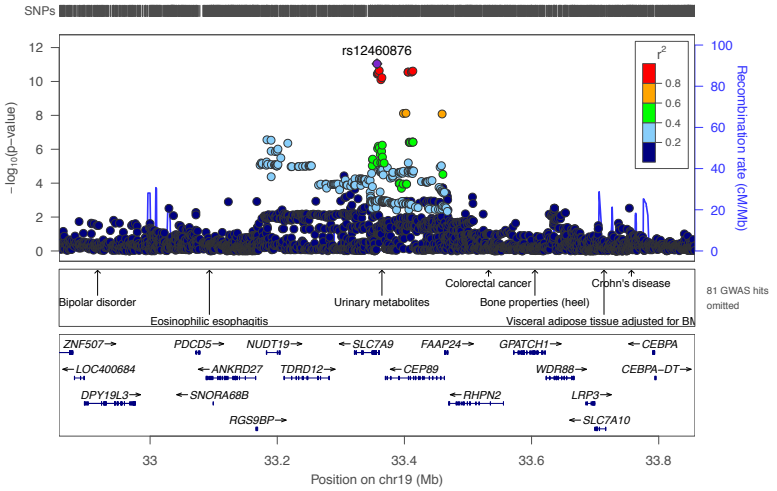
N1,N12−diacetylspermine



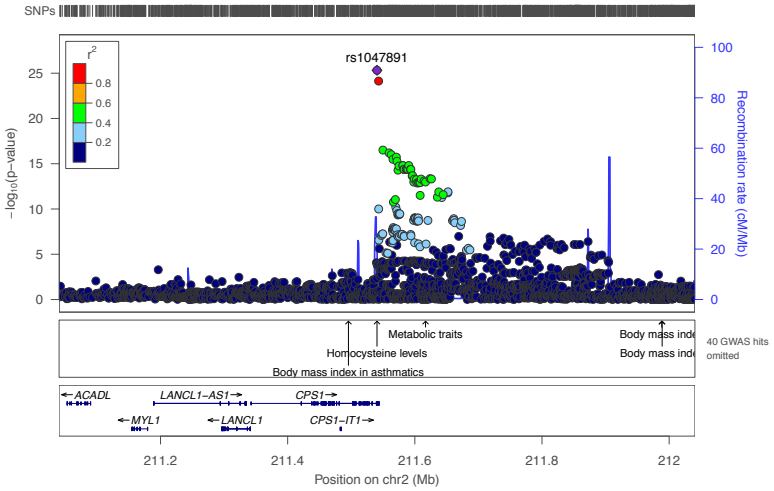
methylsuccinoylcarnitine



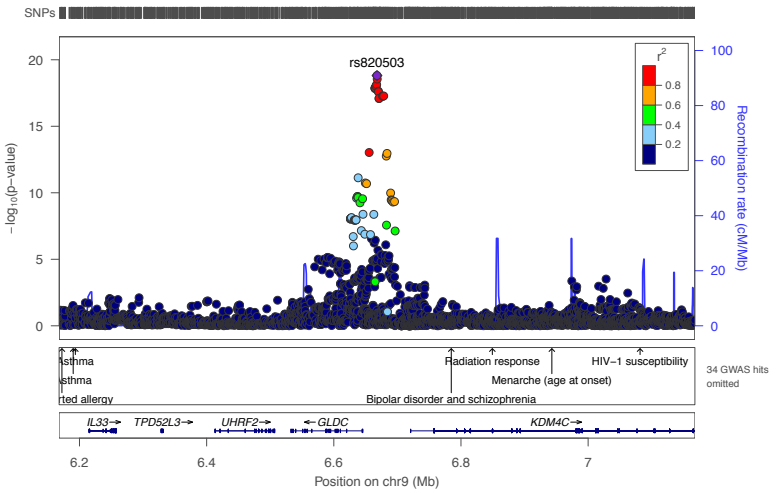
X – 24736



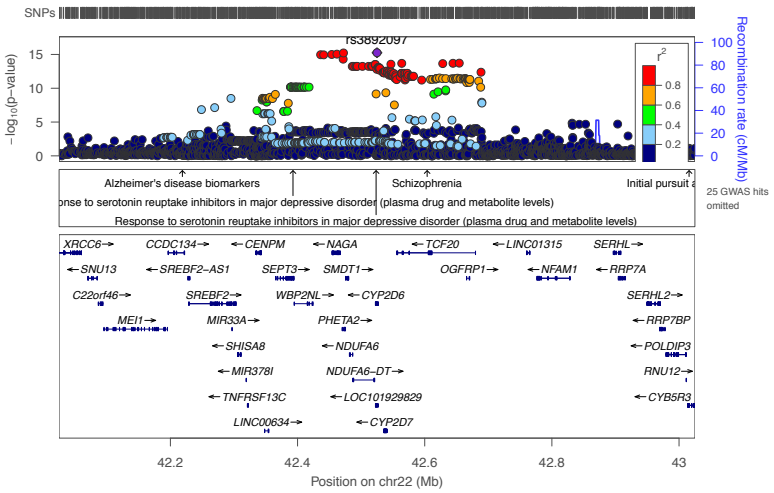
3–methylglutarate/2–methylglutarate



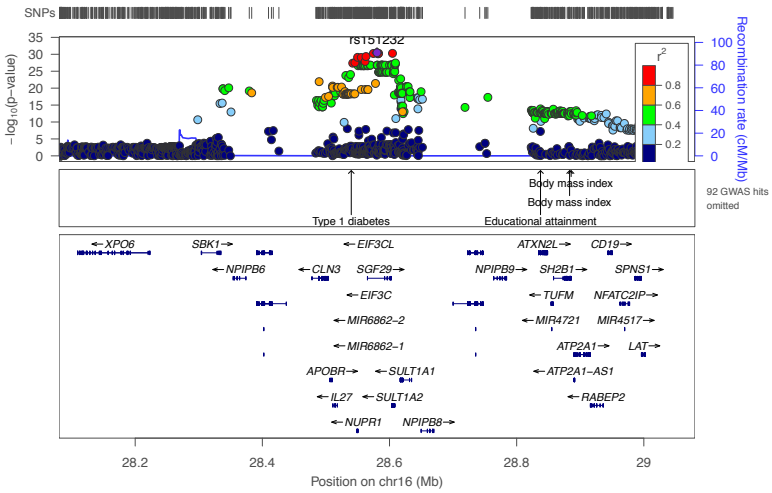
