

In the format provided by the authors and unedited.

Genome-wide association and HLA fine-mapping studies identify risk loci and genetic pathways underlying allergic rhinitis

Johannes Waage^{1,74}, Marie Standl^{2,74}, John A. Curtin³, Leon E. Jessen¹, Jonathan Thorsen¹, Chao Tian⁴, Nathan Schoettler⁵, The 23andMe Research Team⁶, AAGC collaborators⁶, Carlos Flores^{1,7,8}, Abdel Abdellaoui^{9,10}, Tarunveer S. Ahluwalia¹, Alexessander C. Alves¹¹, Andre F. S. Amaral¹², Josep M. Antó^{13,14,15,16}, Andreas Arnold¹⁷, Amalia Barreto-Luis⁷, Hansjörg Baurecht¹⁸, Catharina E. M. van Beijsterveldt⁹, Eugene R. Bleeker¹⁹, Sílvia Bonàs-Guarch²⁰, Dorret I. Boomsma^{9,21}, Susanne Brix^{1,22}, Supinda Bunyavanich^{1,23}, Esteban G. Burchard^{24,25}, Zhanghua Chen²⁶, Ivan Curjuric^{27,28}, Adnan Custovic²⁹, Herman T. den Dekker^{30,31,32}, Shyamali C. Dharmage³³, Julia Dmitrieva³⁴, Liesbeth Duijts^{30,32,35}, Markus J. Ege^{1,36}, W. James Gauderman²⁶, Michel Georges³⁴, Christian Gieger^{2,37}, Frank Gilliland²⁶, Raquel Granell³⁸, Hongsheng Gui³⁹, Torben Hansen^{1,40}, Joachim Heinrich^{2,41}, John Henderson³⁸, Natalia Hernandez-Pacheco^{7,42}, Patrick Holt⁴³, Medea Imboden^{27,28}, Vincent W. V. Jaddoe^{30,31,44}, Marjo-Riitta Jarvelin^{11,45,46,47}, Deborah L. Jarvis¹², Kamilla K. Jensen^{1,48}, Ingileif Jónsdóttir^{1,49,50}, Michael Kabesch⁵¹, Jaakko Kaprio^{1,52,53,54}, Ashish Kumar^{27,28,55}, Young-Ae Lee^{56,57}, Albert M. Levin⁵⁸, Xingnan Li⁵⁹, Fabian Lorenzo-Diaz⁴², Erik Melén^{55,60}, Josep M. Mercader^{1,20,61,62}, Deborah A. Meyers¹⁹, Rachel Myers⁵, Dan L. Nicolae⁵, Ellen A. Nohr⁶³, Teemu Palviainen^{1,53}, Lavinia Paternoster³⁸, Craig E. Pennell⁶⁴, Göran Pershagen^{55,65}, Maria Pino-Yanes^{1,7,8,42}, Nicole M. Probst-Hensch^{27,28}, Franz Rüschenhoff^{1,56}, Angela Simpson³, Kari Stefansson^{49,50}, Jordi Sunyer¹³, Gardar Sveinbjornsson⁴⁹, Elisabeth Thiering^{2,66}, Philip J. Thompson⁶⁷, Maties Torrent⁶⁸, David Torrents^{20,69}, Joyce Y. Tung⁴, Carol A. Wang^{1,64}, Stephan Weidinger¹⁸, Scott Weiss^{1,70}, Gonneke Willemsen⁹, L. Keoki Williams^{39,71}, Carole Ober⁵, David A. Hinds^{1,4}, Manuel A. Ferreira⁷², Hans Bisgaard¹, David P. Strachan⁷³ and Klaus Bønnelykke¹*

¹COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark.

²Institute of Epidemiology, Helmholtz Zentrum München-German Research Center for Environmental Health, Neuherberg, Germany. ³Division of Infection, Immunity and Respiratory Medicine, University of Manchester, Manchester Academic Health Science Centre, Manchester University NHS Foundation Trust, Manchester, UK. ⁴23andMe, Inc., Mountain View, CA, USA. ⁵Department of Human Genetics, University of Chicago, Chicago, IL, USA. ⁶A list of members and affiliations appears in the Supplementary Note. ⁷Research Unit, Hospital Universitario N.S. de Candelaria, Universidad de La Laguna, Tenerife, Spain. ⁸CIBER de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Madrid, Spain. ⁹Department of Biological Psychology, Netherlands Twin Register, VU University, Amsterdam, The Netherlands. ¹⁰Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. ¹¹Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment & Health, School of Public Health, Imperial College London, London, UK. ¹²Population Health and Occupational Disease, National Heart and Lung Institute, Imperial College London, London, UK. ¹³ISGlobal, Barcelona, Spain. ¹⁴IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain. ¹⁵Universitat Pompeu Fabra (UPF), Barcelona, Spain. ¹⁶CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. ¹⁷Clinic and Polyclinic of Dermatology, University Medicine Greifswald, Greifswald, Germany. ¹⁸Department of Dermatology, Venereology and Allergology, University-Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany. ¹⁹Divisions of Pharmacogenomics and Genetics, Genomics and Precision Medicine, Department of Medicine, University of Arizona College of Medicine, Tucson, AZ, USA. ²⁰Barcelona Supercomputing Center (BSC), Joint BSC-CRG-IRB Research Program in Computational Biology, Barcelona, Spain. ²¹APH Amsterdam Public Health, Amsterdam, The Netherlands. ²²Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark. ²³Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ²⁴Department of Medicine, University of California San Francisco, San Francisco, CA, USA. ²⁵Department of Bioengineering & Therapeutic Sciences, University of California

San Francisco, San Francisco, CA, USA. ²⁶Department of Preventive Medicine, University of Southern California, Keck School of Medicine, Los Angeles, CA, USA. ²⁷University of Basel, Basel, Switzerland. ²⁸Swiss Tropical and Public Health Institute, Basel, Switzerland. ²⁹Department of Paediatrics, Imperial College London, London, UK. ³⁰The Generation R Study Group, Erasmus Medical Center, Rotterdam, The Netherlands. ³¹Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. ³²Department of Pediatrics, Division of Respiratory Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. ³³Allergy and Lung Health Unit, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia. ³⁴Laboratory of Animal Genomics, Unit of Medical Genomics, GIGA Institute, University of Liège, Liège, Belgium. ³⁵Department of Pediatrics, Division of Neonatology, Erasmus Medical Center, Rotterdam, The Netherlands. ³⁶LMU Munich, Dr von Hauner Children's Hospital, Munich, and German Center for Lung Research (DZL), Munich, Germany. ³⁷Research Unit of Molecular Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. ³⁸MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. ³⁹Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, USA. ⁴⁰Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Department of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. ⁴¹Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, University of Munich Medical Center, Ludwig-Maximilians-Universität München, Munich, Germany. ⁴²Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna, La Laguna, Tenerife, Spain. ⁴³Telethon Kids Institute (TKI), Perth, Western Australia, Australia. ⁴⁴Department of Pediatrics, Erasmus Medical Center, Rotterdam, The Netherlands. ⁴⁵Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland. ⁴⁶Biocenter Oulu, University of Oulu, Oulu, Finland. ⁴⁷Unit of Primary Care, Oulu University Hospital, Oulu, Finland. ⁴⁸Department of Bio and Health Informatics, Technical University of Denmark, Kongens Lyngby, Denmark. ⁴⁹deCODE genetics/Amgen Inc, Reykjavik, Iceland. ⁵⁰Faculty of Medicine, University of Iceland, Reykjavik, Iceland. ⁵¹Department of Pediatric Pneumology and Allergy, University Children's Hospital Regensburg (KUNO), Regensburg, Germany. ⁵²Department of Public Health, University of Helsinki, Helsinki, Finland. ⁵³Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland. ⁵⁴National Institute for Health and Welfare, Helsinki, Finland. ⁵⁵Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. ⁵⁶Max-Delbrück-Center (MDC) for Molecular Medicine, Berlin, Germany. ⁵⁷Clinic for Pediatric Allergy, Experimental and Clinical Research Center, Charité Universitätsmedizin Berlin, Berlin, Germany. ⁵⁸Department of Public Health Sciences, Henry Ford Health System, Detroit, MI, USA. ⁵⁹Divisions of Genetics, Genomics and Precision Medicine, Department of Medicine, University of Arizona College of Medicine, Tucson, AZ, USA. ⁶⁰Sachs' Children's Hospital, Stockholm, Sweden. ⁶¹Programs in Metabolism and Medical & Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. ⁶²Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA. ⁶³Institute of Clinical Research, University of Southern Denmark, Department of Obstetrics & Gynecology, Odense University Hospital, Odense, Denmark. ⁶⁴School of Medicine and Public Health, Faculty of Medicine and Health, University of Newcastle, Callaghan, New South Wales, Australia. ⁶⁵Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden. ⁶⁶Ludwig-Maximilians-University of Munich, Dr. von Hauner Children's Hospital, Division of Metabolic Diseases and Nutritional Medicine, Munich, Germany. ⁶⁷Institute for Respiratory Health, Harry Perkins Institute of Medical Research, University of Western Australia, Nedlands, Western Australia, Australia. ⁶⁸Ib-Salut, Area de Salut de Menorca, Institut d'Investigació Sanitària Illes Balears (IdISBa), Palma de Mallorca, Spain. ⁶⁹Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. ⁷⁰Channing Division of Network Medicine, Brigham & Women's Hospital and Harvard Medical School, Boston, MA, USA. ⁷¹Department of Internal Medicine, Henry Ford Health System, Detroit, MI, USA. ⁷²QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ⁷³Population Health Research Institute, St George's, University of London, London, UK. ⁷⁴These authors contributed equally: Johannes Waage, Marie Standl. *e-mail: kb@copsac.com

Supplementary Information

Supplementary Note	3
Supplementary Results	3
Description of novel loci	3
Non-allergic rhinitis	4
Association in non-European populations	4
Cohort recruitment details, acknowledgements and funding statements	5
Supplementary References	24
Supplementary Figures	
Supplementary Figure 1: QQ-plot for the discovery phase of the allergic rhinitis GWAS	27
Supplementary Figure 2: Forest plots from the discovery phase – known loci	28
Supplementary Figure 3: Forest plots from the discovery and replication phase – novel loci	50
Supplementary Figure 4: Scatter plot of beta coefficients between 23andme and non-23andme cohorts	77
Supplementary Figure 5: Regional locus plots of allergic rhinitis discovery phase markers	78
Supplementary Figure 6: Enrichment plot (GARFIELD)	130
Supplementary Figure 7: “UpSet” set visualisation plot	131
Supplementary Figure 8: Overlap enriched gene sets allergic rhinitis and allergic sensitization	132
Supplementary Figure 9: Allergic rhinitis genetic risk score and disease prevalence	133
Supplementary Figure 10: Regional locus plot for rs193243426	134
Supplementary Tables	see separate file
Supplementary Table 1: Overview of participating cohorts	ST1
Supplementary Table 2: Extended association results for allergic rhinitis	ST2
Supplementary Table 3: Information for each cohort participating	ST3
Supplementary Table 4: Association results stratified by “doctor diagnosed” or “self-reported” allergic rhinitis	ST4
Supplementary Table 5: Results from MAGENTA analysis	ST5
Supplementary Table 6: List of conditional association results of secondary loci (GCTA)	ST6
Supplementary Table 7: List of locus-to-gene mappings from functional associations	ST7
Supplementary Table 8: List of markers part of the 99% credible set for each locus	ST8
Supplementary Table 9: List of coding and amino acid consequences from the credible sets	ST9
Supplementary Table 10: Association with classical HLA alleles (conditional analyses)	ST10
Supplementary Table 11: Association with classical HLA alleles (unconditional analyses)	ST11

Supplementary Table 12: Association of allergic rhinitis with HLA amino acid variants after iterative conditional analyses	ST12
Supplementary Table 13: Association of allergic rhinitis with HLA amino acid variants (univariable analyses)	ST13
Supplementary Table 14: Functional description of known and novel replicating loci	ST14
Supplementary Table 15: Overlap with reported autoimmune disease variants	ST15
Supplementary Table 16: Results from the GARFIELD regulatory region enrichment analysis	ST16
Supplementary Table 17: Tissue enrichment results (DEPICT)	ST17
Supplementary Table 18: Results from the Ingenuity Pathway enrichment analysis	ST18
Supplementary Table 19: Table of index markers for allergic sensitisation	ST19
Supplementary Table 20: Replication of allergic rhinitis top loci in the allergic sensitization GWAS	ST20
Supplementary Table 21: DEPICT gene set enrichment analysis for allergic rhinitis	ST21
Supplementary Table 22: DEPICT gene set enrichment analysis for allergic sensitization	ST22
Supplementary Table 23: Replication of allergic rhinitis top loci in the non-allergic rhinitis GWAS	ST23
Supplementary Table 24: Results of novel, replicating loci in the random-effects meta-analysis (non-European populations)	ST24
Supplementary Table 25: List of proxies reported for the replication phase	ST25

Supplementary Note

Supplementary Results

Description of novel loci

Three of the 20 loci not previously associated with AR have previously been associated with other atopic traits. These include the locus at 5p13.2 previously associated with eczema¹. This index SNP is in strong LD with a missense variation in *IL7R* and with meQTLs in blood and lung tissue. IL-7R is necessary for V(D)J recombination of T and B cell receptors, for cellular activation by the type 2 immune inducer TSLP, and *IL7R* polymorphisms have been associated with the autoimmune disease multiple sclerosis^{2,3}. A second locus at 12q24.12 has been associated with blood eosinophil count⁴. A possible functional candidate at this locus, supported by a blood cis-eQTL and a correlated missense variation, is *SH2B3* which is involved in inflammatory pathways and T Cell activation⁵. The third locus at 19q13.11 has previously been associated with eosinophilic esophagitis.⁶ We found evidence of enhancer-promotor interaction with *CEBPA* and *CEBPG* in macrophages, monocytes and neutrophils. CEBPA is essential for myeloid cell lineage differentiation⁷ while CEBPG function as an enhancer binding protein in *IL4*, *IL6*, *IL8* and *IGH*, the immunoglobulin heavy chain locus⁸.

The majority of the remaining novel loci imply genes with a known role in the immune system. The locus at 11q23 is located near *CXCR5*, a plausible causal candidate supported by cis eQTL and enhancer-promotor interaction in B-cells. *CXCR5* encodes a chemokine receptor that is present on B cells and in a subset of follicular T cells and is involved in B cell migration and facilitates B and T cell interactions within the lymph node⁹. This locus has previously been associated with autoimmune diseases^{10,11} and lymphoma¹². The lead signal at 1q23.3 is located near *FCER1G* and associated with expression of FCER1G in blood and lung tissue. *FCER1G* encodes the gamma chain of the high affinity IgE receptor, a central component in allergic responses, and is thereby a plausible candidate conferring risk of AR at this locus. The locus at 4q24 harbors cis-eQTLs for *NFKB1* in monocytes and whole blood and meQTLs in lung tissue. *NFKB1* encodes a subunit of the NF-kappaB complex playing an important role for activation of multiple inflammatory pathways.¹³ NF-kappaB activation might also be implicated in the association signal at 10q24.32 with evidence of enhancer-promoter interaction with *NFKB2*, encoding another subunit of the NF-kappaB complex. Other immune related candidates include *BACH2* at 6q15 with a role in antigen-induced formation of memory B cells and memory T cells^{14,15}, *TYRO3* and *LTK* at 15q15.1 modulating TLR signaling¹⁶ and T helper 2 immunity¹⁷, respectively, *VPRBP* at 3p21.2 required for optimal T cell proliferation after antigen recognition and involved in V(D)J recombination during B cell development¹⁸, *SPPL3* and *OASL* at 12q24.31 with potential roles for NK cell maturation¹⁹ and IFN-alpha signaling²⁰ respectively, *RORA* at 15q22.2 involved in type 2 innate lymphoid cell development and allergic inflammation²¹, and *TNFSF11* at 13q14.11 with a role in T cell activation by dendritic cells²².

Finally, a number of the novel AR loci imply genes with no clear function in AR pathogenesis. These include one of the strongest associated loci in this meta-analysis at 12q24.31 with the top-signal located between *CDK2AP1* and *C12orf65*. This locus harbors cis-eQTLs in blood and/or lung tissue for several genes (*ABCB9*, *ARL6IP4*, *C12orf65*, *CDK2AP1*, *MPHOSPH9*, *OGFOD2*, *PITPNM2*, *RILPL2*, *SBNO1*, *SETD8*, and *SNRNP35*) and shows evidence for enhancer-promoter interaction with *DDX55* in various immune cells, but none of these genes have an obvious function related to AR. The same is the case for *NEGR1* at 1p31.1, *JAZF1* at 7p15.1,

FOSL2 at 2p23.2, and *RERE* at 1p36.32, all supported by cis-eQTLs or me-QTLs in blood and/or lung tissue. No cis-eQTL was identified for the locus at 2q36.3 near *DAW1*. Further studies of these genes and loci might help to understand AR pathogenesis and identify novel drug targets.

Non-allergic rhinitis

Non-allergic rhinitis (NAR), defined as rhinitis symptoms without evidence of allergic sensitization, is a common but poorly understood disease entity.²³ We performed the first GWAS on this phenotype hypothesizing that this might reveal specific rhinitis mechanisms. The analysis included 2,028 cases and 9,606 controls from 9 studies but did not identify any risk loci at the genome-wide significance level. Look-up of the 41 AR top-loci in the non-allergic rhinitis GWAS indicated considerable overlap of susceptibility loci with 33 of 41 loci showing consistent direction of effect and 8 AR loci showing nominal significance ($p < 0.05$) in the non-allergic rhinitis GWAS (Supplementary Table 23). In line with this, we found that calculating the population-attributable risk fraction (PARF) for NAR using the risk score from AR showed a highly statistically significant PARF of 48% (95% CI 23%-65%), indicating that a significant proportion of NAR cases can be attributed to AR susceptibility loci. This suggests some genetic commonality between the two phenotypes that can be due to methodological issues, e.g. that some cases of non-allergic rhinitis misdiagnosed due to sensitization not captured by the tests performed, or might have a biological explanation, e.g. the presence of local allergic mechanisms in the nose without presence of systemic allergic sensitization.

Association in non-European populations

We only had data from 3 populations of non-European descent, including populations of African-American (SAPPHIRE, SAGEII) and Latin-American (GALAI) descent. The combined number of individuals was 6,313 providing insufficient statistical power to perform genome-wide association analysis. Look up of the identified novel AR loci was possible for 19 variants. Of these, 13 showed the same direction of effect and two (near *RERE* and *BACH2*) reached nominal statistical significance in the random-effects model (Supplementary Table 24). However, the limited statistical power in the non-European sample set does not allow solid interpretations about the potential overlap of AR loci between ethnicities.

Cohort recruitment details, acknowledgements and funding statements

23andme

Recruitment

All individuals included in the analyses provided informed consent and answered surveys online according to our human subjects protocol, which was reviewed and approved by Ethical & Independent Review Services, a private institutional review board (<http://www.eandireview.com>). Samples were drawn from 23andMe research participants who reported via web-based questionnaires and answered questions about allergic symptoms.

Case/Control definition

23andMe research participants were able to fill out web-based questionnaires whenever they logged into their 23andMe accounts.

For GWAS:

The "allergic rhinitis" phenotype combines the questions from multiple surveys

(returning all consistent responses, inconsistent responses are set to NA)

1. Survey: "Asthma"

Question: "Have you ever had any of the following? Please check all that apply: "Allergic rhinitis (stuffed or dripping nose caused by allergies)"

2. Samples from 'Allergies and Asthma' and "Allergies" surveys that controls are those with no allergies to a broad range of allergens (e.g seasonal_allergies, environmental_allergy, food_allergy and drug_allergy); cases report the allergy to any of these:

['grasses', 'trees', 'weeds', 'cats', 'dogs', 'dust_mite', 'mold'],

and had symptoms 'itchy or runny nose'.

2.1 Survey: "Allergies and Asthma"

Question: "Have you had an allergic reaction to any of the following types of plants?" "Grasses"

2.2 Survey: "Allergies and Asthma"

Question: "What type of reaction did you have after being exposed to grasses? Please check all that apply: "Itchy or runny nose"

2.3 Survey: "Allergies (allergies)"

Question: "Plant Allergies"/"Have you ever had an allergic reaction to any of these types of plants?" "Grasses"

2.4 Survey: "Allergies(allergies)",

Question: "What type of reaction did you have after being exposed to grass? Please check all that apply." : "Itchy or runny nose"

For replication analysis:

The "allergic rhinitis" phenotype combines the questions from multiple surveys

(returning all consistent responses, inconsistent responses are set to NA)

1. Survey: "Asthma"

Question: "Have you ever had any of the following? Please check all that apply: "Allergic rhinitis (stuffed or dripping nose caused by allergies)"

2. New Survey: "Asthma(asthma_background)"

Question: "Have you ever had any of the following? Please check all that apply: "Allergic rhinitis (stuffed or dripping nose caused by allergies)"

3. New Survey: "Asthma(asthma_meds)",

Question: "Have you ever had any of the following? Please check all that apply: "Allergic rhinitis (stuffed or dripping nose caused by allergies)"

4. Samples from allergy surveys that controls are those with no allergies to a broad range of allergens (e.g seasonal_allergies, environmental_allergy, food_allergy and drug_allergy); cases report the allergy to any of these:

['grasses', 'trees', 'weeds', 'cats', 'dogs', 'dust_mite', 'mold'],

and had symptoms 'itchy or runny nose'.

4.1 Survey: "Allergies and Asthma"

Question: "Have you had an allergic reaction to any of the following types of plants?" "Grasses"

4.2 Survey: "Allergies and Asthma"

Question: "What type of reaction did you have after being exposed to grasses? Please check all that apply: "Itchy or runny nose"

4.3 Survey: "Allergies (allergies)"

Question: "Have you ever had an allergic reaction to any of these types of plants?" "Grasses"

4.4 Survey: "Allergies(allergies)",

Question: "What type of reaction did you have after being exposed to grass? Please check all that apply." : "Itchy or runny nose"

Acknowledgements and Funding

We thank the research participants of 23andMe who answered surveys, as well as the employees of 23andMe, who together made this research possible. We particularly thank the following members of the 23andMe Research Team: Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, Nicholas A. Furlotte, Bethann S. Hromatka, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A. M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Vladimir Vacic, and Catherine H. Wilson. This work was supported in part by the National Heart, Lung, and Blood Institute of the US National Institutes of Health under grant 1R43HL115873-01, and the National Human Genome Research Institute of the National Institutes of Health under grant number R44HG006981.

AAGC (QIMR)

Recruitment

As part of the AAGC, we performed a GWAS of allergic rhinitis in unrelated individuals of European ancestry ascertained from the Australian population. Participants provided informed consent to participate in this study, which was approved by the respective ethics committees.

Acknowledgements and Funding

The AAGC was funded by a project grant from the NHMRC (613627).

AAGC members

Graham Jones¹, Patrick Danoy², Svetlana Baltic³, Desiree Mészáros⁴, Catherine Hayden⁵, Sarah E Medland⁶, Andrew J. Kemp⁷, Faang Cheah³, Dale R. Nyholt⁶, Melissa C. Southey⁸, Mary Roberts⁹, Scott D. Gordon⁶, Euan R. Tovey¹, Loren Price³, Margaret J. Wright⁶, James Markos¹⁰, Anjali K. Henders⁶, Graham Giles¹¹, Li P. Chung³, Paul S. Thomas¹², Ian Feather¹³, Pamela A. Madden¹⁴, Suzanna Temple³, Stephen Morrison¹⁵, Chalermchai Mitrpant³, Brad Shelton³, Andrew C. Heath¹⁴, Mark Jenkins², Warwick J. Britton¹⁶, John L. Hopper¹⁷, Stephen R. Leeder¹⁸, Haydn Walters⁴, Michael J. Abramson¹⁹, Matthew A Brown², Shyamali C. Dharmage¹⁷

¹ Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia.

² University of Queensland Diamantina Institute, Princess Alexandra Hospital, Brisbane, Australia.

³ Lung Institute of WA and Centre for Asthma, Allergy and Respiratory Research, University of WA, Perth, Australia.

⁴ Menzies Research Institute, Hobart, Australia.

⁵ School of Paediatrics and Child Health, Princess Margaret Hospital for Children, Perth, Australia.

⁶ The Queensland Institute of Medical Research, Brisbane, Australia.

⁷ The Children's Hospital, Westmead, Sydney, Australia.

⁸ Department of Pathology, The University of Melbourne, Melbourne, Australia.

⁹ Department of Respiratory Medicine, Royal Children's Hospital, Parkville, Australia.

¹⁰ Launceston General Hospital, Launceston, Australia.

¹¹ Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Australia.

¹² Faculty of Medicine, University of New South Wales, Sydney, Australia.

¹³ Gold Coast Hospital, Southport, Australia.

¹⁴ Washington University School of Medicine, St Louis, United States.

¹⁵ University of Queensland, Brisbane, Australia.

¹⁶ Centenary Institute of Cancer Medicine & Cell Biology, Royal Prince Alfred Hospital, Camperdown, Australia.

¹⁷ Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Australia.

¹⁸ Australian Health Policy Institute, University of Sydney, Sydney, Australia

ALSPAC

Recruitment

ALSPAC recruited 15,247 pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992, resulting in 14,775 live births and 14,701 children who were alive at 1 year of age. Enrolment is described in more detail in the cohort profile paper²⁴ and via the website <http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/>. Biological samples including DNA have been collected for 10,121 of the children from this cohort. Ethical approval for the study was obtained

from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees and written and informed consent was provided by the parents.

Acknowledgements and Funding

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and the Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. GWAS data was generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. FLG genotyping was funded by an MRC centenary award (awarded to L Paternoster).

B58C (British 1958 birth cohort)

Recruitment

The British 1958 birth cohort is an ongoing follow-up of all persons born in England, Scotland and Wales during one week in 1958. At ages 7, 11, 16, 23, 33 and 42 years, a history of hay fever or allergic rhinitis was obtained by interview with a parent (aged 7-16) or the cohort member (ages 23-42). At the age of 44-45 years, the cohort were followed up with a biomedical examination and blood sampling, from which a DNA collection was established as a nationally representative reference panel, and the presence of circulating IgE antibodies to house dust mite, mixed grass pollen and cat fur was ascertained by Hytec enzyme immunoassay, with a detection threshold of 0.35 kU/L.

Acknowledgements and Funding

We acknowledge use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). The B58C-GABRIEL genotyping was supported by a contract from the European Commission Framework Programme 6 (018996) and grants from the French Ministry of Research.

The funding agencies did not have any influence on design and conduct of the study; collection, management, and interpretation of the data; or preparation, review, or approval of the manuscript

BAMSE

Recruitment

BAMSE is a Swedish birth cohort study. A total number of 4,089 newborn infants were recruited between 1994 and 1996 in the Stockholm area¹⁹. The first questionnaire data, dealing with parental allergic diseases, socio-economic status and residential characteristics, was obtained when the children were about 2 months. Similar questionnaires, with a focus on the children's symptoms related to asthma and allergic diseases

including eczema, were answered by the parents when the children were approximately 1, 2, 4, 8 and 12 years old. At 8 years of age, all children were invited to clinical testing, and blood samples were obtained from 2,480 children (~60%). DNA was extracted from 2,033 samples after exclusion of samples with too little blood, lack of questionnaire data, or if parental consent to genetic analysis of the sample was not obtained.

Acknowledgements and Funding

BAMSE was supported by The Swedish Research Council, The Swedish Heart-Lung Foundation, Stockholm County Council (ALF) and the Strategic Research Programme (SFO) in Epidemiology at Karolinska Institutet. Genome-wide genotyping in BAMSE was funded by the European Commission as part of GABRIEL (A multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community) contract number 018996 under the Integrated Program LSH-2004-1.2.5-1.

CAMP

Recruitment

The Childhood Asthma Management Program (CAMP) population is composed of non-Hispanic white subjects from a multicenter clinical trial that followed 1,041 children with asthma for four years and 84% of the original participants for 12 years. Stringent inclusion criteria ensured that participants had mild to moderate asthma, which was assessed as having asthma symptoms at least twice per week, using asthma medication daily, or using an inhaled bronchodilator twice per week for six or more months of the year prior to recruitment. CAMP subjects had increased airway responsiveness, as established by a bronchoprovocation test of up to 12.5mg/dl of methacholine resulting in 20% or greater forced expiratory volume in one second (FEV1) reduction.

Acknowledgements and Funding

We thank all CAMP subjects for their ongoing participation in this study. We acknowledge the CAMP investigators and research team, supported by NHLBI, for collection of CAMP Genetic Ancillary Study data. All work on data collected from the CAMP Genetic Ancillary Study was conducted at the Channing Laboratory of the Brigham and Women's Hospital under appropriate CAMP policies and human subject's protections. The CAMP Genetics Ancillary Study is supported by U01 HL075419, U01 HL65899, P01 HL083069 R01 HL086601, from the National Heart, Lung and Blood Institute, National Institutes of Health, and Supinda Bunyavanich is funded by the US National Institutes of Health (NIH) R01 AI118833 and NIH K08 AI093538. In addition Supinda Bunyavanich is funded by US National Institutes of Health (NIH) R01 AI118833 and NIH K08 AI093538.

CEDAR

Recruitment

The CEDAR cohort comprises 323 healthy individuals of European descent visiting the University of Liège Hospital as part of a national screening for colon cancer. They provided blood samples and intestinal biopsies under full informed consent. The experimental protocol was approved by the ethics committee of the University of Liège Academic Hospital. Informed consent was obtained prior to donation in agreement with the recommendations of the declaration of Helsinki for experiments involving human subjects. We refer to this cohort as CEDAR for Correlated Expression and Disease Association Research.

Acknowledgements and funding

Assembling the CEDAR cohort was supported by grants from WELBIO (CAUSIBD), BELSPO (BeMGI), and Horizon 2020 (SYSCID). Computational resources at ULg have been provided by GIGA and the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique

(F.R.S.-FNRS) under Grant No. 2.5020.11. We thank Benoit Charlotiaux, Wouter Coppieters, François Crins, Valérie Deffontaine, Elisa Docampo, Mahmoud Elansary, Anne-Stefan Gori, Rob Mariman, Myriam Mni, Yukihide Momozawa, Emilie Théâtre and the members of the GIGA genomics platform for their contributions.

COPSAC2000

Recruitment

The COPSAC2000 birth cohort study is a prospective clinical study of a birth cohort of 411 infants born to mothers with a history of asthma. The newborns were enrolled at the age of 1 month, the recruitment of which was previously described in detail.²⁵ The study was approved by the Ethics Committee for Copenhagen (KF 01- 289/96) and The Danish Data Protection Agency (2008-41-1754) and informed consent was obtained from both parents. For this study, only the parents were included.

Acknowledgements and Funding

COPSAC is funded by private and public research funds all listed on www.copsac.com. The Lundbeck Foundation (Grant no R16-A1694); The Danish Ministry of Health (Grant no 903516); Danish Council for Strategic Research (Grant no 0603-00280B); The Danish Council for Independent Research (Grant no 10-082884 and 271-08-0815); The Capital Region Research Foundation (No grant number) and NIH-NHLBI R01 HL129735 have provided core support for COPSAC.

deCODE

Recruitment

Icelandic patients with allergic rhinitis, based on physicians diagnosis, who attended an asthma clinic or emergency room at the National University Hospital of Iceland or the Icelandic Medical Center (Laeknasetrid) during the years 1977 to 2016 were recruited.^{26,27} All participating subjects who donated biospecimens provided informed consent. Personal identities of the participants and biological samples were encrypted by a third-party system approved and monitored by the Icelandic Data Protection Authority. The study, Genetics of asthma and allergies, was approved by the National Bioethics Committee (VSN-14-099, VS00-135) and the Data Protection Authority (no. PV_2014060841/PS) in Iceland.

Acknowledgements and Funding

We thank all the participants who provided data and samples for this study. We acknowledge the clinical investigators and staff of the Patient Recruitment Center for collection clinical information and biological samples. We thank the scientists at deCODE genetics and the staff deCODE's core facilities for genotyping, processing and analysis of genome sequences data.²⁸ The study is funded by deCODE genetics.

ECRHS

Study description

The European Community Respiratory Health Survey (ECRHS) is an international cohort study designed to assess the prevalence of asthma and allergic disease and identify their risk factors (Burney et al. The European Community Respiratory Health Survey.²⁹ Young adults of European descent were randomly recruited from community-based sampling frames in the ECRHS I (1991-1993) and followed up twice in the 20 years after the first assessment (ECRHS II: 1998-2002; ECRHS III: 2008-2013). Genotyping was performed on blood samples collected at ECRHS II. The current analysis was restricted to participants from centres that took part in the GABRIEL genotyping initiative and were representative of the random sample taken at baseline (i.e. this sample is not enriched with asthmatics). Ethical approval from local research ethics committees and written consent from participants were obtained.

Acknowledgements and Funding

The authors would like to thank the participants, field workers and researchers who have participated in the ECRHS study for their time and cooperation. This work was supported by a contract from the European Commission (018996), Fondo de Investigación Sanitaria (91/0016-060-05/E, 92/0319, 93/0393, 97/0035-01, 99/0034-01 and 99/0034-02), Hospital General de Albacete, Hospital General Ramón Jiménez, Consejería de Sanidad del Principado de Asturias, CIRIT (1997SGR 00079, 1999SGR 00241), and Servicio Andaluz de Salud, SEPAR, Public Health Service (R01 HL62633-01), RCESP (C03/09), Red RESPIRA (C03/011), Basque Health Department, Swiss National Science Foundation, Swiss Federal Office for Education and Science, Swiss National Accident Insurance Fund (SUVA), GSF-National Research Centre for Environment and Health, Deutsche Forschungsgemeinschaft (DFG) (FR 1526/1-1, MA 711/4-1), Programme Hospitalier de Recherche Clinique-DRC de Grenoble 2000 no. 2610, Ministry of Health, Direction de la Recherche Clinique, Ministère de l'Emploi et de la Solidarité, Direction Générale de la Santé, CHU de Grenoble, Comité des Maladies Respiratoires de l'Isère. UCB-Pharma (France), Aventis (France), Glaxo France. Estonian Science Foundation, and Asthma UK (formerly known as National Asthma Campaign UK).

EVE (CAG/CSGA)

Recruitment

Subjects were from the Chicago Asthma Genetics Study (CAG) and the NHLBI Collaborative Studies on the Genetics of Asthma (CSGA).

CAG included subjects recruited from a) families ascertained through asthma affected sib pairs, b) asthma affected children, c) adults and children with severe persistent asthma, and d) non-asthma controls over 18 years of age. Asthma was defined as 1) a physical diagnosis, 2) the presence of at least two self-reported symptoms, 3) current use of medications, and 4) bronchial hyperresponsiveness to methacholine or reversibility to inhaled bronchodilator. Controls had no self-reported personal or family history (1st degree relatives) of asthma.

CSGA subjects were recruited from the NHLBI Collaborative Studies on the Genetics of Asthma (CSGA). CSGA cases were recruited and diagnosed as described above for CAG.

Acknowledgements and Funding

EVE was funded by an NIH grant RC2 HL101651.

EVE (CHS)

Recruitment

Two cohorts in the Children's Health Study (CHS) including a total of 3000 children (1512 from cohort 1, and 1488 from cohort 2) who were either Hispanic (n=1398) or non-Hispanic white (n=1602) were included in the analysis. The first cohort included children between the ages of 8-14 years recruited from schools in 12 southern California communities in 1993 and 1996. The second cohort included children between the ages of 5-8 years recruited from schools in 13 communities in 2002. A total of 1249 cases were classified as having asthma if the adult completing the questionnaire reported that a doctor had "ever diagnosed the child as having asthma". A total of 1751 controls were selected using frequency matching with cases on sex, ethnicity, cohort, and follow-up time from children without the asthma outcome.