3–methylglutarate/2–methylglutarate



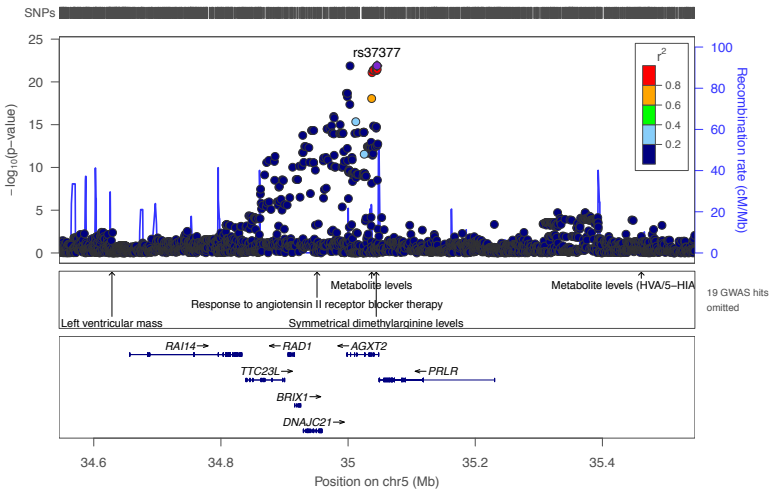
alpha–hydroxymetoprolol



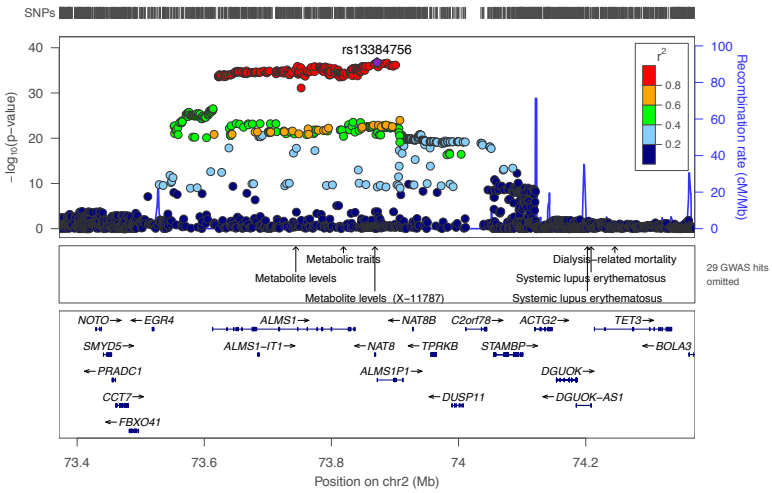
furaneol sulfate



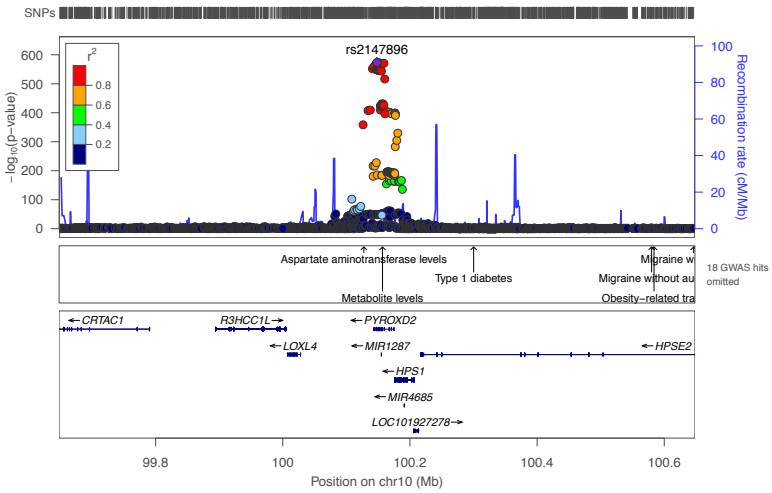
X – 24796



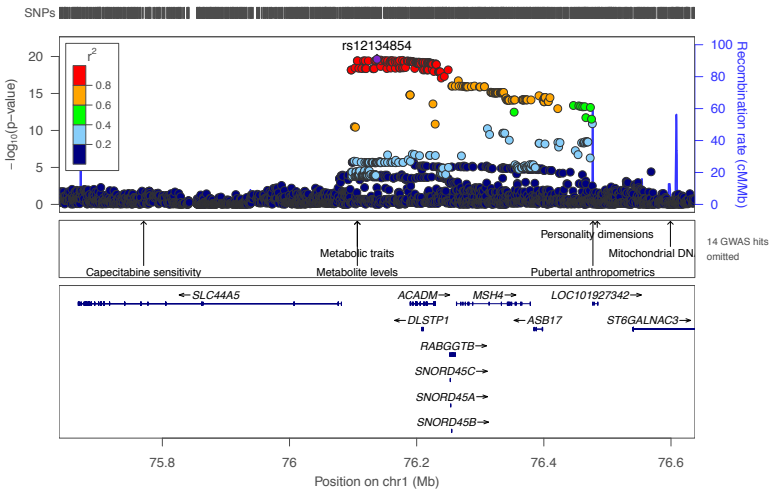
X – 24809



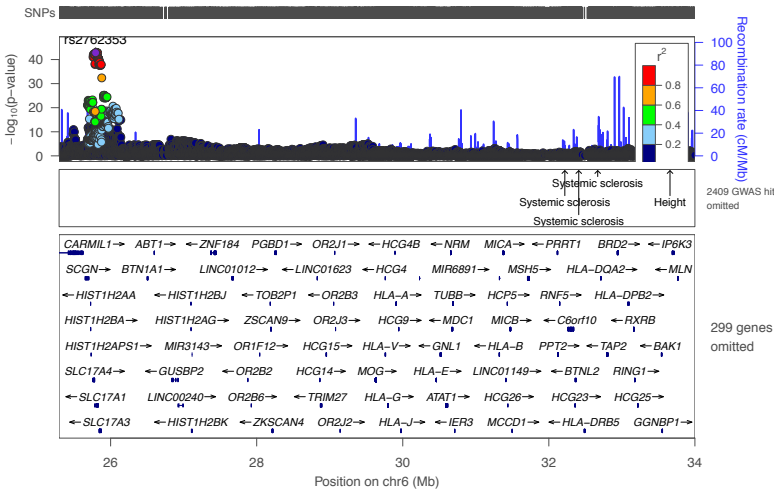
X – 24809



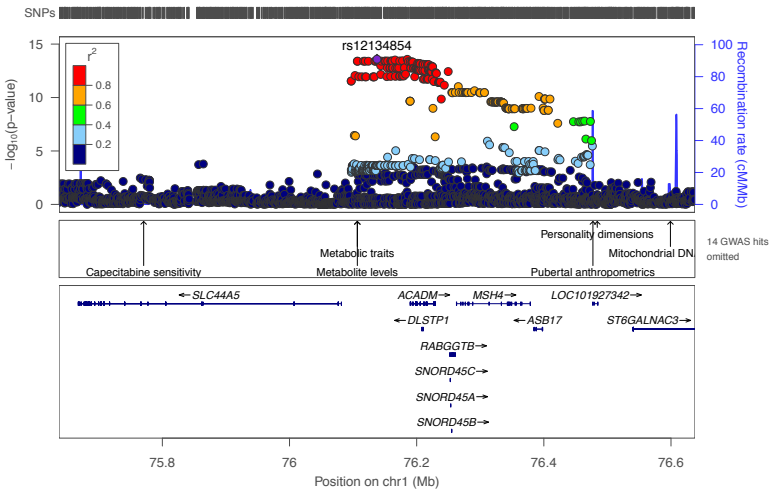
hexanoylglutamine



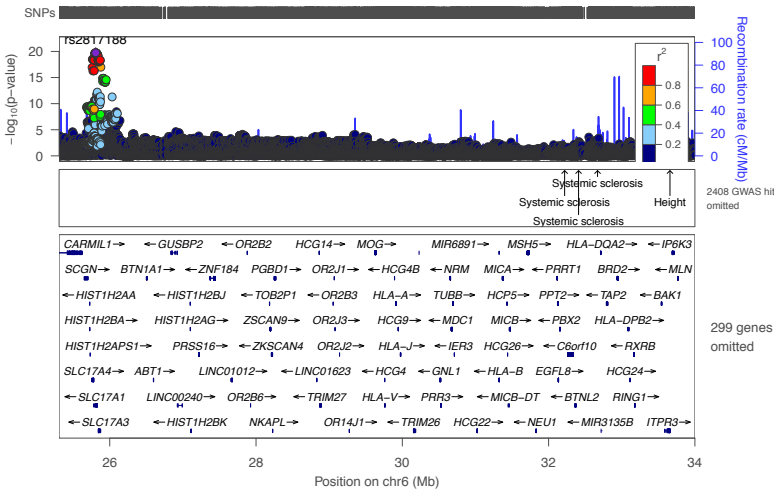
hexanoylglutamine



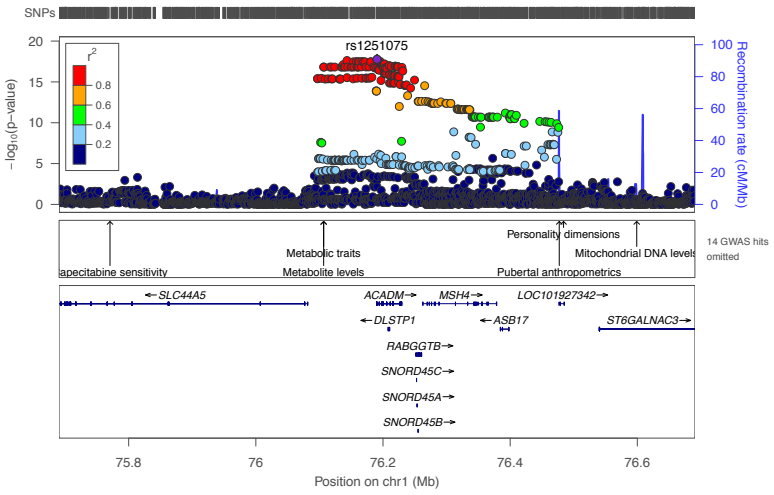
heptanoylglutamine



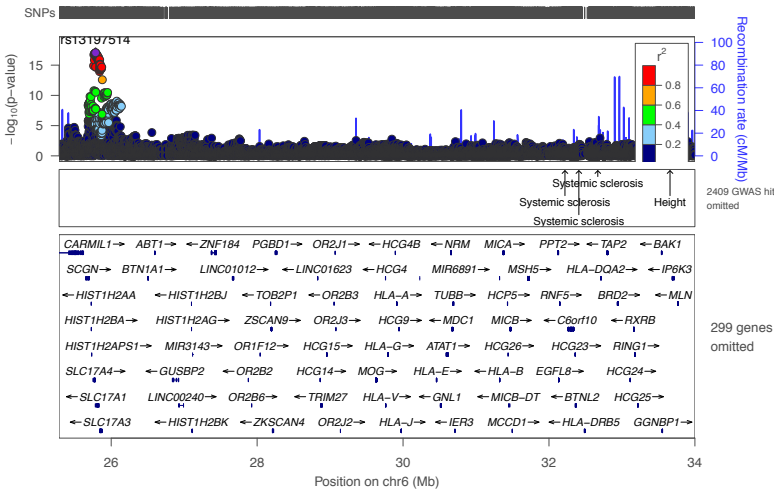
heptanoylglutamine



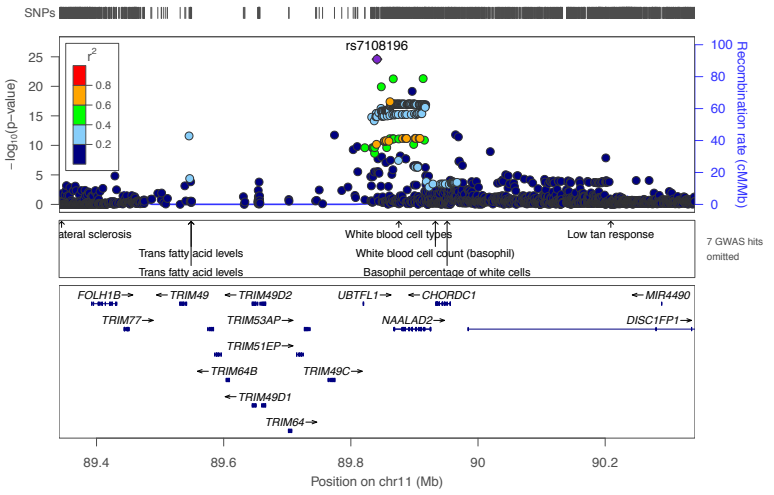
N−octanoylglutamine



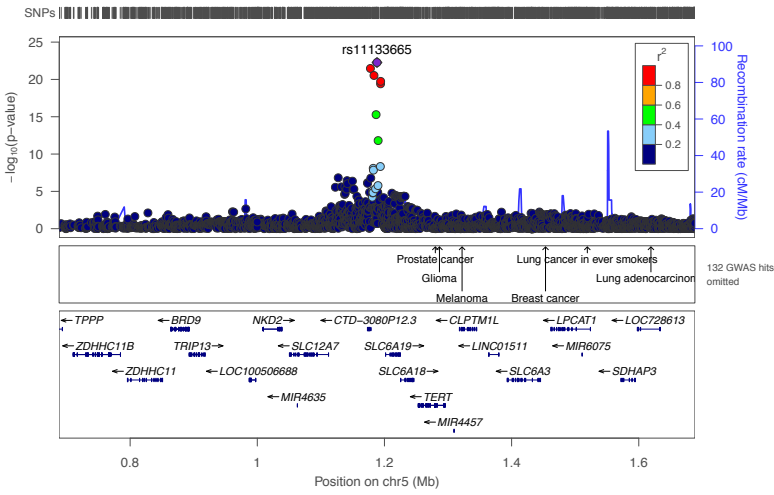