Acknowledgements and Funding

The CHS was supported by the Southern California Environmental Health Sciences Center (grant P30ES007048); National Institute of Environmental Health Sciences (grants 5P01ES011627, ES021801,

ES023262); P01ES009581, P01ES011627, P01ES022845, R01 ES016535, R03ES014046, P50 CA180905, R01HL061768, R01HL076647, R01HL087680, and RC2HL101651), the Environmental Protection Agency (grant #s RD83544101, R826708, RD831861, and R831845), and the Hastings Foundation.

FINNTWIN (FTC)

Recruitment

The total of 2373 subjects originate within the Finnish Twin Cohort (FTC).^{30,31} Altogether 1161 of the subjects belong to the Older Finnish Twin Cohort consisting of adult twins born in 1938-1957.³² Altogether 687 of the subjects belong to the FinnTwin12 cohort, a population-based longitudinal study of five consecutive birth cohorts (1983-1987) of Finnish twins. Altogether 525 of the subjects belong to the FinnTwin16 cohort, a population-based longitudinal study of five consecutive birth cohorts (1975-1979) of Finnish twins and their families.

Acknowledgements and Funding

We warmly thank the participating twin pairs and their family members for their contribution. We would like to express our appreciation to the skilled study interviewers A-M Iivonen, K Karhu, H-M Kuha, U Kulmala-Gråhn, M Mantere, K Saanakorpi, M Saarinen, R Sipilä, L Viljanen and E Voipio. Anja Häppölä and Kauko Heikkilä are acknowledged for their valuable contribution in recruitment, data collection, and data management.

Phenotyping and genotyping of the Finnish twin cohorts has been supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506, 129680), the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 263278 and 264146 to J. Kaprio), National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145, and AA-09203 to R. J. Rose and AA15416 and K02AA018755 to D. M. Dick), Sigrid Juselius Foundation (to J. Kaprio), Global Research Award for Nicotine Dependence, Pfizer Inc. (to J. Kaprio), and the Wellcome Trust Sanger Institute, UK. Antti-Pekka Sarin and Samuli Ripatti are acknowledged for genotype data quality controls and imputation.

The funding agencies did not have any influence on design and conduct of the study; collection, management, and interpretation of the data; or preparation, review, or approval of the manuscript.

GABRIELA

Recruitment

The GABRIELA study (GABRIEL Advanced Studies) was a cross-sectional population-based study in 5 rural regions of Europe: Baden-Württemberg and Bavaria in Germany, North/Central Switzerland, Tyrol in Austria, and Lower Silesia in Poland. In phase 1, a recruitment questionnaire was sent out to 132,518 children age 5 to 13 years through their primary schools. Completed questionnaires were returned by 79,888 (60.3%). With a documented parental consent to dust sampling, blood sampling, and genetic analyses 34,491 children were eligible for phase 2. A stratified random selection process was applied to enrich informative observations. In a first step, 9668 children were selected for phase 2 within 3 exposure strata per center, i.e. farm children, non-farm children with and without farm exposure. In the second step, 1708 children from the Austrian, German, and Swiss centers were selected for genotyping within the mentioned exposure strata and simultaneously stratified for outcome, i.e. asthma, atopic sensitization, and unaffected controls.³³

Acknowledgements and Funding

The GABRIELA study was supported by the European Commission as part of GABRIEL (a multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community) contract

number 018996 under the Integrated Program LSH-2004-1.2.5-1. Michael Kabesch was supported by the German Ministry for Education and Science (BMBF) with the fund SYSINFAME, 01ZX1306E.

The GABRIELA study group

Silvia Apprich, PhD, Andrzej Boznanski, MD, PhD, Charlotte Braun-Fahrlander, MD,

Gisela Büchele, MPH, William Cookson, MD, DPhil, Paul Cullinan, MD, Hanna Danielewicz, MD, Anna Debinska, Martin Depner, PhD, Markus J. Ege, MD, Urs Frey, MD, PhD, Oliver Fuchs, MD, Jon Genuneit, Dick Heederik, PhD, Elisabeth Horak, MD, Anne Hyvärinen, PhD, Sabina Illi, PhD, Michael Kabesch, MD, Katalin Kovacs, Aleksandra Kosmeda, PhD, Wolfgang Kneifel, PhD, Philipp Latzin, MD, PhD, Roger Lauener, MD, Georg Loss, Stephanie MacNeill, MSc, Bernhard Morass, MD, Anne-Cécile Normand, Ilka Noss, PhD, Renaud Piarroux, MD, PhD, Helena Rintala, PhD, Mascha K. Rochat, MD, Nikolaos Sitaridis, Barbara Sozanska, MD, David Strachan, MD, Christine Strunz-Lehner, MPH, Bertrand Sudre, MD, PhD, Erika von Mutius, MD, MSc, Marco Waser, PhD, Juliane Weber, MD, and Inge Wouters, PhD.

From Innsbruck and Vienna, Austria; Munich, Regensburg, and Ulm, Germany; Kuopio, Finland; Besancon and Lausanne, France, Utrecht, the Netherlands; Wroclaw, Poland; Basel, Bern, Davos, and St. Gallen, Switzerland; London, UK.

GALA II

Recruitment

The Genes-environments & Admixture in Latino Americans (GALA II) study is an ongoing case-control study of asthma in Latinos that are recruited in five centers throughout the U.S. (Chicago, Illinois; Bronx, New York; Houston, Texas; San Francisco Bay Area, California; and San Juan, Puerto Rico).³⁴ Participants were eligible if they were aged between 8 to 21 years, had <10 pack-years of smoking history and self-identified all four grandparents as Latino. Subjects were excluded if they reported any of the following: (1) 10 or more pack-years of smoking; (2) any smoking within one year of recruitment date; (3) pregnancy in the third trimester; or (4) history of one of the following conditions: sickle cell disease, cystic fibrosis, sarcoidosis, cerebral palsy, or history of heart or chest surgery. Both children and young adults with physician-diagnosed asthma and non-allergic, non-asthmatic controls were recruited. The presence of allergic rhinitis was defined as a positive response to the question “Has a doctor ever diagnosed the child with hay fever or allergic rhinitis?” According to this, a total of 755 individuals were classified as cases and 2,929 as controls.

All local institutional review boards approved the study and all subjects/parents provided written assent/consent, respectively.

Acknowledgements and Funding

The authors acknowledge the families and patients for their participation and thank the numerous health care providers and community clinics for their support and participation in GALA II. In particular, the authors thank the study coordinator Sandra Salazar and the recruiters who obtained the data: Duanny Alva, MD, Gaby Ayala-Rodriguez, Ulysses Burley, Lisa Caine, Elizabeth Castellanos, Jaime Colon, Denise DeJesus, Iliana Flexas, Blanca Lopez, Brenda Lopez, MD, Louis Martos, Vivian Medina, Juana Olivo, Mario Peralta, Esther Pomares, MD, Jihan Quraishi, Johanna Rodriguez, Shahdad Saeedi, Dean Soto, Ana Taveras, Emmanuel Viera. Some computations were performed using the UCSF Biostatistics High Performance Computing System. We also acknowledge the principal investigators of the centers where the GALA II and SAGE II participants were recruited: Kelley Meade, Harold J. Farber, Pedro C. Avila, Denise Serebrisky, Shannon M. Thyne, Emerita

Brigino-Buenaventura, William Rodriguez-Cintron, Saunak Sen, Rajesh Kumar, Michael Lenoir, Luisa N. Borrell, and Jose R. Rodriguez-Santana.

The GALA II study was funded by the Sandler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II, National Institutes of Health 1R01HL117004, R01HL128439, R01HL135156, 1X01HL134589, National Institute of Health and Environmental Health Sciences R01ES015794, R21ES24844, the National Institute on Minority Health and Health Disparities 1P60MD006902, RL5GM118984, 1R01MD010443, and the Tobacco-Related Disease Research Program under Award Number 24RT-0025 to E.G.B. MP-Y was supported by the Ramón y Cajal Program (RYC-2015-17205) by the Spanish Ministry of Economy, Industry and Competitiveness; an award (AC15/00015) by the ISCIII through Strategic Action for Health Research (AES) and European Community (EC) within the Active and Assisted Living (AAL) Programme framework (MP-Y); and the SysPharmPedia grant from the ERACoSysMed 1st Joint Transnational Call from the European Union under the Horizon 2020.

The content of this publication is solely the responsibility of the authors and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government and the funding agencies.

GENEVA-KORA

Recruitment

A total of 1895 individuals were obtained from an atopic dermatitis case collection recruited from tertiary dermatology clinics based at three centres (Technische Universität Munich, as part of the GENEVA study, University of Kiel, University of Bonn) as well as from two German population-representative cohorts, the PopGen biorepository³⁵ and the KORA study.³⁶

Acknowledgements and Funding

The project received infrastructure support through the DFG Clusters of Excellence “Inflammation at Interfaces” (grants EXC306 and EXC306/2), and was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (sysINFLAME, grant # 01ZX1306A), and the PopGen 2.0 network (01EY1103). The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

GENR

Recruitment

The Generation R Study is a population-based prospective cohort study of pregnant women and their children from fetal life onwards in Rotterdam, The Netherlands. All children were born between April 2002 and January 2006, and currently followed until young adulthood. Of all eligible children in the study area, 61% were participating in the study at birth. Cord blood samples including DNA have been collected at birth. The study protocol was approved by the Medical Ethical Committee of the Erasmus Medical Centre, Rotterdam (MEC 217.595/2002/20). Written informed consent was obtained from parents of all participants.

Acknowledgements and Funding

The Generation R Study gratefully acknowledges the contributions of the children and their parents, the general practitioners, the hospitals and the midwives and pharmacies in Rotterdam. We would like to thank A. Abuseiris, K. Estrada, Dr. T. A. Knoch, and R. de Graaf as well as their institutions Biophysical Genomics, Erasmus MC Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technology under grants #01 AK 803 A-H and # 01 IG 07015 G, for access to their grid resources. We thank M. Jhamai, M. Ganesh, P. Arp, M. Verkerk, L. Herrera and M. Peters for their help in creating, managing and performing quality control for the genetic database. Also, we thank K. Estrada and C. Medina-Gomez for their support in the creation and analysis of imputed data. The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the School of Law and the Faculty of Social Sciences of Erasmus University Rotterdam, the Municipal Health Service, Rotterdam area, the Rotterdam Homecare Foundation and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC; Rotterdam). The generation and management of genotype data for the Generation R Study were performed at the Genetic Laboratory of the Department of Internal Medicine at Erasmus Medical Center. The Generation R Study is made possible by financial support from the Erasmus Medical Center (Rotterdam), the Erasmus University Rotterdam and the Netherlands Organization for Health Research and Development (ZonMw; 21000074). The project received funding from the European Union's Horizon 2020 research and innovation programme (LIFECYCLE project, grant agreement no 733206; 2016). Dr Vincent Jaddoe received an additional grant from the Netherlands Organization for Health Research and Development (VIDI 016.136.361) and the European Research Council (Consolidator Grant, ERC-2014-CoG-648916). Dr Liesbeth Duijts received an additional grant from the Lung Foundation Netherlands (no 3.2.12.089; 2012) and from the European Union's Horizon 2020 co-funded programme ERA-Net on Biomarkers for Nutrition and Health (ERA HDHL) (ALPHABET project (no 696295; 2017), ZonMW The Netherlands (no 529051014; 2017)). The researchers are independent from the funders. The study sponsors had no role in the study design, data analysis, interpretation of data, or writing of this report.

GENUFAD

Recruitment

All patients were recruited at Charité Universitätsmedizin Berlin for the GENUFAD study (GEnetic analysis of NUclear Families with Atopic Dermatitis) which has been described in a previous GWAS.³⁷ Briefly, 270 complete nuclear German families were recruited through two siblings affected by eczema with an age of onset below two years. For all family members, eczema, allergic rhinitis, and asthma status was determined by a physician at recruitment. To ensure independent, unrelated cases, all parents with allergic rhinitis were included in this study. If the parents were unaffected, only one child with allergic rhinitis per family was included. Allergic sensitization status was determined for all cases.

German control individuals originated from the Study of Health in Pomerania (SHIP), a population-based cohort from North-Eastern Germany.³⁸

Acknowledgements and Funding

We thank all individuals and families for their participation in this study. We thank all physicians and nurses involved in patient recruitment for their valuable contribution to the study. We are grateful to the laboratory technicians C. Flachmeier and T. Thuss for their work. The study was funded by the German Ministry of Education and Research (BMBF) through the Clinical Research Group for Allergy at Charité Berlin, the National Genome Research Network (NGFN).

SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED) funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012).

GINI/LISA

Recruitment

The influence of Life-style factors on the development of the Immune System and Allergies in East and West Germany PLUS the influence of traffic emissions and genetics (LISApplus) Study is a population based birth cohort study. A total of 3094 healthy, full-term neonates were recruited between 1997 and 1999 in Munich, Leipzig, Wesel and Bad Honnef. The participants were not preselected based on family history of allergic diseases.³⁹ A total of 5991 mothers and their newborns were recruited into the German Infant study on the influence of Nutrition Intervention PLUS environmental and genetic influences on allergy development (GINIplus) between September 1995 and June 1998 in Munich and Wesel. Infants with at least one allergic parent and/or sibling were allocated to the interventional study arm investigating the effect of different hydrolysed formulas for allergy prevention in the first year of life.⁴⁰ All children without a family history of allergic diseases and children whose parents did not give consent for the intervention were allocated to the non-interventional arm. Detailed descriptions of the LISApplus and GINIplus studies have been published elsewhere^{1,2}. DNA was collected at the age 6 and 10 years and 1511 children from the Munich study center from both studies were genotyped. For both studies, approval by the local Ethics Committees and written consent from participant's families were obtained.

Acknowledgements and Funding

The authors thank Martina Müller-Nurasyid for her help with imputation and analysis, all families for participation in the studies and the LISApplus and GINIplus study teams for their excellent work. The GINIplus study team wishes to acknowledge the following: Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg (Heinrich J, Brüske I, Schulz H, Flexeder C, Zeller C, Standl M, Schnappinger M, Ferland M, Thiering E, Tiesler C); Department of Pediatrics, Marien-Hospital, Wesel (Berdel D, von Berg A); Ludwig-Maximilians-University of Munich, Dr von Hauner Children's Hospital (Koletzko S); Child and Adolescent Medicine, University Hospital rechts der Isar of the Technical University Munich (Bauer CP, Hoffmann U); IUF- Environmental Health Research Institute, Düsseldorf (Schikowski T, Link E, Klümper C). The LISApplus study team wishes to acknowledge the following: Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Epidemiology I, Munich (Heinrich J, Schnappinger M, Brüske I, Ferland M, Lohr W, Schulz H, Zeller C, Standl M); Department of Pediatrics, Municipal Hospital "St. Georg", Leipzig (Borte M, Gnodtke E); Marien Hospital Wesel, Department of Pediatrics, Wesel (von Berg A, Berdel D, Stiers G, Maas B); Pediatric Practice, Bad Honnef (Schaaf B); Helmholtz Centre of Environmental Research – UFZ, Department of Environmental Immunology/Core Facility Studies, Leipzig (Lehmann I, Bauer M, Röder S, Schilde M, Nowak M, Herberth G, Müller J, Hain A); Technical University Munich, Department of Pediatrics, Munich (Hoffmann U, Paschke M, Marra S); Clinical Research Group Molecular Dermatology, Department of Dermatology and Allergy, Technische Universität München (TUM), Munich (Ollert M).

GOA

Recruitment

Subjects were drawn from the Genetics of Asthma study (GOA) in the Spanish population that have been previously described.⁴¹⁴² Cases with rhinitis were diagnosed by a physician, aged ≥ 5 years, taking medications to control their asthma symptoms, and had an extensive clinical characterization, including allergic sensitization by skin prick tests and specific IgE measurements. Control subjects self-reported absence of personal or familial history of chronic diseases, and were obtained from the National Genotyping Center (CEGEN-PRB2-ISCI), Universidad de Santiago de Compostela Node.

Acknowledgements and Funding

This work was supported by Instituto de Salud Carlos III (FIS PI11/00623 and FI12/00493) and co-financed by the European Regional Development Funds, “A way of making Europe” from the European Union, and by a grant from the 7th Framework Programme (FP7-REGPOT-2012-CT2012-31637-IMBRAIN).

GOYA

Recruitment

The GOYA study (Genetics of Overweight Young Adults) is derived from the genome wide association study (GWAS) on obesity with a case-cohort (obese-non obese) design comprising young men and women of Danish origin.⁴³

The GOYA (Male) cohort is a longitudinal case-cohort (obese, non-obese) study comprising a randomly (1%) selected control group and all extremely overweight men identified among 362,200 Caucasian men examined at the mean age of 20 years at the draft boards in Copenhagen and its surrounding areas during 1943–1977. Obesity was defined as 35% overweight relative to a local standard in use at the time (mid 1970's), corresponding to a BMI ≥ 31.0 kg/m², which proved to be above the 99th percentile. All of the obese and 50% of the random sampled controls, who were still living in the region, were invited to a follow-up survey in 1992–94 at the mean age of 46 years, at which time the blood samples were taken and genotyping were performed for a total of 673 extremely overweight and 792 controls. With a sampling fraction of 0.5% (50% of 1%), the controls represent about 158,000 men among whom the case group was the most obese.

The GOYA (Female) cohort is nested within the Danish National Birth Cohort (DNBC⁴³) which is a collection of data on 92,274 pregnant women recruited between 1996 and 2002, from their first antenatal visit to their general practitioner. Women participated in four telephone interviews (16 and 30 weeks gestation and 6 and 18 months after birth). They also provided two blood samples during pregnancy. The GOYA females were drawn from the 67,863 women within the DNBC who provided information about prepregnancy BMI, gave birth to a live born infant and provided a blood sample during pregnancy. A case sample of the 3.6% most obese women (n=2451) was defined as those with the largest residuals from the regression of BMI on age and parity (all entered as continuous variables). The BMI for these women ranged from 32.6 to 64.4. From the remaining cohort we selected a random sample of similar size (n=2450). Of these, 1,960 extremely overweight and 1,948 control women were genotyped and passed quality control. Information on mother's year of birth and age at first birth derived from national registers. The study was approved by the regional scientific ethics committee and by the Danish Data Protection Board.