N−octanoylglutamine



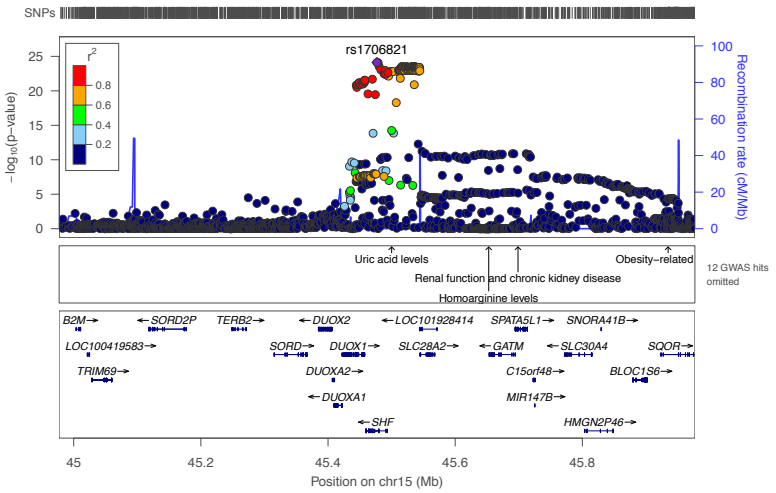
beta−citrylglutamate



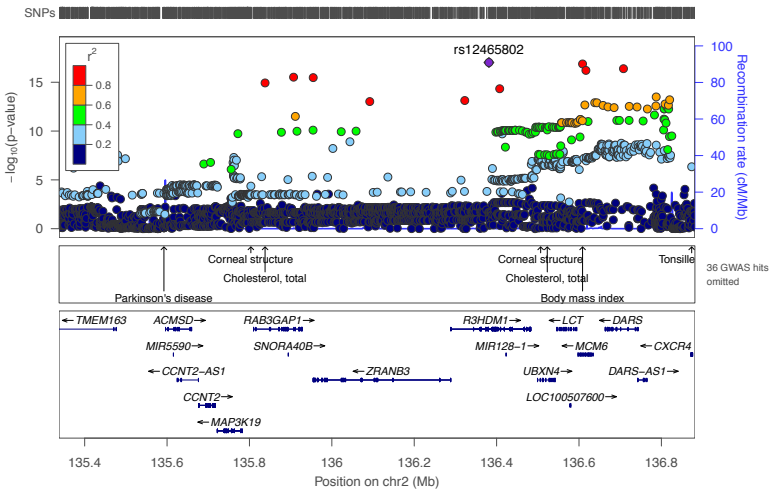
tryptophan



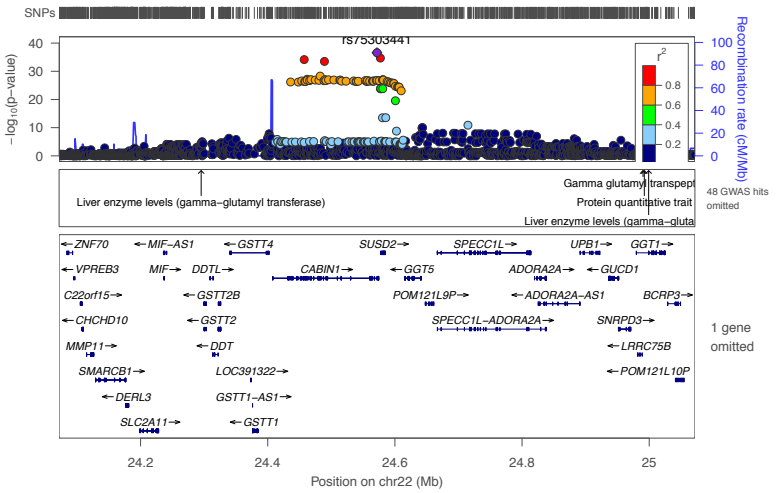
adenosine



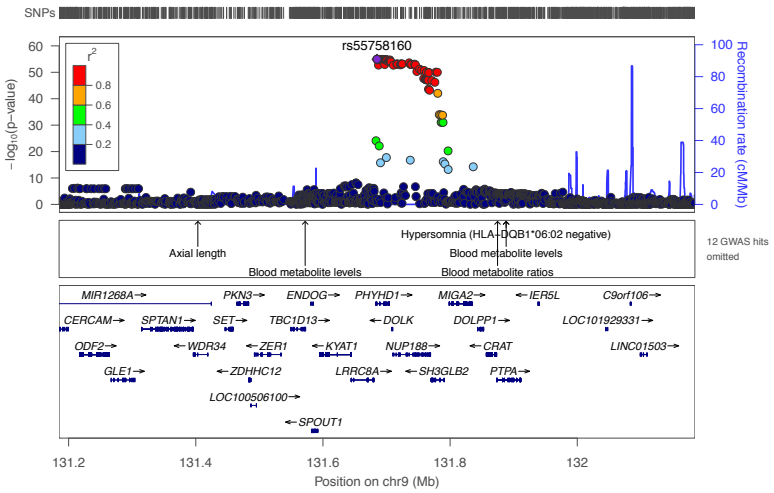
lactose



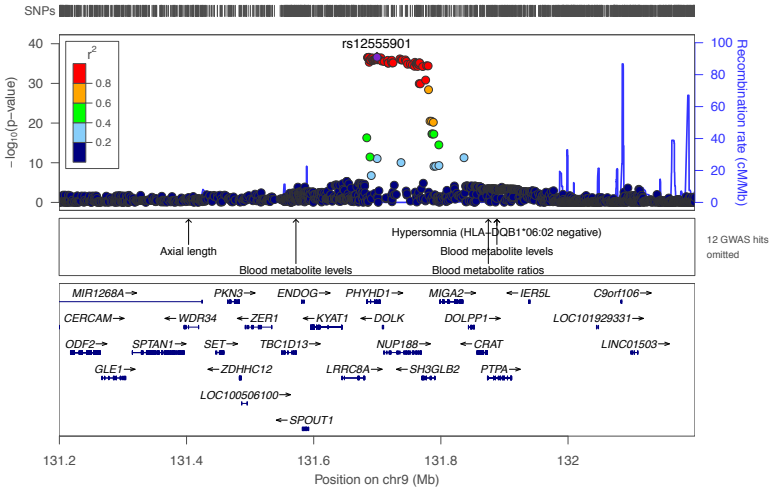
glycerophosphoserine*