Acknowledgements and Funding

This GOYA male study was conducted as part of the activities of the Gene-diet Interactions in Obesity project (GENDINO, www.gendino.dk) and the MRC Centre for Causal Analyses in Translational Epidemiology (MRC CAITE). TSA received his Post-Doctoral Research Grant from the GENDINO project and acknowledges the same.

The GOYA female part of the study was sampled as a case-cohort study within the Danish National Birth Cohort. The establishment of this cohort was based on a major grant from the Danish National Research Foundation. Additional support for the Danish National Birth Cohort was obtained from the Pharmacy Foundation, the Egmont Foundation, the March of Dimes Birth Defects Foundation and the Augustinus Foundation. Genotyping for the GOYA female study was funded by the Wellcome Trust (Grant ref: 084762MA).

We thank all the participants of the GOYA study.

INMA

Recruitment

Population-based birth cohorts were established as part of the INMA – INfancia y Medio Ambiente [Environment and Childhood] Project in several regions of Spain following a common protocol. This analysis uses the INMA subcohort of Menorca established between 1997 and 1998 in Menorca island. This project aims to study the associations between pre- and postnatal environmental exposures and growth, health, and development from early fetal life until adolescence and has been described previously in detail.⁴⁴ Pregnant women were enrolled during the 1st trimester of pregnancy at public primary health care centers or public hospitals and children were followed after birth until 18y of age. Informed consent was obtained from all participants and the study was approved by the Hospital Ethics Committee.

Acknowledgements and Funding

This study was funded by grants from Instituto de Salud Carlos III (97/0588, 00/0021-02, G03/176, 04/1705, 06/1756, 09/1958 and PS09/00432), Fundació La Marató de TV3, Obra Social la Caixa (IV Convocatoria de ayudas a la Investigación en Enfermedades Neurodegenerativas) and European Community (Contract No. QLK4-CT-2000-00263). DNA extractions and genotyping was performed at the Spanish National Genotyping Centre (CEGEN-Barcelona). The authors would particularly like to thank all the participants for their generous collaboration. The authors are grateful to Mireia Garcia, Maria Victoria Estraña, Maria Victoria Iturriaga, Cristina Capo and Josep LLuch for their assistance in contacting the families and administering the questionnaires. A full roster of the INMA Project Investigators can be found at http://www.proyectoinma.org/presentacion-inma/listado-investigadores/en_listado-investigadores.html.

MAAS

Recruitment

The Manchester Asthma and Allergy Study is an unselected (i.e. population-based), birth cohort study. The setting is the maternity catchment area of Wythenshawe and Stepping Hill Hospitals, comprising of 50 square miles of South Manchester and Cheshire, UK, a stable mixed urban-rural population. Study was approved by the Local Research Ethics Committee. Informed consent was obtained from all parents.

Screening & Recruitment

All pregnant women were screened for eligibility at antenatal visits (8th-10th week of pregnancy). The study was explained to the parents, and informed consent for initial questionnaires and skin prick testing was obtained. Both parents completed a questionnaire about their and their partner's history of asthma and allergic diseases and smoking habits. If the pregnant woman's partner was not present at the antenatal clinic visit, an invitation was sent for him to attend an open-access evening clinic for skin prick testing and questionnaire. Once both parents had completed questionnaires and skin prick testing, a full explanation of the proposed future follow-up for the child was given. Of the 1499 couples who met the inclusion criteria

(<10 weeks of pregnancy, maternal age >18 years, questionnaire and skin test data available for both parents), 288 declined to take part in the study. A total of 1185 participants had at least some evaluable data.

Follow-up

The children have been followed prospectively, and attended review clinics at ages 1, 3, 5, 8 and 11 years.

Acknowledgements and Funding

We would like to thank the children and their parents for their continued support and enthusiasm. We greatly appreciate the commitment they have given to the project. We would also like to acknowledge the hard work and dedication of the study team (post-doctoral scientists, research fellows, nurses, physiologists, technicians and clerical staff). This report includes independent research supported by National Institute for Health Research Respiratory Clinical Research Facility at Manchester University NHS Foundation Trust (Wythenshawe). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. MAAS was supported by the Asthma UK Grants No 301 (1995-1998), No 362 (1998-2001), No 01/012 (2001-2004), No 04/014 (2004-2007), the BMA James Trust and Medical Research Council, UK (G0601361) and The Moulton Charitable Foundation (2004-current); the Medical Research Council (MRC) Grants G0601361, MR/K002449/1 and MR/L012693/1, and Angela Simpson is supported by the NIHR Manchester Biomedical Research Centre. The authors would like to acknowledge the North West Lung Centre Charity for supporting this project.

NFBC1966

Recruitment

The Northern Finland Birth Cohort 1966 is a prospective follow-up study of children from the two northernmost provinces of Finland. Women with expected delivery dates in 1966 were recruited through maternity health centers. Cohort members living in northern Finland or in the capital area were invited to a clinical examination as well as questionnaire at age 31 years. DNA was extracted from blood samples given at the clinical examination. Informed consent for the use of the data including DNA was obtained from all subjects. The study was approved by the ethics committees in Oulu (Finland) and Oxford (UK) universities in accordance with the Declaration of Helsinki.

Acknowledgements and Funding

We thank the late Professor Paula Rantakallio (launch of NFBC1966), and Ms Outi Tornwall and Ms Minttu Jussila (DNA biobanking). The authors would like to acknowledge the contribution of the late Academician of Science Leena Peltonen.

The NFBC resource has been supported by grants from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, 24300796, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), NHLBI grant 5R01HL087679-02 (1RL1MH083268-01), NIH/NIMH (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, EU FP7 EurHEALTHAgeing -277849, the Medical Research Council, UK (G0500539, G0600705, G1002319, PrevMetSyn/SALVE) and the MRC, Centenary Early Career Award. H2020 DynaHEALTH (European Union's Horizon 2020 research and innovation programme under grant agreement No 633595); Exposomic, Genomic and Epigenomic Approach to Prediction of Metabolic and Cardiorespiratory function and Ill-Health (EGEA), Academy of Finland, Grant No 285547; ALEC Study (funded by the European Union's Horizon 2020 Research and Innovation programme under grant agreement No. 633212); H2020 / Marie Skłodowska-Curie Actions, CAPICE (Marie Curie Grant agreement Number 721567); National Public Health Institute, Biomedicum Helsinki, Finland.

NTR

Recruitment

In the Young Netherlands Twin Register (YNTR), twins are followed from birth onwards. Around every two to three years, a survey is sent out inquiring about physical and mental health. At ages 1, 2, 3, 5, 7, 10 and 12 years, the surveys are completed by the parents and/or teachers, from age 14 onward, twins and their non-twin siblings are asked to complete the surveys by themselves. At age 5 years of the twins, parents are asked a series of questions about allergies and rhinitis.⁴⁵ Adult and adolescent twins (ANTR) have been recruited through city councils and media, and complete a self-rating survey every few years. Surveys contain a series of questions about doctor diagnosed diseases and disorders. Subsamples of twin families have taken part in biobank and genotyping studies.

Acknowledgements and Funding

The Netherlands Organization for Scientific Research (NWO 480-15-001/674): Netherlands Twin Registry Repository: researching the interplay between genome and environment; Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007); Amsterdam Public Health (APH); KNAW Academy Professor Award (PAH/6635); the Avera Institute, Sioux Falls, South Dakota (USA) and Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995).

RAINE

Recruitment

The Western Australian Pregnancy Cohort (Raine) Study is a prospective pregnancy cohort where 2900 were recruited from King Edward Memorial Hospital between 1989 and 1991. Data were collected throughout pregnancy and the children have been followed-up at ages 1, 2, 3, 5, 8, 10, 14, 17, 18, 20, and 23. Ethics approval for this study was obtained from King Edward Memorial Hospital and Princess Margaret Hospital. Participants were consented to being involved in this study prior to each follow-up.

Acknowledgements and Funding

This study was supported by the National Health and Medical Research Council of Australia [grant numbers 572613, 403981 and 003209] and the Canadian Institutes of Health Research [grant number MOP-82893]. The authors are grateful to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the NH&MRC for their long term contribution to funding the study over the last 29 years and also the following Institutions for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA), Curtin University, Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health Sciences, The Telethon Kids Institute, Women and Infants Research Foundation (King Edward Memorial Hospital), Murdoch University, The University of Notre Dame (Australia), and Edith Cowan University. The authors gratefully acknowledge the assistance of the Western Australian DNA Bank (National Health and Medical Research Council of Australia National Enabling Facility).

This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

SAGE II

Recruitment

The Study of African Americans, Asthma, Genes & Environments (SAGE II) is an ongoing study of asthma in children and young adults coordinated from the University of California San Francisco. Recruitment protocols and the definition of allergic rhinitis were similar to GALA II with the only differences being that the

recruitment was restricted to San Francisco Bay Area and participants self-identified all four grandparents as African American.⁴⁶ In the current study, 385 cases of allergic rhinitis and 978 controls were analyzed.

Acknowledgements and Funding

The authors acknowledge the patients, families, numerous recruiters, health care providers and community clinics for their participation in the SAGE II study.

N.H-P was funded by a fellowship (FI16/00136) from the Instituto de Salud Carlos III (ISCIII) (NH-P) and co-funded by the European Social Funds from the European Union (ESF) "ESF invests in your future". The SAGE II study was funded by the Sandler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II, National Institutes of Health 1R01HL117004, R01HL128439, R01HL135156, 1X01HL134589, National Institute of Health and Environmental Health Sciences R01ES015794, R21ES24844, the National Institute on Minority Health and Health Disparities 1P60MD006902, U54MD009523, 1R01MD010443, and the Tobacco-Related Disease Research Program under Award Number 24RT-0025 to E.G.B. The content of this publication is solely the responsibility of the authors and does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. We also acknowledge the principal investigators of the centers where the GALA II and SAGE II participants were recruited: Kelley Meade, Harold J. Farber, Pedro C. Avila, Denise Serebrisky, Shannon M. Thyne, Emerita Brigino-Buenaventura, William Rodriguez-Cintron, Saunak Sen, Rajesh Kumar, Michael Lenoir, Luisa N. Borrell, and Jose R. Rodriguez-Santana.

SAPALDIA

Recruitment

SAPALDIA is a population-based cohort from Switzerland that recruited subjects aged 18 to 60 from population registries in eight communities, representing the three largest language groups (German, French, Italian) as well as different levels of air pollution and degrees of urbanization. Subjects underwent spirometry and answered a detailed questionnaire on respiratory health, allergies, smoking history, and lifestyle factors in the baseline (year 1991) and follow-up (year 2002) examination. 6,055 subjects participated in both examinations and agreed to provide blood for genetic analysis. All subjects that were part of a nested asthma case-control sample subjected for genomewide genotyping in the context of the GABRIEL genome-wide association study on asthma formed the basis of the current study. In the current study, subjects were considered to be a case if they ever reported to have had allergic rhinitis at the baseline or follow-up examination. The analysis included 526 cases of allergic rhinitis and 921 controls.

Acknowledgements

Study directorate: NM Probst-Hensch (PI; e/g); T Rochat (p), C Schindler (s), N Künzli (e/exp), JM Gaspoz (c)

Scientific team: JC Barthélémy (c), W Berger (g), R Bettschart (p), A Bircher (a), C Brombach (n), PO Bridevaux (p), L Burdet (p), Felber Dietrich D (e), M Frey (p), U Frey (pd), MW Gerbase (p), D Gold (e), E de Groot (c), W Karrer (p), F Kronenberg (g), B Martin (pa), A Mehta (e), D Miedinger (o), M Pons (p), F Roche (c), T Rothe (p), P Schmid-Grendelmeyer (a), D Stolz (p), A Schmidt-Trucksäss (pa), J Schwartz (e), A Turk (p), A von Eckardstein (cc), E Zemp Stutz (e).

Scientific team at coordinating centers: M Adam (e), I Aguilera (exp), S Brunner (s), D Carballo (c), S Caviezel (pa), I Curjuric (e), A Di Pascale (s), J Dratva (e), R Ducret (s), E Dupuis Lozeron (s), M Eeftens (exp), I Eze (e),

E Fischer (g), M Foraster (e), M Germond (s), L Grize (s), S Hansen (e), A Hensel (s), M Imboden (g), A Ineichen (exp), A Jeong (g), D Keidel (s), A Kumar (g), N Maire (s), A Mehta (e), R Meier (exp), E Schaffner (s), T Schikowski (e), M Tsai (exp).

(a) allergology, (c) cardiology, (cc) clinical chemistry, (e) epidemiology, (exp) exposure, (g) genetic and molecular biology, (m) meteorology, (n) nutrition, (o) occupational health, (p) pneumology, (pa) physical activity, (pd) pediatrics, (s) statistics.

Acknowledgements

The study could not have been done without the help of the study participants, technical and administrative support and the medical teams and field workers at the local study sites.

Local fieldworkers : Aarau: S Brun, G Giger, M Sperisen, M Stahel, Basel: C Bürli, C Dahler, N Oertli, I Harreh, F Karrer, G Novicic, N Wyttenbacher, Davos: A Saner, P Senn, R Winzeler, Geneva: F Bonfils, B Blicharz, C Landolt, J Rochat, Lugano: S Boccia, E Gehrig, MT Mandia, G Solari, B Viscardi, Montana: AP Bieri, C Darioly, M Maire, Payerne: F Ding, P Danieli A Vonnez, Wald: D Bodmer, E Hochstrasser, R Kunz, C Meier, J Rakic, U Schafroth, A Walder.

Administrative staff: N Bauer Ott, C Gabriel, R Gutknecht.

Funding

The Swiss National Science Foundation (grants no 33CS30-148470/1&2, 33CSCO-134276/1, 33CSCO-108796, 324730_135673, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, PMPDP3_129021/1, PMPDP3_141671/1), the Federal Office for the Environment, the Federal Office of Public Health, the Federal Office of Roads and Transport, the canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Valais, and Zürich, the Swiss Lung League, the canton's Lung League of Basel Stadt/ Basel Landschaft, Geneva, Ticino, Valais, Graubünden and Zurich, Stiftung ehemals Bündner Heilstätten, SUVA, Freiwillige Akademische Gesellschaft, UBS Wealth Foundation, Talecris Biotherapeutics GmbH, Abbott Diagnostics, European Commission 018996 (GABRIEL), Wellcome Trust WT 084703MA, Exposomics EC FP7 grant(Grant agreement No: 308610).

SAPPHIRE

Recruitment

The Study of Asthma Phenotypes and Pharmacogenomic Interactions by Race-ethnicity (SAPPHIRE) is an ongoing study that was approved by the Institutional Review Board of Henry Ford Health System. Study individuals included in the current study were members of a large health system, which serves southeast Michigan and all of the Detroit metropolitan statistical area. Individuals were 12-56 years of age and no prior diagnosis of congestive heart failure or chronic obstructive pulmonary disease. Patient recruitment included patients with and without a clinical diagnosis of asthma. Written informed consent was required at the time of enrolment as a condition for study participation. The exam at the time of enrollment included both a staff-administered questionnaire and lung function testing. For the current study, we analyzed only those SAPPHIRE participants who identified themselves as African American.

Acknowledgements and Funding

This work was supported by grants from the Fund for Henry Ford Hospital, the American Asthma Foundation, and the following institutes of the National Institutes of Health: National Institute of Allergy and Infectious Diseases (R01AI079139 and R01AI061774), the National Heart Lung and Blood Institute (R01HL118267 and R01HL079055), and the National Institute of Diabetes and Digestive and Kidney diseases (R01DK064695 and R01DK113003).

UK Biobank

Recruitment

The UK Biobank study is a prospective study of 502,682 volunteer participants aged 40-69 recruited at 22 centers across the UK between 2006 and 2010.⁴⁷ A subset of 152,566 individuals with available genotype data were included in the analysis. All participants have given informed consent for the storage and use of their data for research.

Acknowledgements and Funding

This research has been conducted using the UK Biobank Resource. L Paternoster was supported by a UK Medical Research Council Fellowship (MR/J012165/1) and works in a unit which receives funding from the UK Medical Research Council (MC_UU_12013/1 & MC_UU_12013/4).

Barcelona Supercomputing Center

Acknowledgements and Funding

Josep M. Mercader was supported by Sara Borrell Fellowship from the Instituto Carlos III and Beatriu de Pinós fellowship from the Agency for Management of University and Research Grants (AGAUR). Sílvia Bonàs was FI-DGR Fellowship from FI-DGR 2013 from Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR, Generalitat de Catalunya). We acknowledge PRACE for awarding us access to MareNostrum supercomputer, based in Spain at Barcelona.

Supplementary References

1. Paternoster, L. *et al.* Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat. Genet.* **47**, 1449–1456 (2015).
2. Puel, A., Ziegler, S. F., Buckley, R. H. & Leonard, W. J. Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency. *Nat. Genet.* **20**, 394–397 (1998).
3. Lundmark, F. *et al.* Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. *Nat. Genet.* **39**, 1108–1113 (2007).
4. Gudbjartsson, D. F. *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* **41**, 342–347 (2009).
5. Mori, T. *et al.* Lnk/Sh2b3 controls the production and function of dendritic cells and regulates the induction of IFN- γ -producing T cells. *J. Immunol.* **193**, 1728–1736 (2014).
6. Sleiman, P. M. A. *et al.* GWAS identifies four novel eosinophilic esophagitis loci. *Nat. Commun.* **5**, 5593 (2014).
7. Scott, L. M., Civin, C. I., Rorth, P. & Friedman, A. D. A novel temporal expression pattern of three C/EBP family members in differentiating myelomonocytic cells. *Blood* **80**, 1725–1735 (1992).
8. Gao, H., Parkin, S., Johnson, P. F. & Schwartz, R. C. C/EBP gamma has a stimulatory role on the IL-6 and IL-8 promoters. *J. Biol. Chem.* **277**, 38827–38837 (2002).
9. León, B. *et al.* Regulation of T(H)2 development by CXCR5+ dendritic cells and lymphotoxin-expressing B cells. *Nat. Immunol.* **13**, 681–690 (2012).
10. Zhang, J. *et al.* Three SNPs in chromosome 11q23.3 are independently associated with systemic lupus erythematosus in Asians. *Hum. Mol. Genet.* **23**, 524–533 (2014).
11. International Multiple Sclerosis Genetics Consortium *et al.* MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain* **136**, 1778–1782 (2013).
12. Charbonneau, B. *et al.* CXCR5 polymorphisms in non-Hodgkin lymphoma risk and prognosis. *Cancer Immunol. Immunother.* **62**, 1475–1484 (2013).
13. Lawrence, T. The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb. Perspect. Biol.* **1**, a001651 (2009).
14. Shinnakasu, R. *et al.* Regulated selection of germinal-center cells into the memory B cell compartment. *Nat. Immunol.* **17**, 861–869 (2016).
15. Roychoudhuri, R. *et al.* BACH2 regulates CD8(+) T cell differentiation by controlling access of AP-1 factors to enhancers. *Nat. Immunol.* **17**, 851–860 (2016).
16. Rothlin, C. V., Ghosh, S., Zuniga, E. I., Oldstone, M. B. A. & Lemke, G. TAM receptors are pleiotropic inhibitors of the innate immune response. *Cell* **131**, 1124–1136 (2007).
17. Chan, P. Y. *et al.* The TAM family receptor tyrosine kinase TYRO3 is a negative regulator of type 2 immunity. *Science* **352**, 99–103 (2016).
18. Kassmeier, M. D. *et al.* VprBP binds full-length RAG1 and is required for B-cell development and V(D)J recombination fidelity. *EMBO J.* **31**, 945–958 (2012).
19. Hamblet, C. E., Makowski, S. L., Tritapoe, J. M. & Pomerantz, J. L. NK Cell Maturation and Cytotoxicity Are Controlled by the Intramembrane Aspartyl Protease SPPL3. *J. Immunol.* **196**, 2614–2626 (2016).
20. Andersen, J. B., Strandbygård, D. J., Hartmann, R. & Justesen, J. Interaction between the 2'-5' oligoadenylate synthetase-like protein p59 OASL and the transcriptional repressor methyl CpG-binding

protein 1. *Eur. J. Biochem.* **271**, 628–636 (2004).

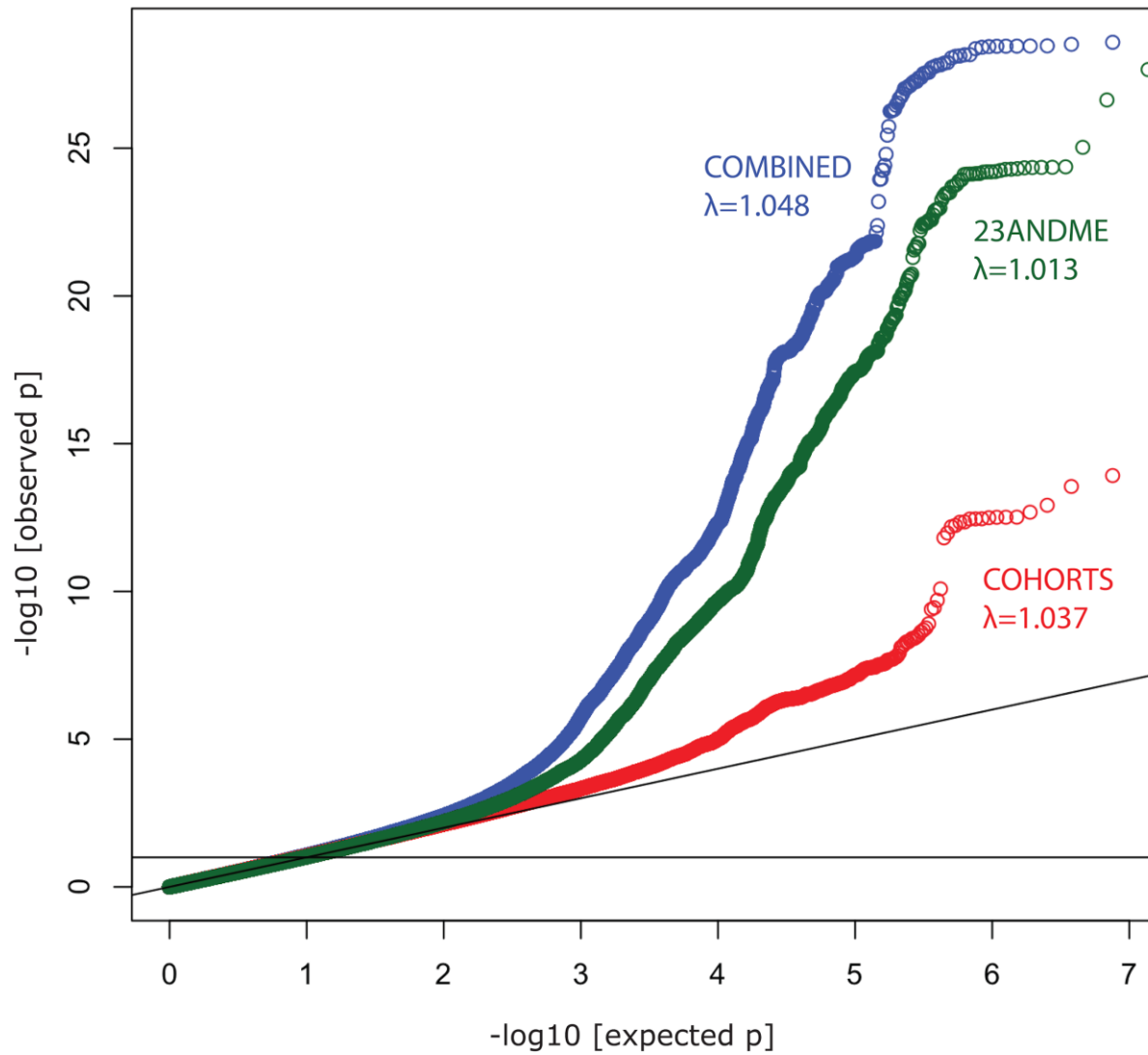
21. Halim, T. Y. F. *et al.* Retinoic-acid-receptor-related orphan nuclear receptor alpha is required for natural helper cell development and allergic inflammation. *Immunity* **37**, 463–474 (2012).
22. Anderson, D. M. *et al.* A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature* **390**, 175–179 (1997).
23. Bousquet, J. *et al.* Important research questions in allergy and related diseases: nonallergic rhinitis: a GA2LEN paper. *Allergy* **63**, 842–853 (2008).
24. Boyd, A. *et al.* Cohort Profile: the ‘children of the 90s’--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
25. Bisgaard, H. The Copenhagen Prospective Study on Asthma in Childhood (COPSAC): design, rationale, and baseline data from a longitudinal birth cohort study. *Ann. Allergy Asthma Immunol.* **93**, 381–389 (2004).
26. Gudbjartsson, D. F. *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* **41**, 342–347 (2009).
27. Smith, D. *et al.* A rare IL33 loss-of-function mutation reduces blood eosinophil counts and protects from asthma. *PLoS Genet.* **13**, e1006659 (2017).
28. Gudbjartsson, D. F. *et al.* Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435–444 (2015).
29. Burney, P. G., Luczynska, C., Chinn, S. & Jarvis, D. The European Community Respiratory Health Survey. *Eur. Respir. J.* **7**, 954–960 (1994).
30. Kaprio, J. Twin studies in Finland 2006. *Twin Res. Hum. Genet.* **9**, 772–777 (2006).
31. Kaprio, J. The Finnish Twin Cohort Study: an update. *Twin Res. Hum. Genet.* **16**, 157–162 (2013).
32. Kaprio, J. & Koskenvuo, M. Genetic and environmental factors in complex diseases: the older Finnish Twin Cohort. *Twin Res.* **5**, 358–365 (2002).
33. Ege, M. J. *et al.* Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. *J. Allergy Clin. Immunol.* **127**, 138–44, 144.e1–4 (2011).
34. Pino-Yanes, M. *et al.* Genetic ancestry influences asthma susceptibility and lung function among Latinos. *J. Allergy Clin. Immunol.* **135**, 228–235 (2015).
35. Krawczak, M. *et al.* PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet.* **9**, 55–61 (2006).
36. Wichmann, H.-E., Gieger, C., Illig, T. & MONICA/KORA Study Group. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1**, S26–30 (2005).
37. Esparza-Gordillo, J. *et al.* A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nat. Genet.* **41**, 596–601 (2009).
38. Völzke, H. *et al.* Cohort profile: the study of health in Pomerania. *Int. J. Epidemiol.* **40**, 294–307 (2011).
39. Heinrich, J. *et al.* Allergens and endotoxin on mothers’ mattresses and total immunoglobulin E in cord blood of neonates. *Eur. Respir. J.* **20**, 617–623 (2002).
40. Berg, A. v. *et al.* Impact of early feeding on childhood eczema: development after nutritional intervention compared with the natural course - the GINIplus study up to the age of 6 years. *Clin. Exp. Allergy* **40**, 627–636 (2010).
41. Barreto-Luis, A. *et al.* Genome-wide association study in Spanish identifies ADAM metallopeptidase

with thrombospondin type 1 motif, 9 (ADAMTS9), as a novel asthma susceptibility gene. *J. Allergy Clin. Immunol.* **137**, 964–966 (2016).

42. Barreto-Luis, A. *et al.* A pathway-based association study reveals variants from Wnt signalling genes contributing to asthma susceptibility. *Clin. Exp. Allergy* **47**, 618–626 (2017).
43. Paternoster, L. *et al.* Genome-wide population-based association study of extremely overweight young adults--the GOYA study. *PLoS One* **6**, e24303 (2011).
44. Guxens, M. *et al.* Cohort Profile: the INMA--Infancia y Medio Ambiente--(Environment and Childhood) Project. *Int. J. Epidemiol.* **41**, 930–940 (2012).
45. van Beijsterveldt, C. E. M. & Boomsma, D. I. Genetics of parentally reported asthma, eczema and rhinitis in 5-yr-old twins. *Eur. Respir. J.* **29**, 516–521 (2007).
46. White, M. J. *et al.* Novel genetic risk factors for asthma in African American children: Precision Medicine and the SAGE II Study. *Immunogenetics* **68**, 391–400 (2016).
47. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).

Supplementary Figure 1

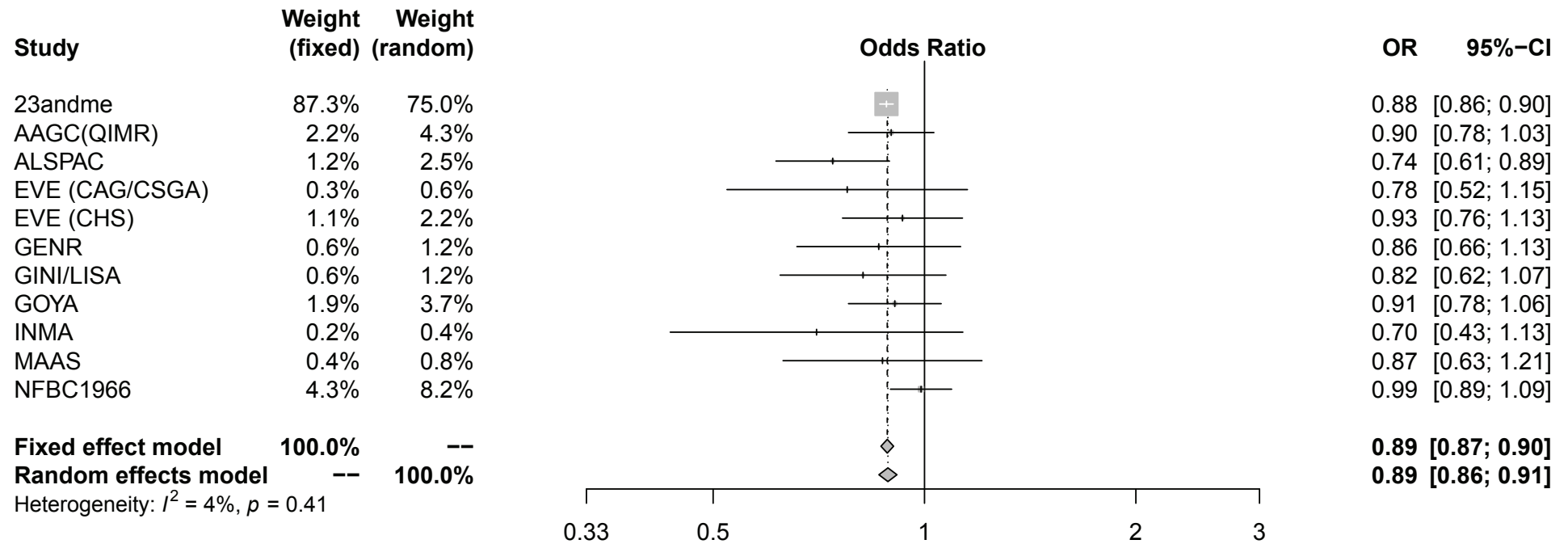
QQ-plot for the discovery phase of the allergic rhinitis GWAS, based on p-values from an inverse variance weighted fixed-effect meta-analysis of genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals). λ indicates the genomic inflation rate.



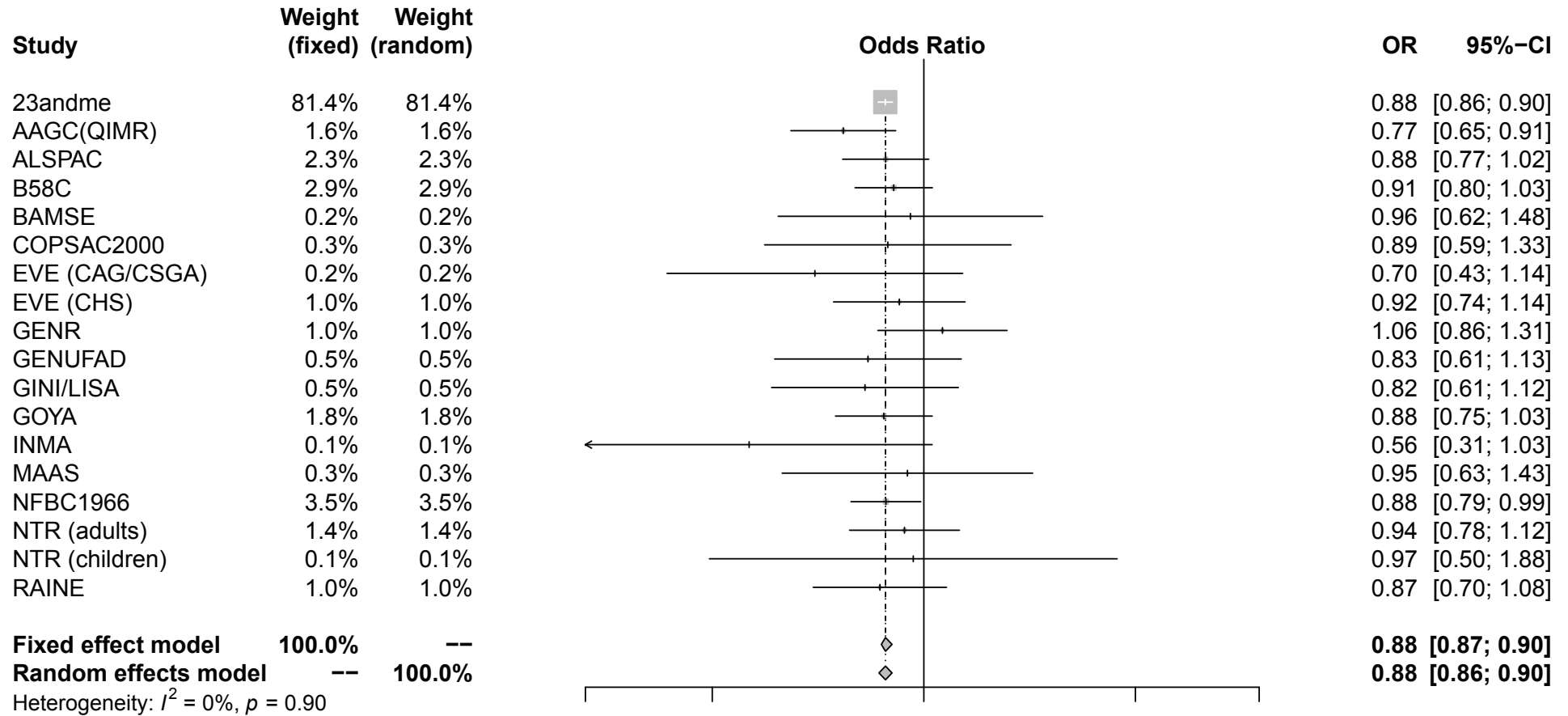
Supplementary Figure 2

Forest plots from the discovery phase for markers previously associated to allergy, including weights, estimates and 95% confidence intervals for inverse variance weighted fixed and random effects meta-analyses from genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals.

rs34004019[G]