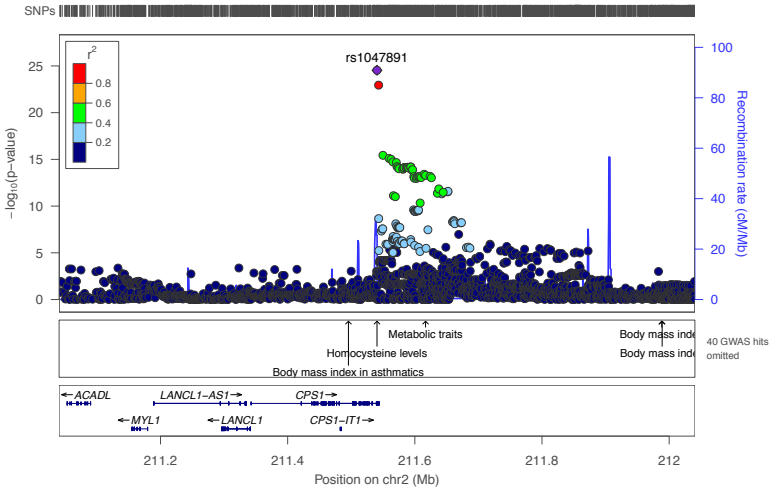
2'−O−methylcytidine



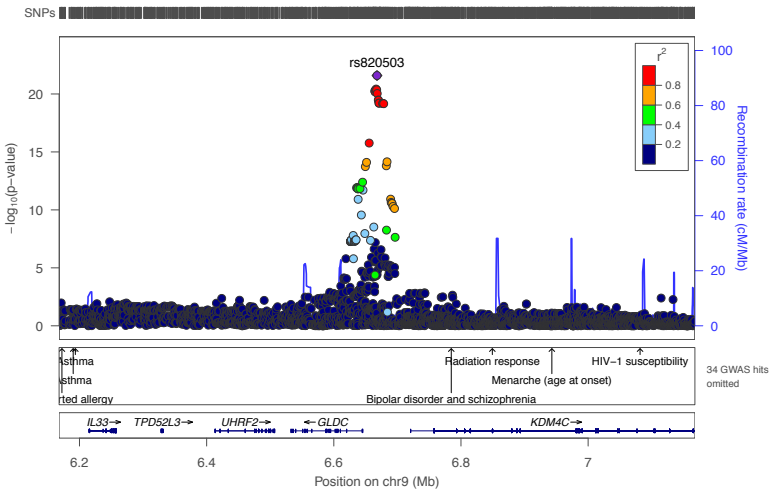
2'−O−methyluridine



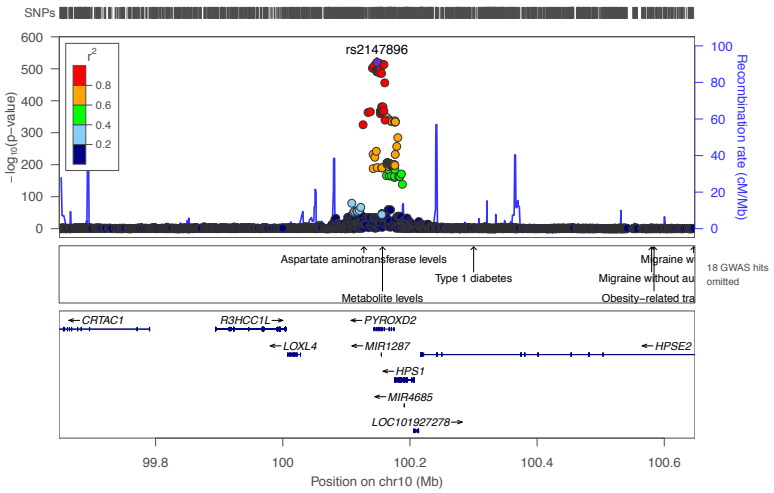
3−methylglutaconate



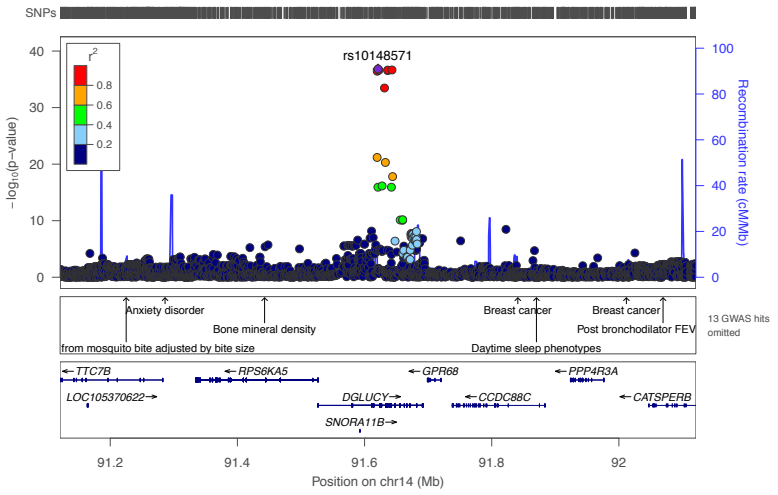
3−methylglutaconate



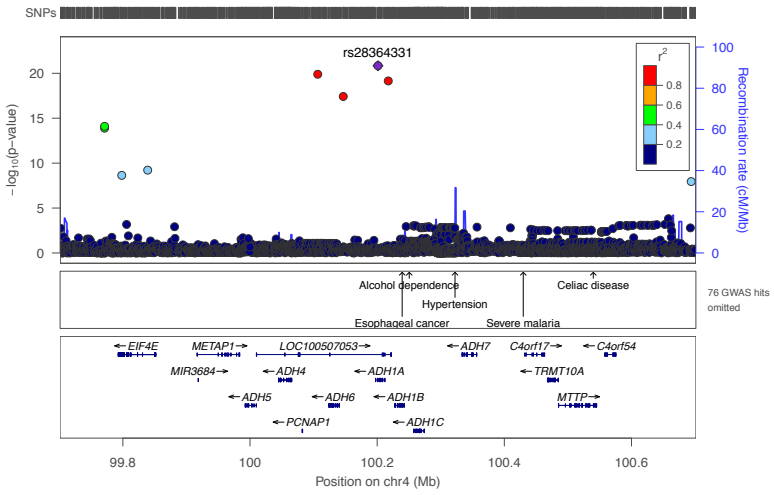
X − 24983