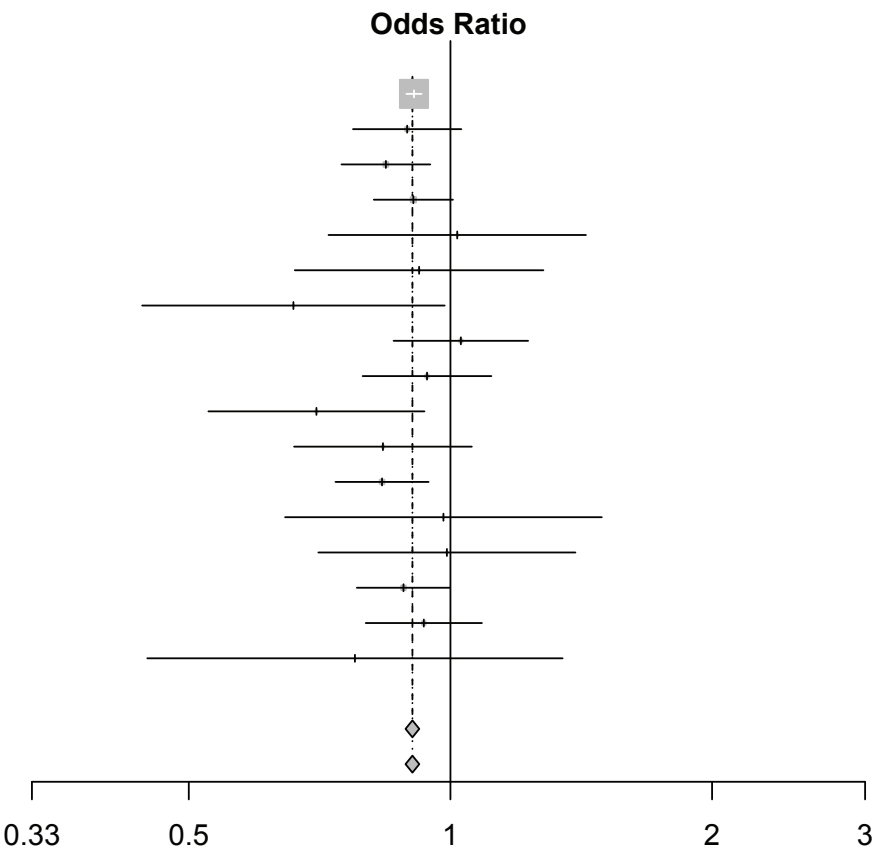


rs950881[T]

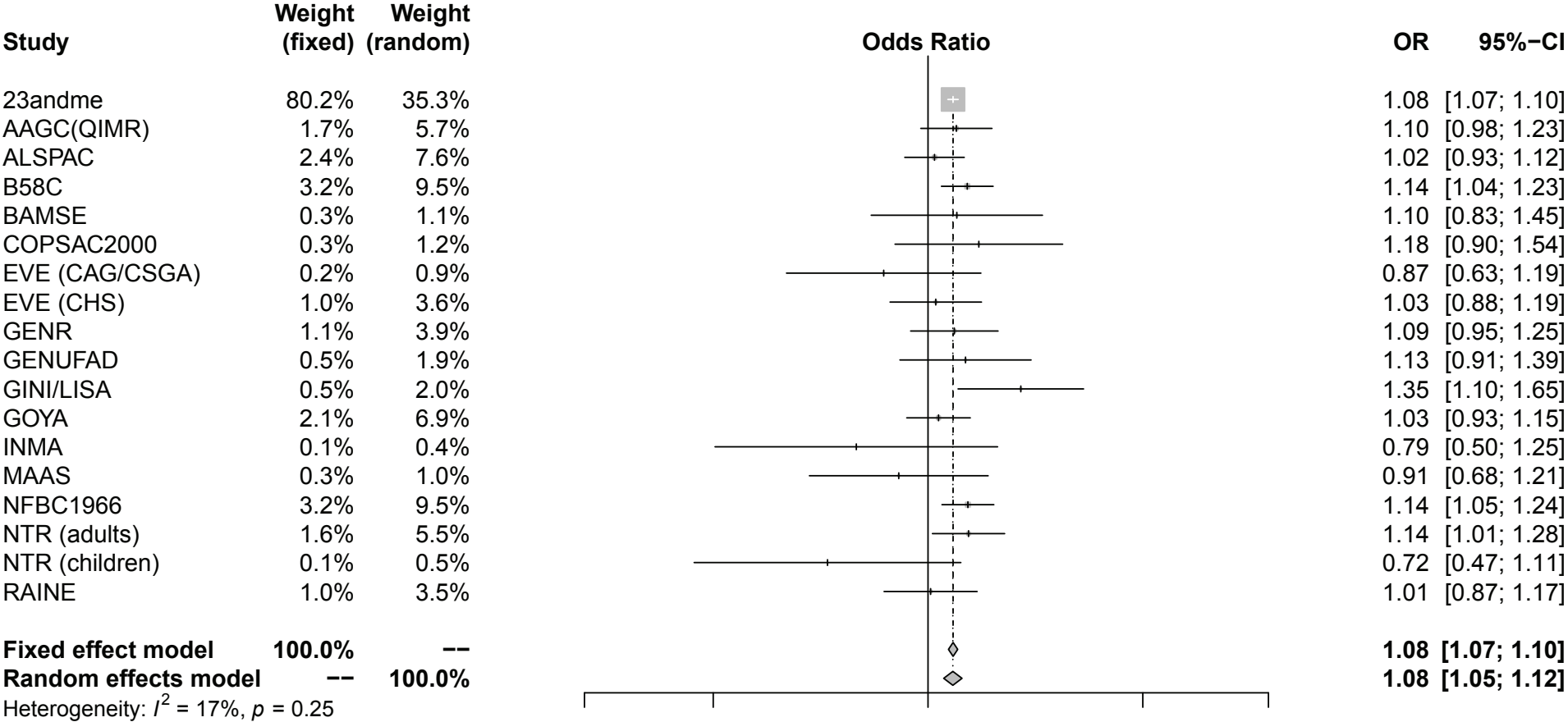


rs5743618[A]

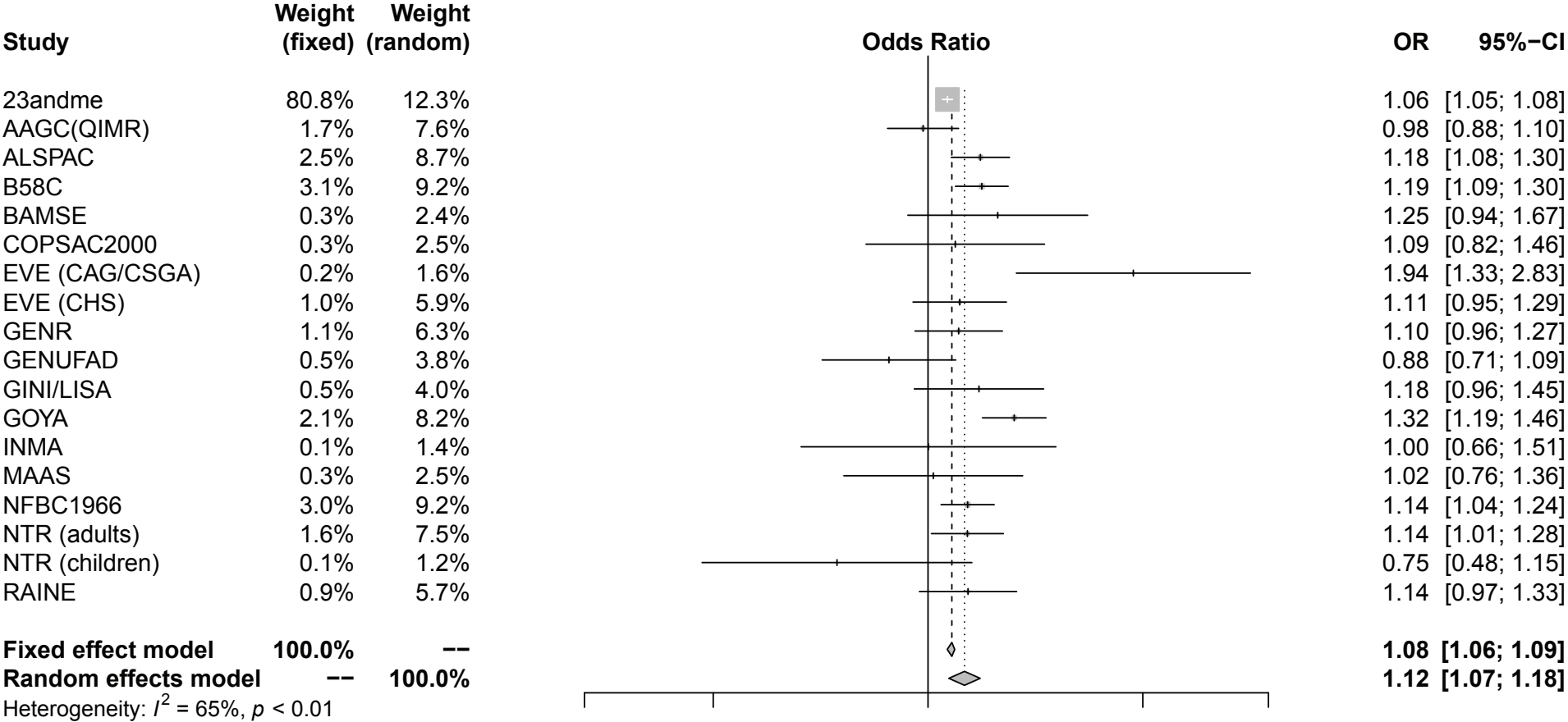
Study	Weight (fixed)	Weight (random)
23andme	82.5%	82.5%
AAGC(QIMR)	1.6%	1.6%
ALSPAC	2.4%	2.4%
B58C	3.0%	3.0%
BAMSE	0.3%	0.3%
COPSAC2000	0.3%	0.3%
EVE (CAG/CSGA)	0.2%	0.2%
EVE (CHS)	1.1%	1.1%
GENR	1.2%	1.2%
GENUFAD	0.4%	0.4%
GINI/LISA	0.6%	0.6%
GOYA	2.2%	2.2%
INMA	0.2%	0.2%
MAAS	0.3%	0.3%
NFBC1966	2.2%	2.2%
NTR (adults)	1.4%	1.4%
NTR (children)	0.1%	0.1%
Fixed effect model	100.0%	--
Random effects model	--	100.0%
Heterogeneity: $I^2 = 0\%$, $p = 0.69$		



rs1438673[C]

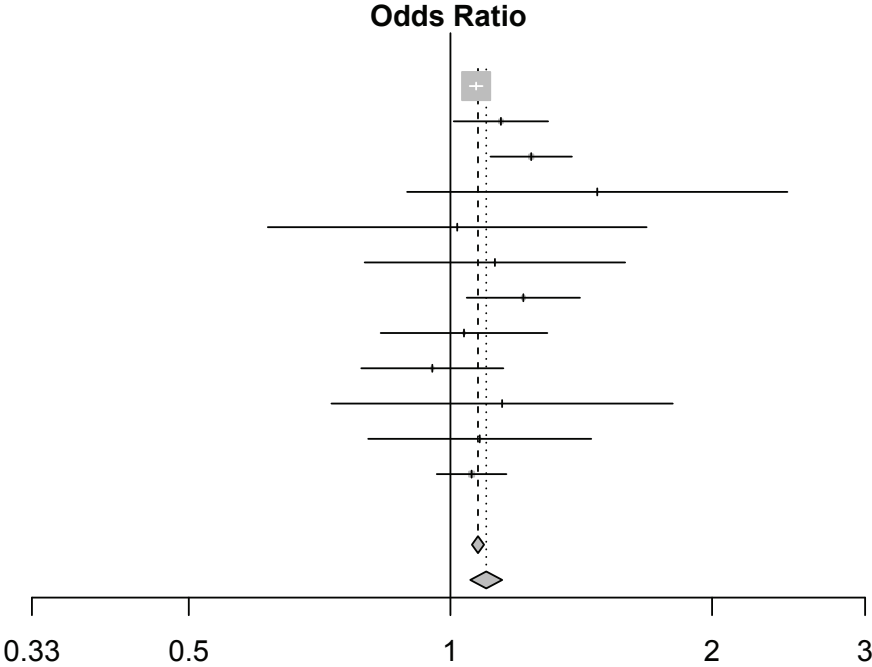


rs7936323[A]

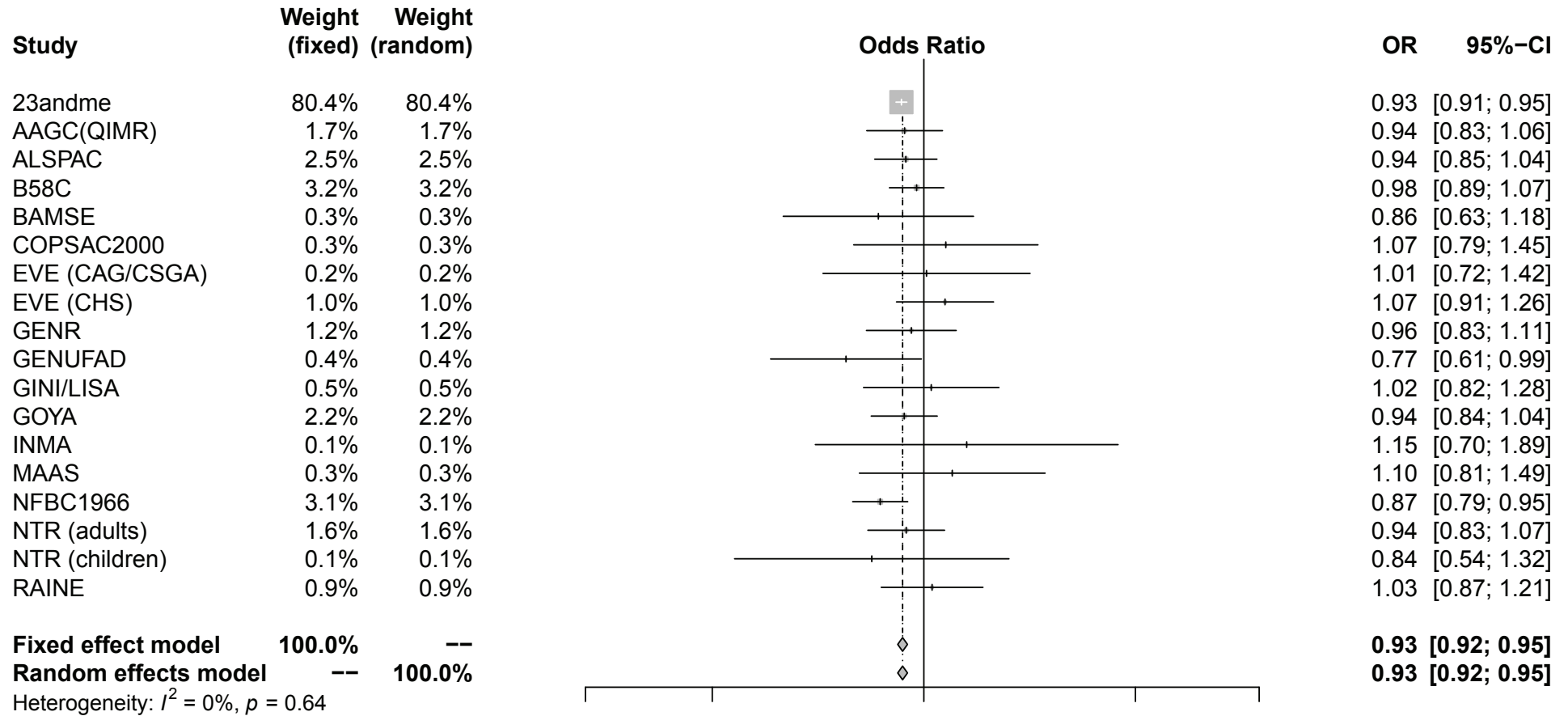


rs2428494[A]

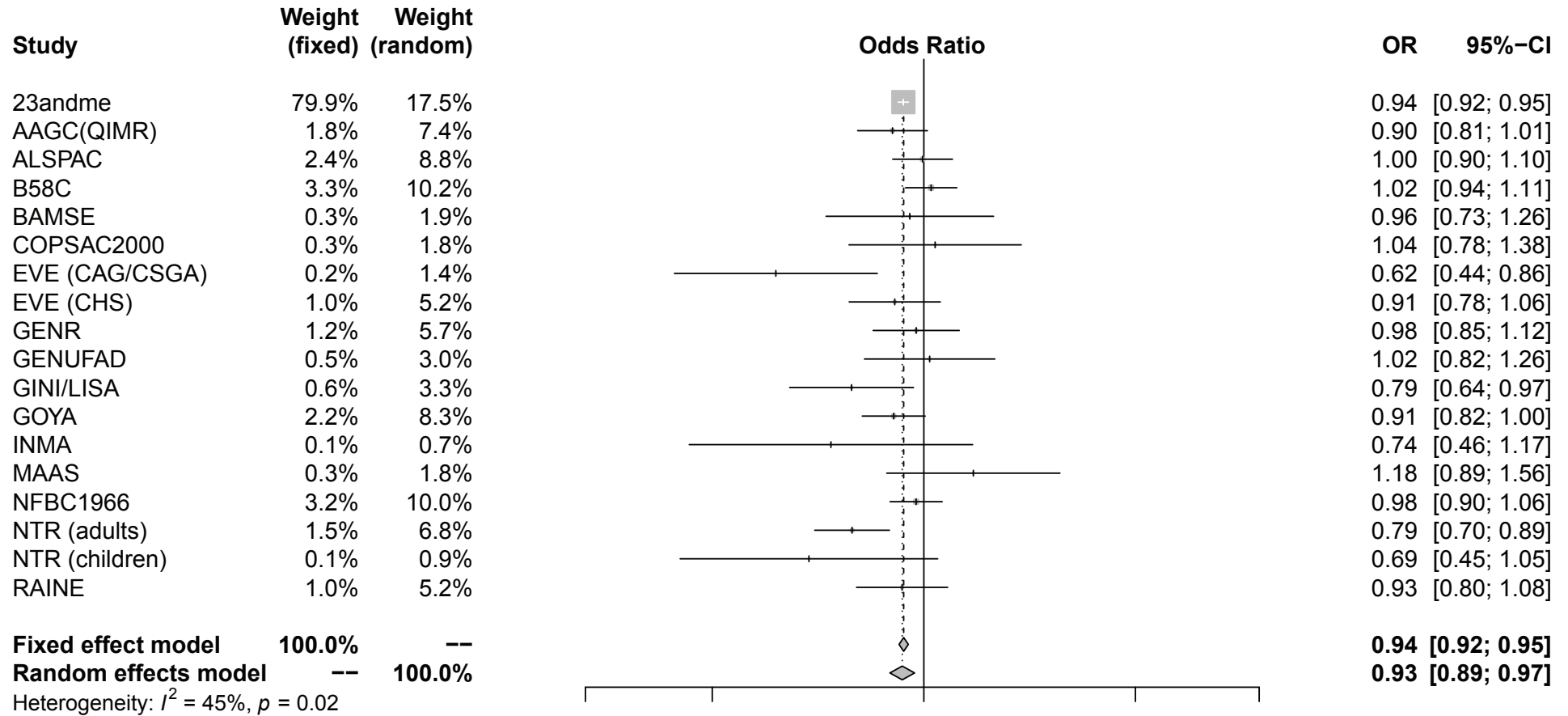
Study	Weight (fixed)	Weight (random)
23andme	89.8%	44.1%
AAGC(QIMR)	1.7%	9.2%
ALSPAC	2.2%	11.6%
BAMSE	0.1%	0.7%
COPSAC2000	0.1%	0.7%
EVE (CAG/CSGA)	0.2%	1.5%
EVE (CHS)	1.2%	6.8%
GINI/LISA	0.5%	3.4%
GOYA	0.7%	4.6%
INMA	0.1%	0.9%
MAAS	0.3%	2.0%
NFBC1966	3.0%	14.6%
Fixed effect model	100.0%	--
Random effects model	--	100.0%
Heterogeneity: $I^2 = 20\%$, $p = 0.24$		



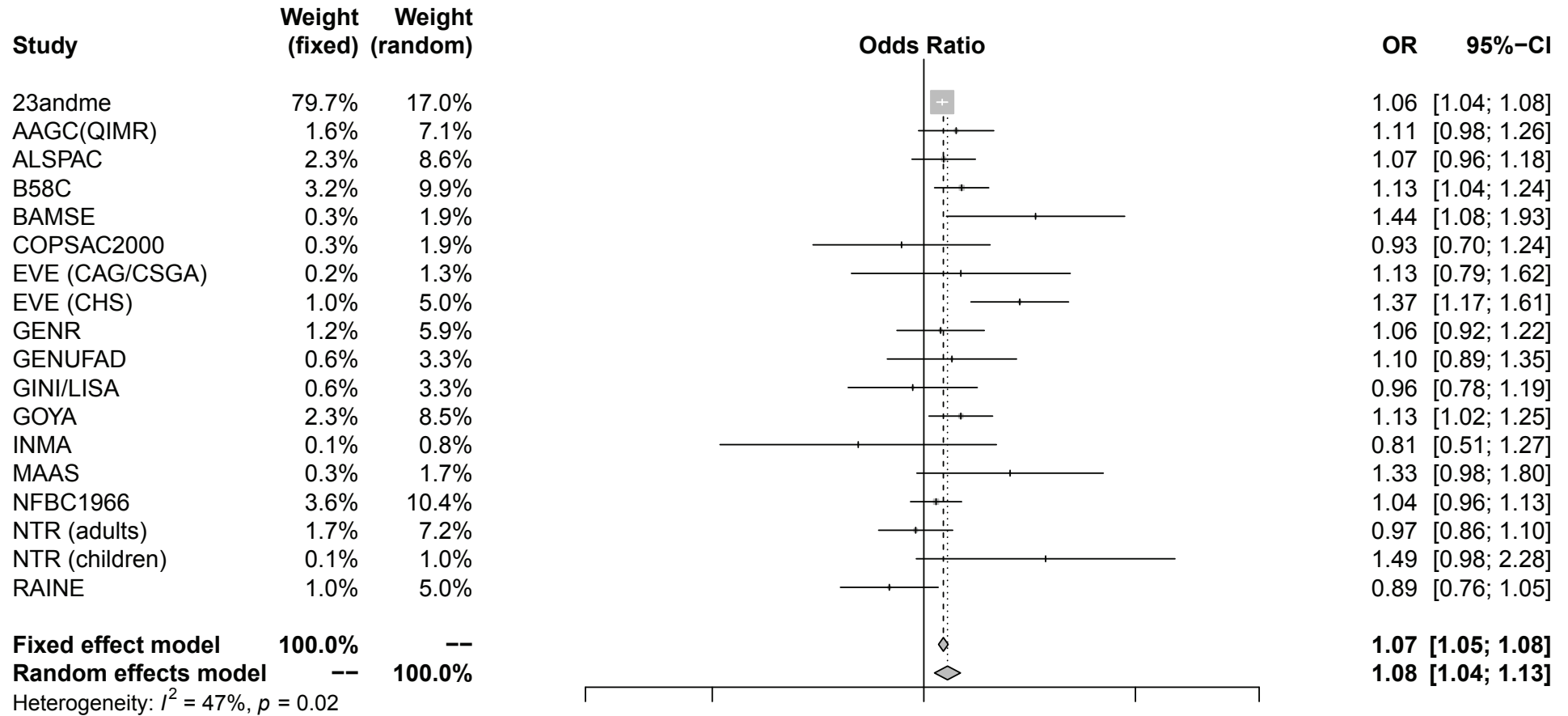
rs11644510[T]



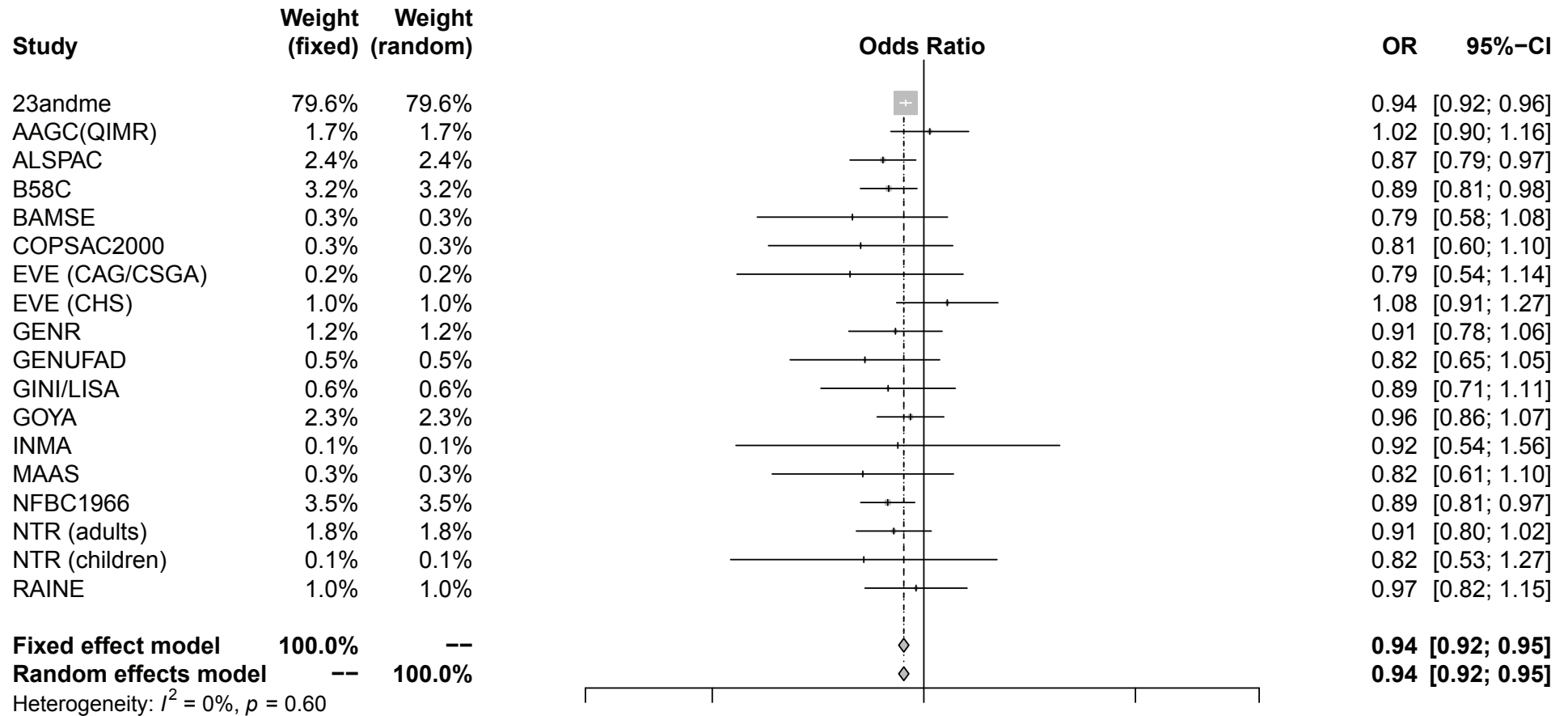
rs12939457[C]



rs148505069[G]

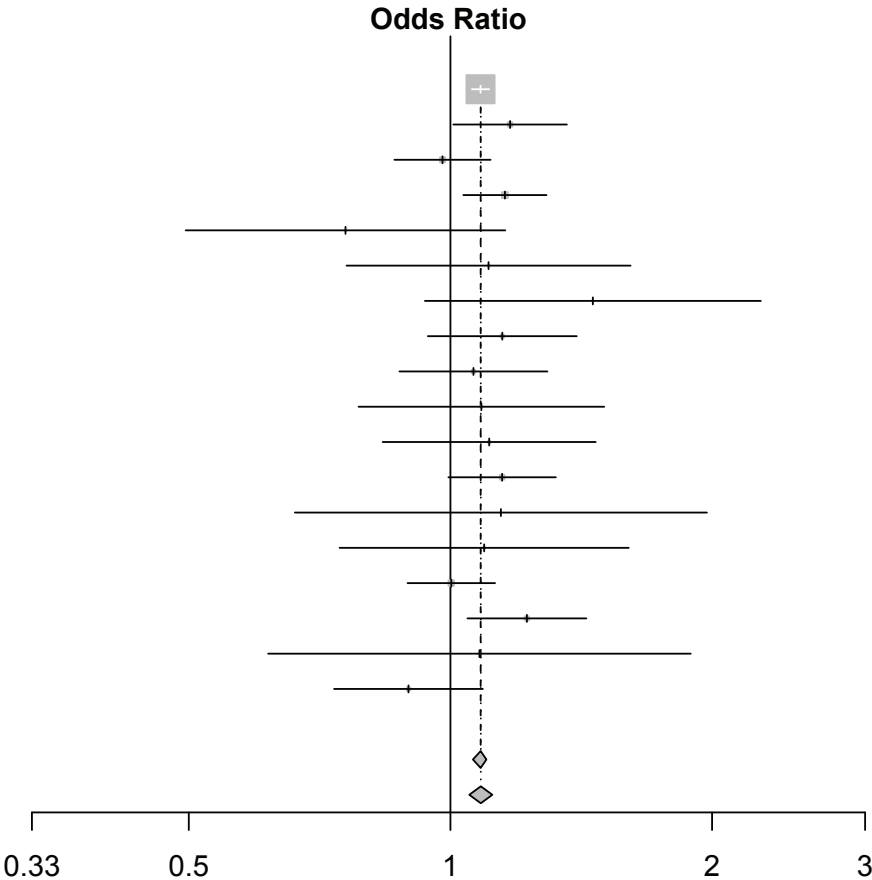


rs13395467[G]

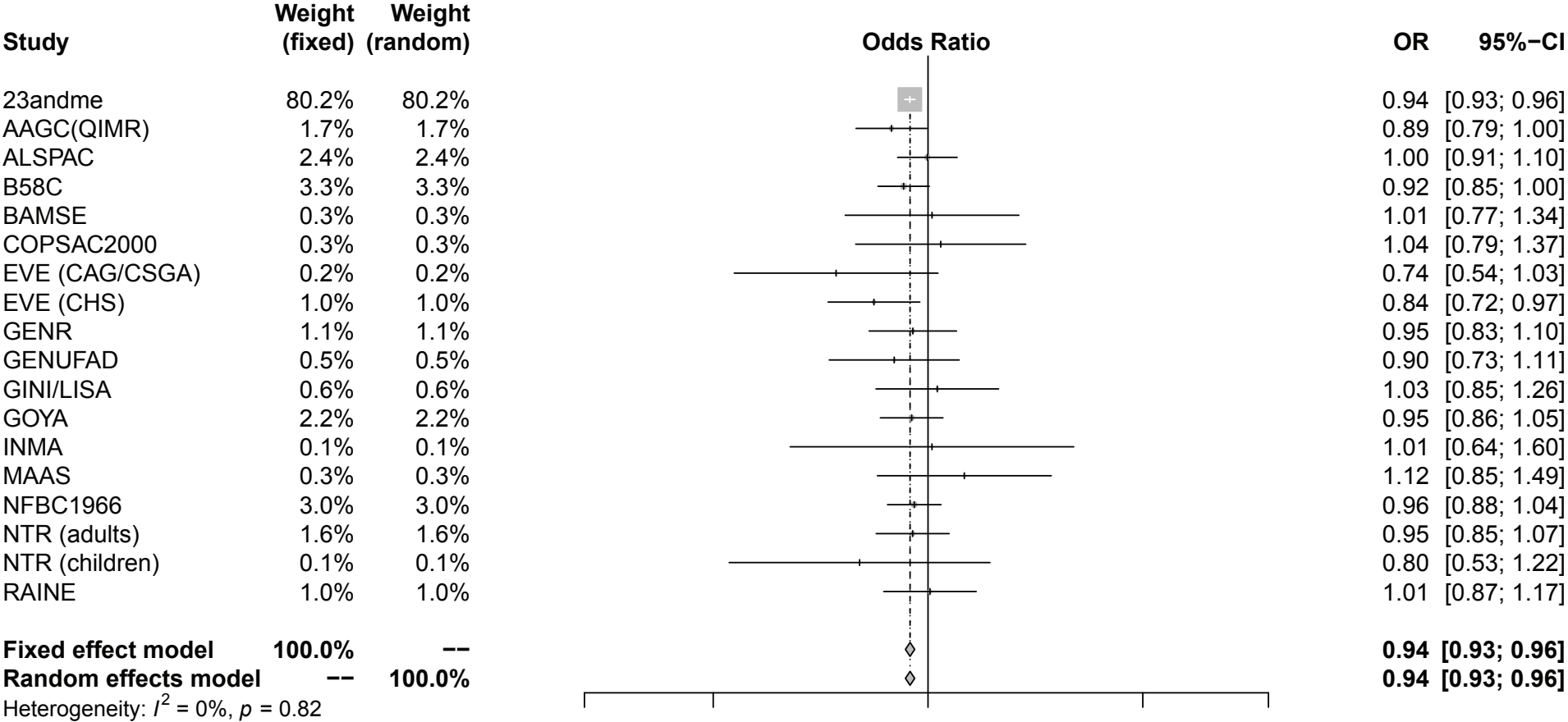


rs9775039[A]

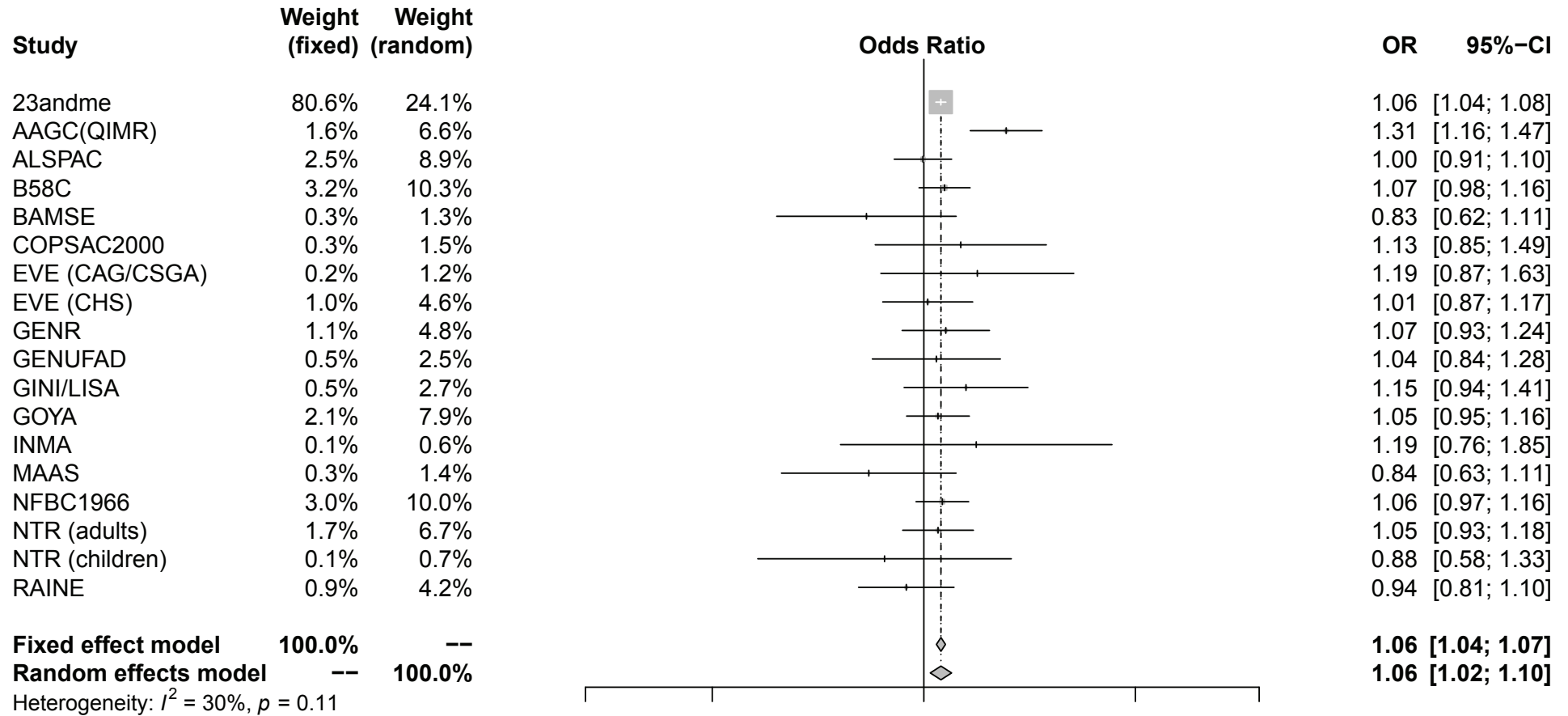
Study	Weight (fixed)	Weight (random)
23andme	79.7%	57.3%
AAGC(QIMR)	1.9%	3.9%
ALSPAC	2.6%	5.4%
B58C	3.5%	7.1%
BAMSE	0.2%	0.5%
COPSAC2000	0.3%	0.7%
EVE (CAG/CSGA)	0.2%	0.5%
EVE (CHS)	1.1%	2.3%
GENR	1.1%	2.4%
GENUFAD	0.4%	0.9%
GINI/LISA	0.5%	1.2%
GOYA	2.1%	4.4%
INMA	0.1%	0.3%
MAAS	0.3%	0.6%
NFBC1966	3.1%	6.4%
NTR (adults)	1.7%	3.6%
NTR (children)	0.1%	0.3%
RAINE	1.1%	2.3%
Fixed effect model	100.0%	--
Random effects model	--	100.0%
Heterogeneity: $I^2 = 5\%$, $p = 0.39$		