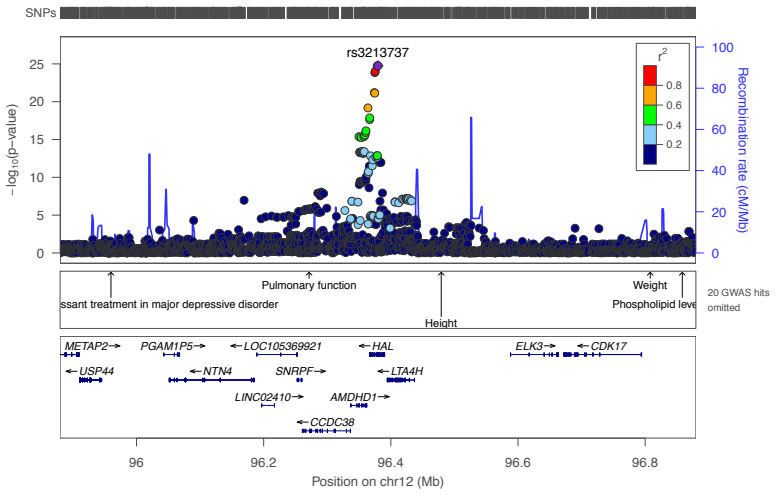
glutamate



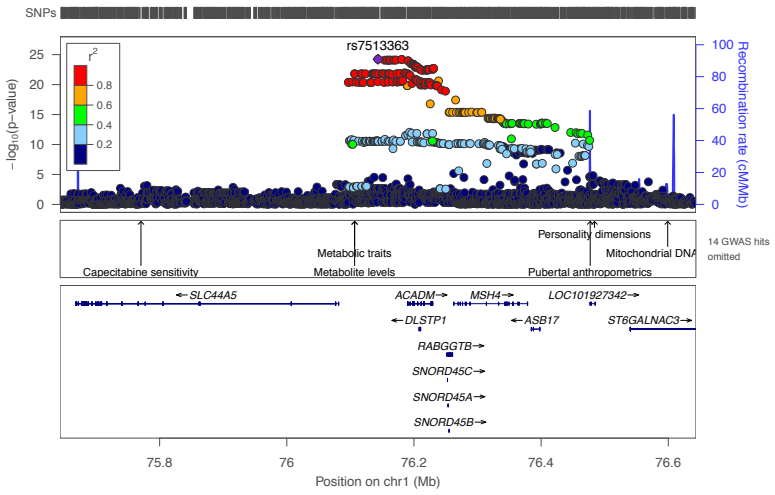
gluconate



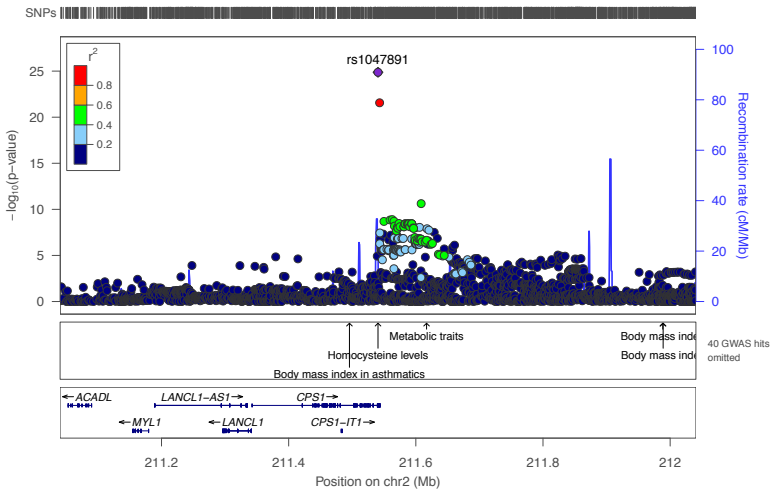
trans−urocanate



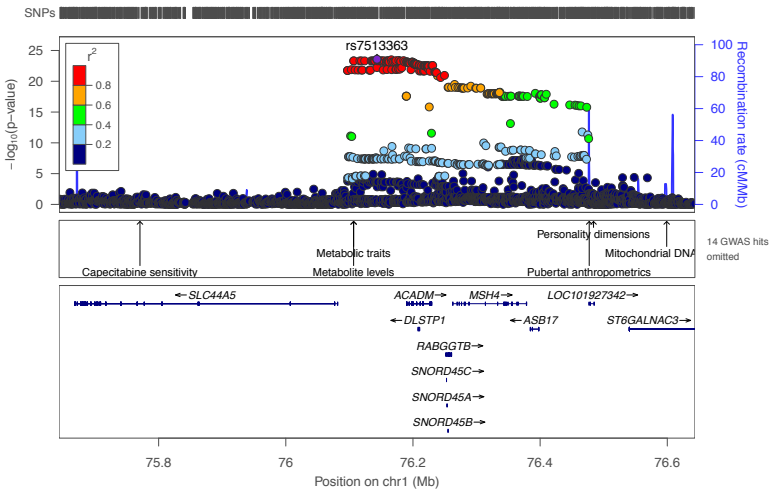
isocaproylglycine



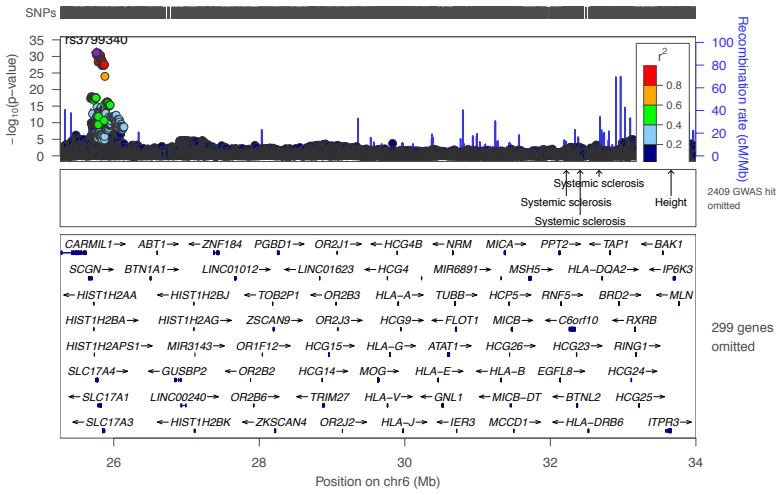
isocaproylglycine



isocaproylglutamine

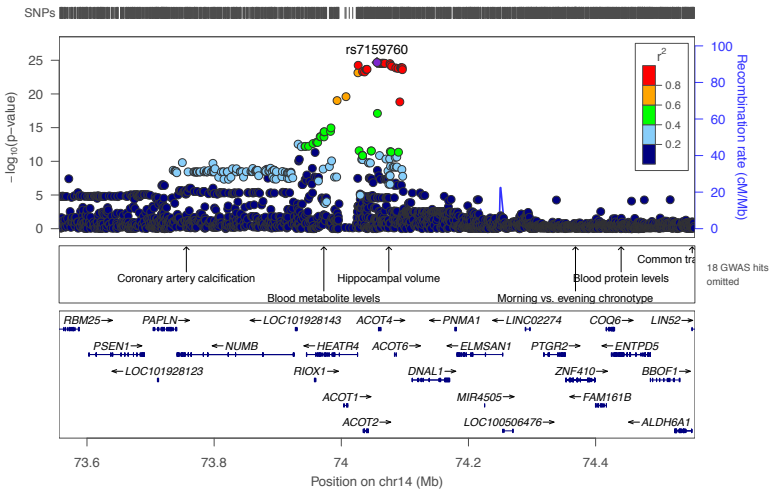


isocaproylglutamine

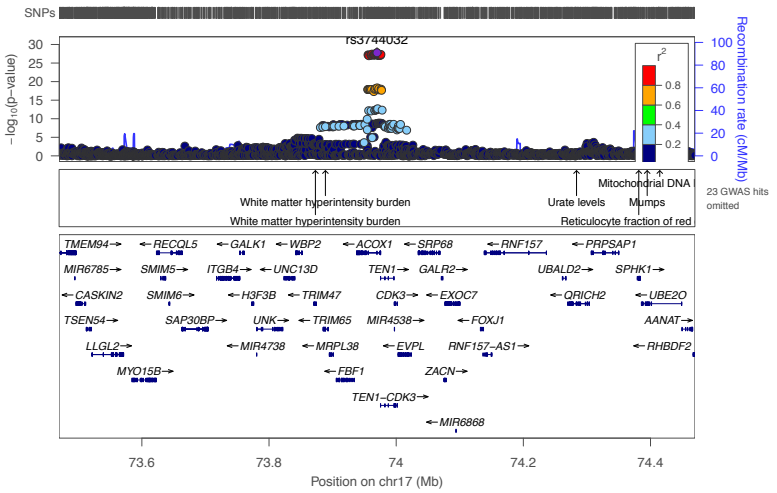


299 genes omitted

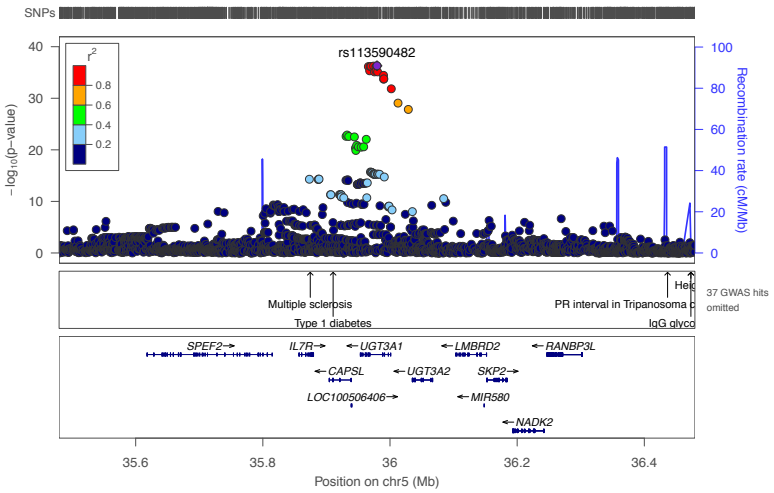
nonenedioate (C9:1−DC)*



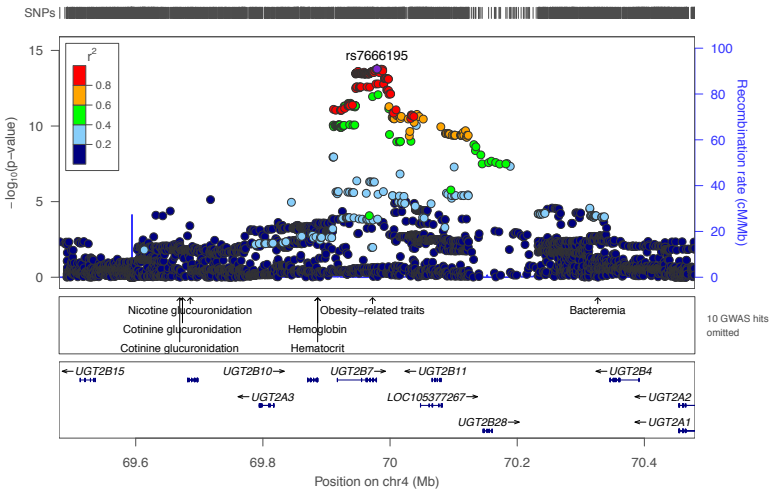
nonenedioate (C9:1−DC)*



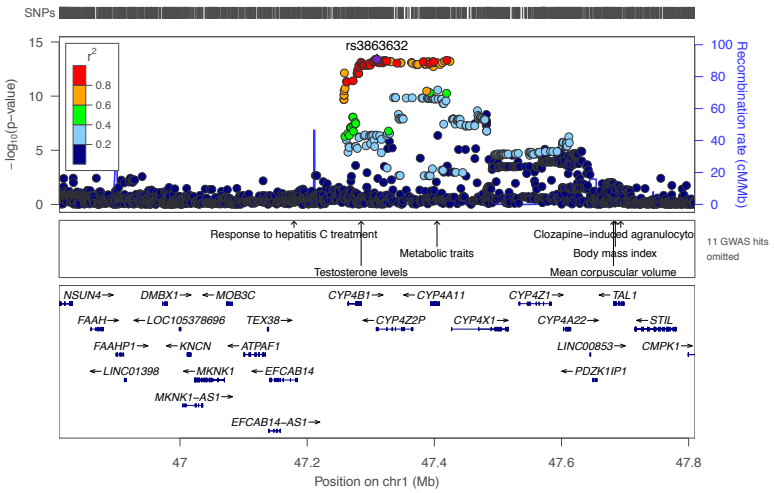
N−acetylglucosamine conjugate of C24H40O4 bile acid**



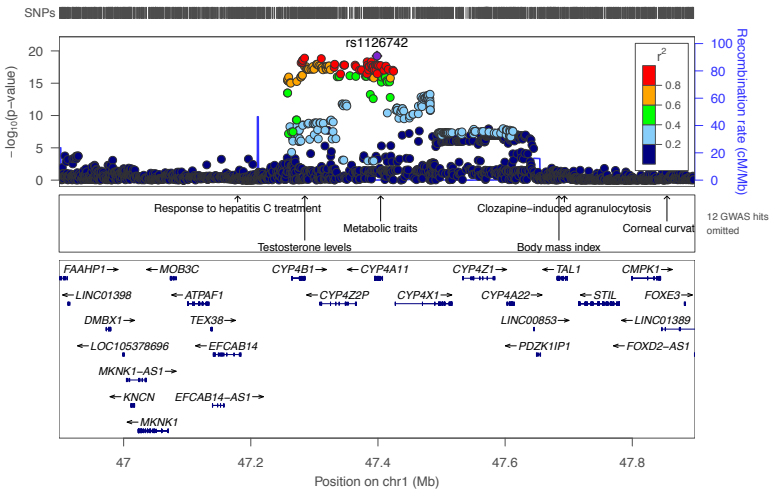
glucuronide of C12H22O4 (2)*



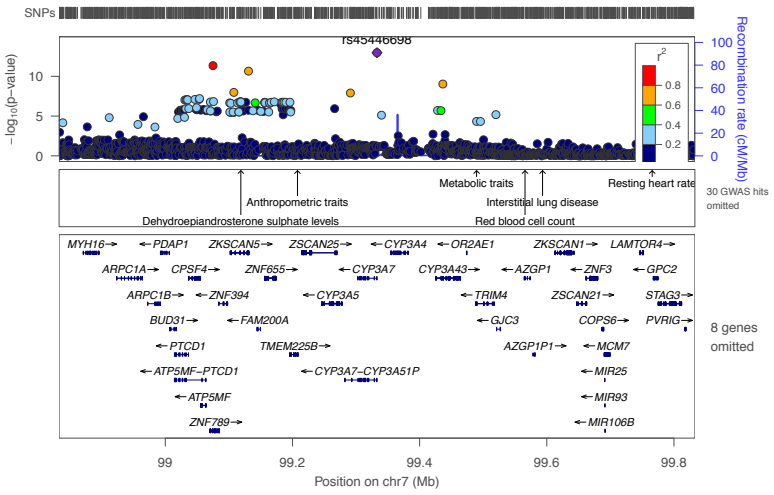
glucuronide of C10H18O2 (11)*



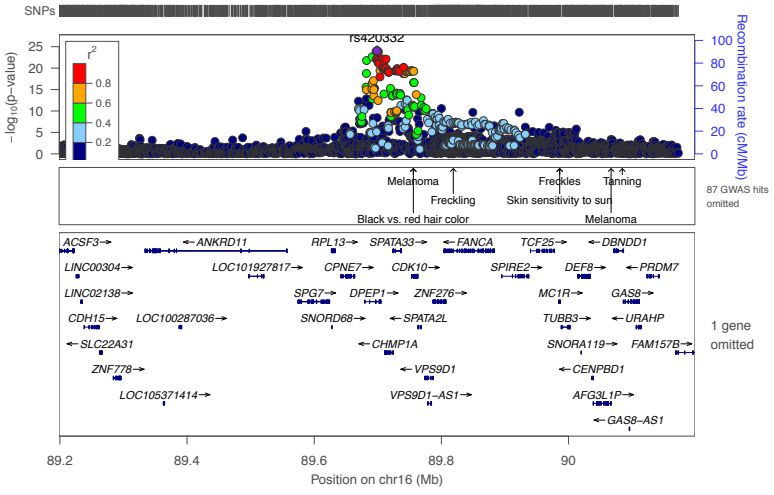
glucuronide of C10H18O2 (12)*



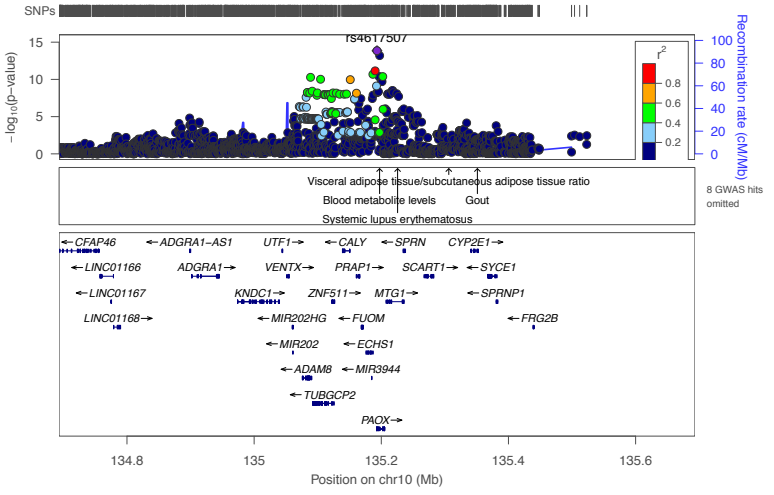
glyco−beta−muricholate**



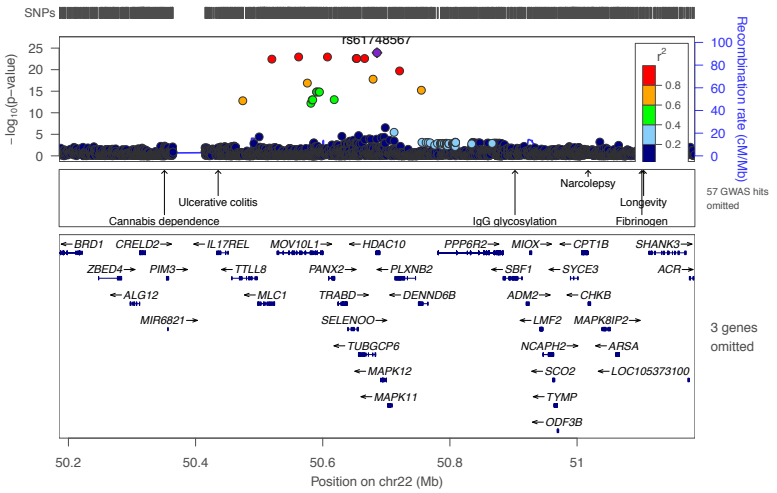
cysteinylglycine disulfide*



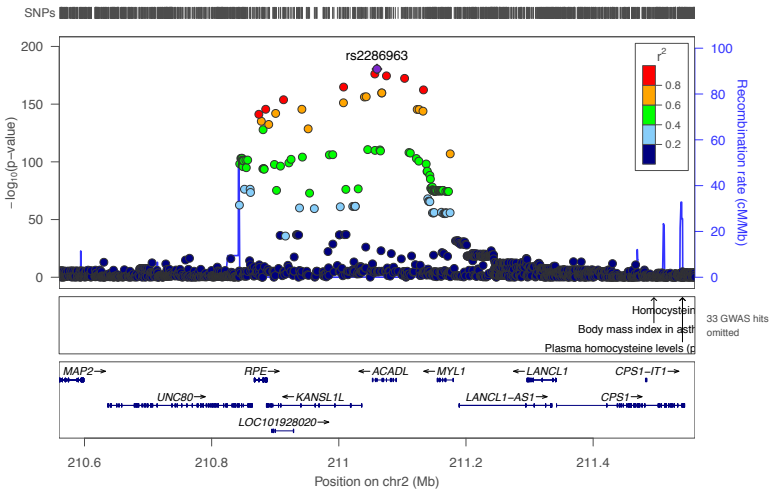
diacetylspermidine*



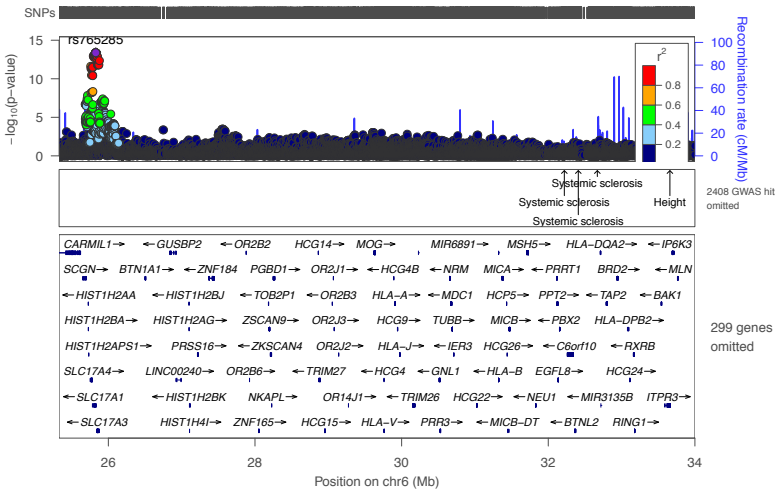
diacetylspermidine*



glycine conjugate of C9H16O2*



gamma−CEHC sulfate*



phenylalanine