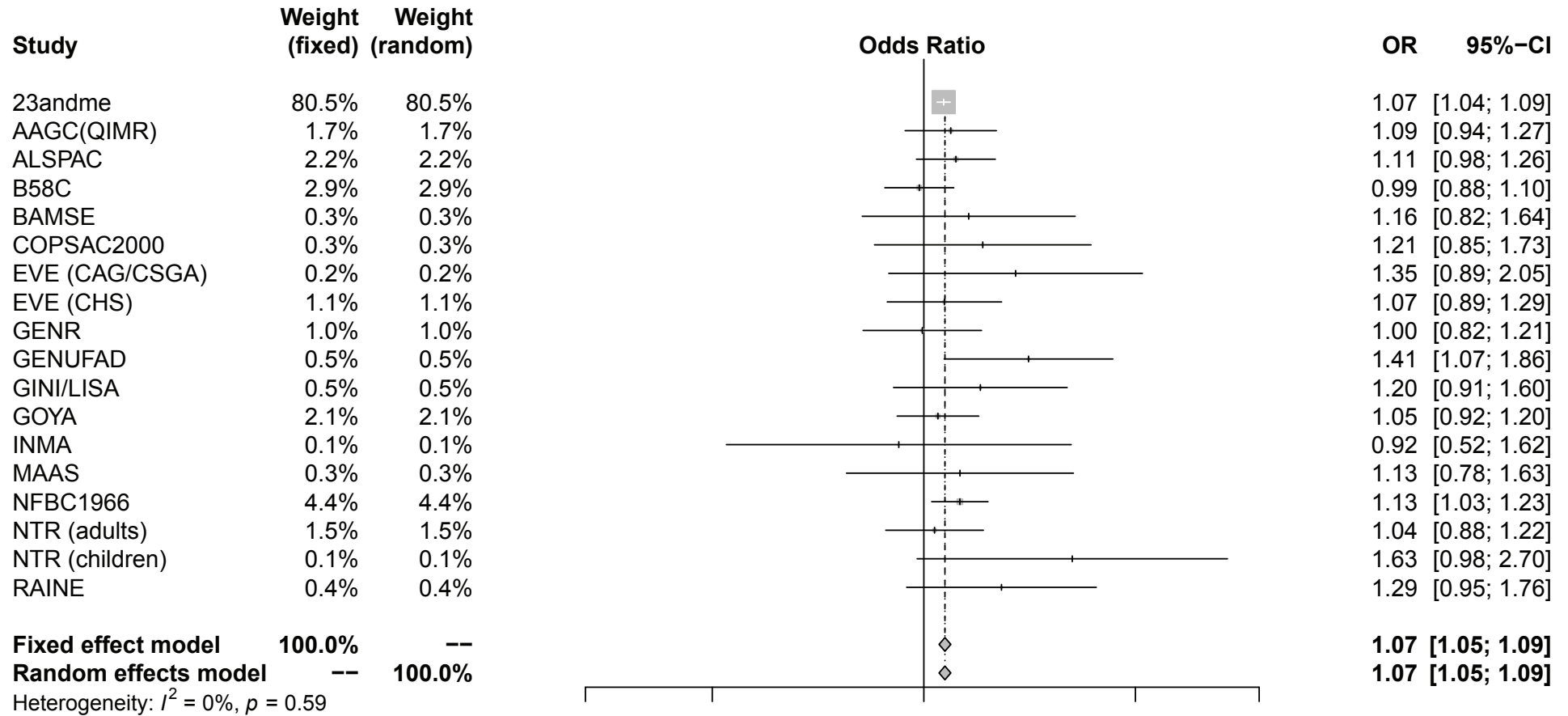
rs2164068[A]



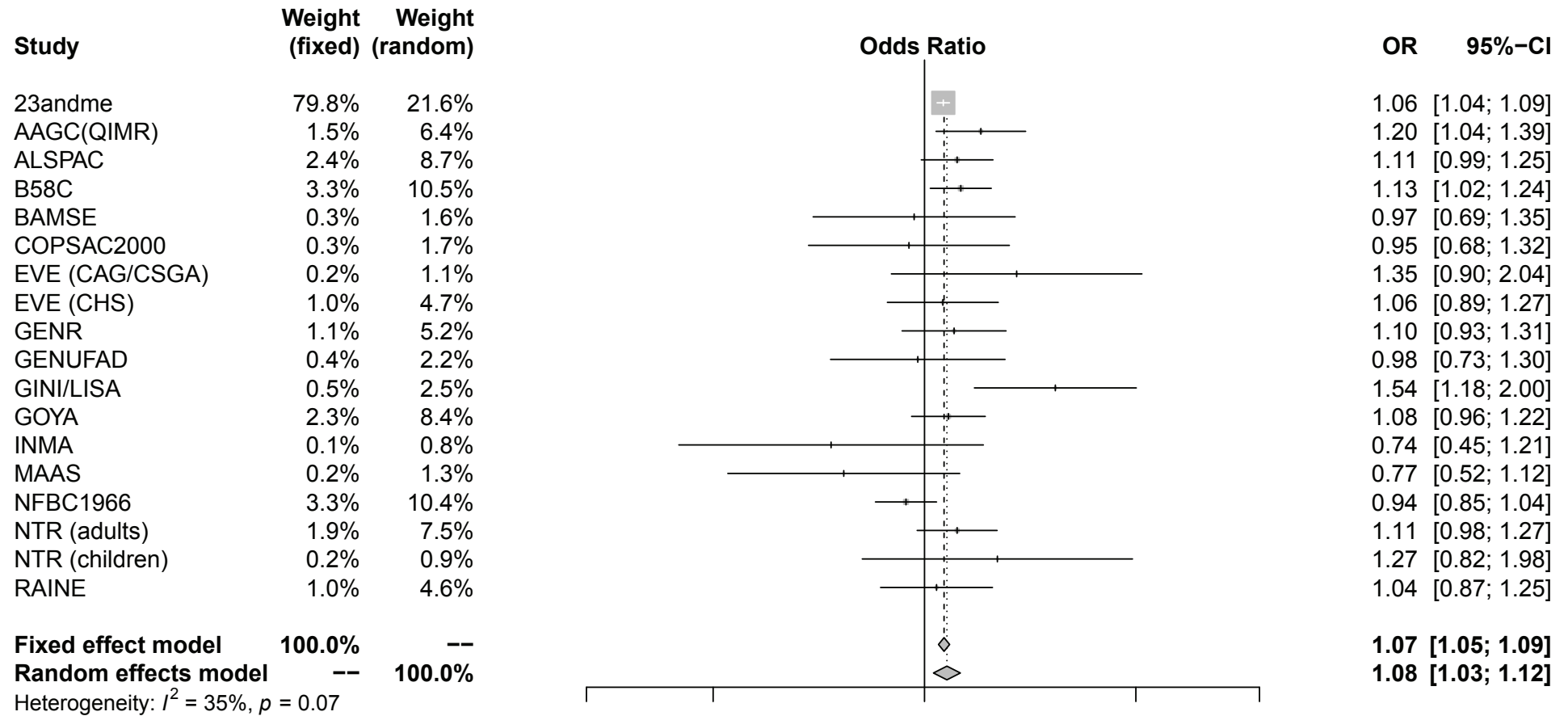
rs2030519[G]



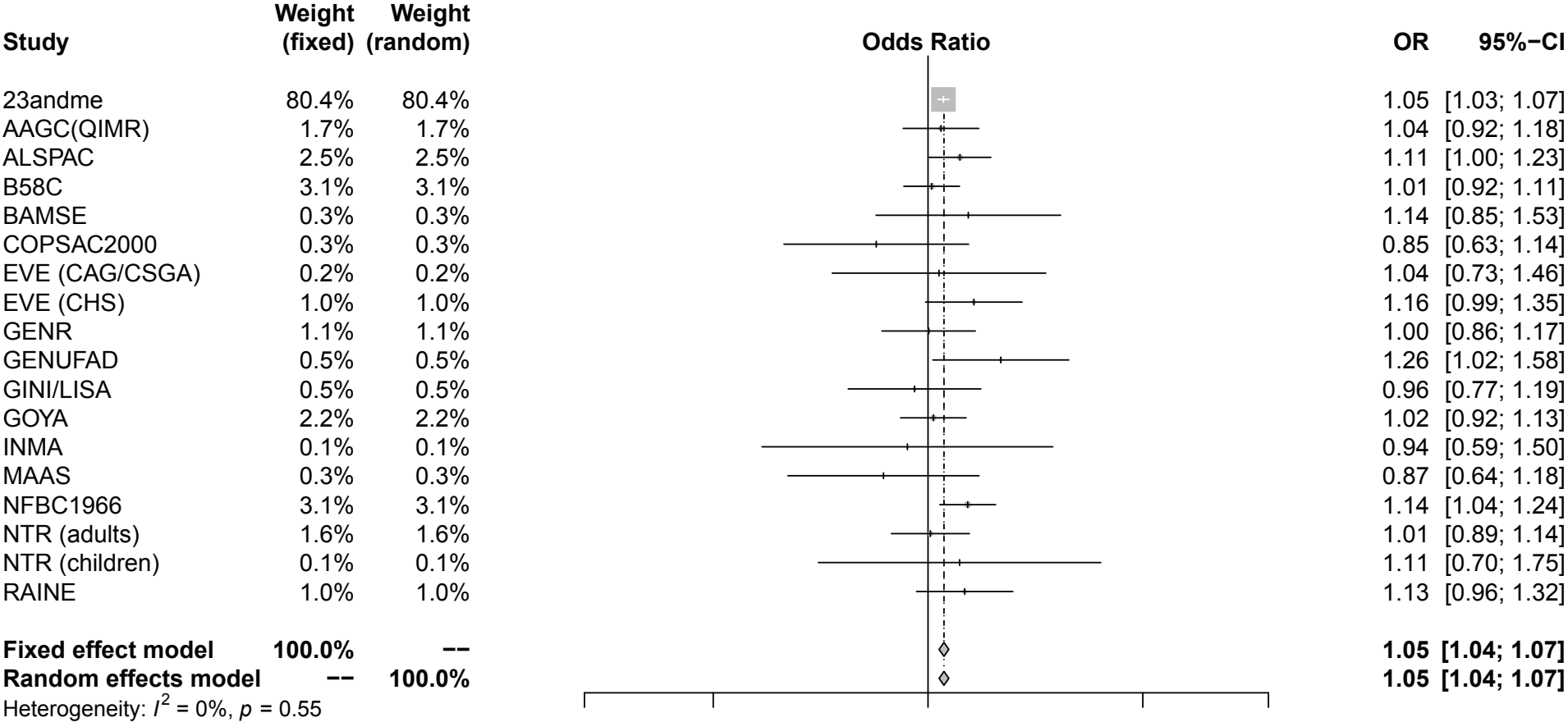
rs11256017[T]



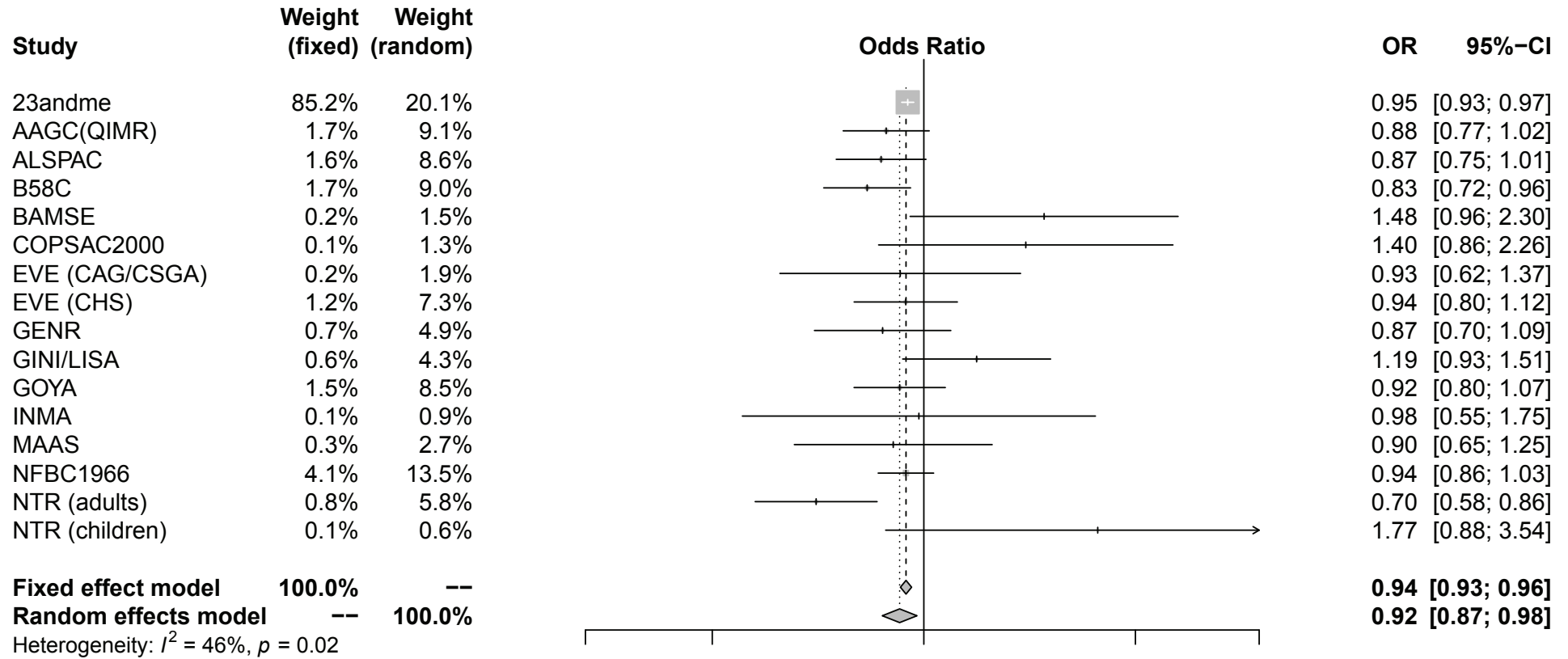
rs17294280[G]



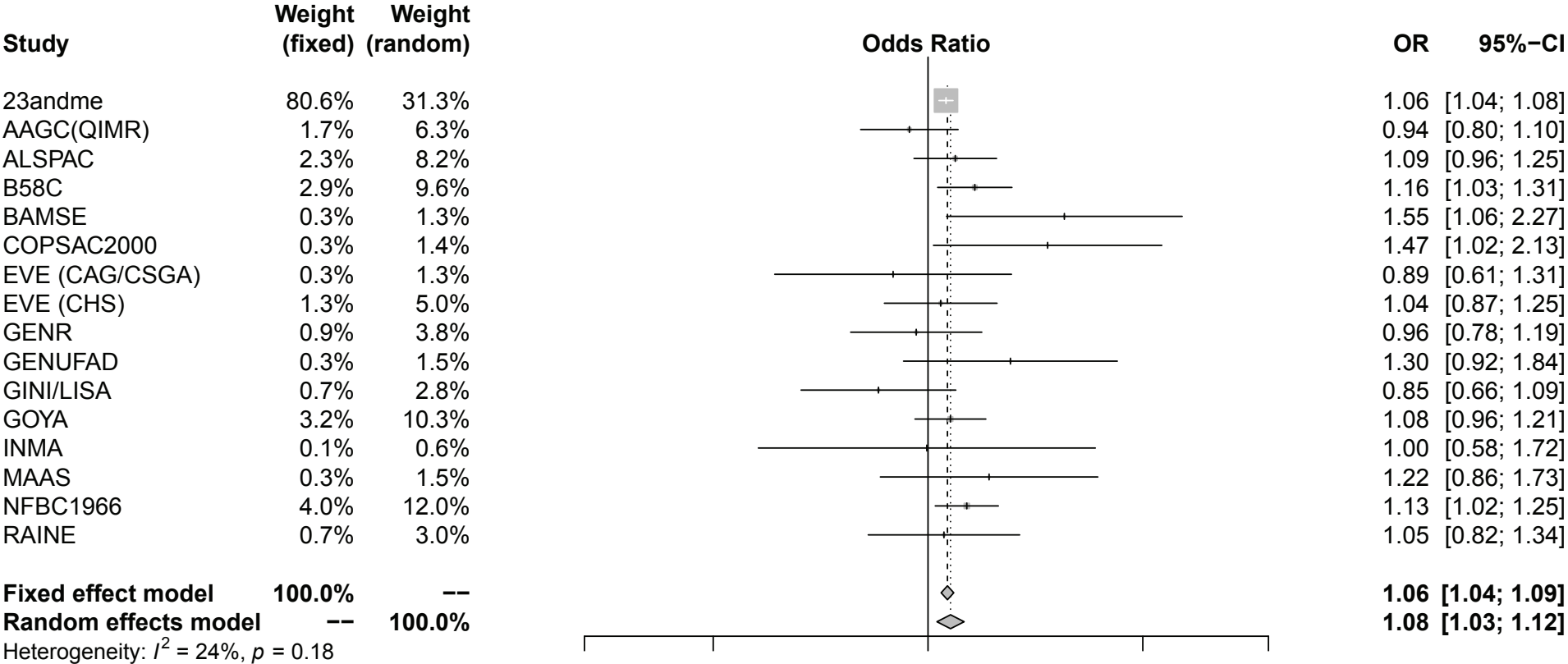
rs7824993[A]



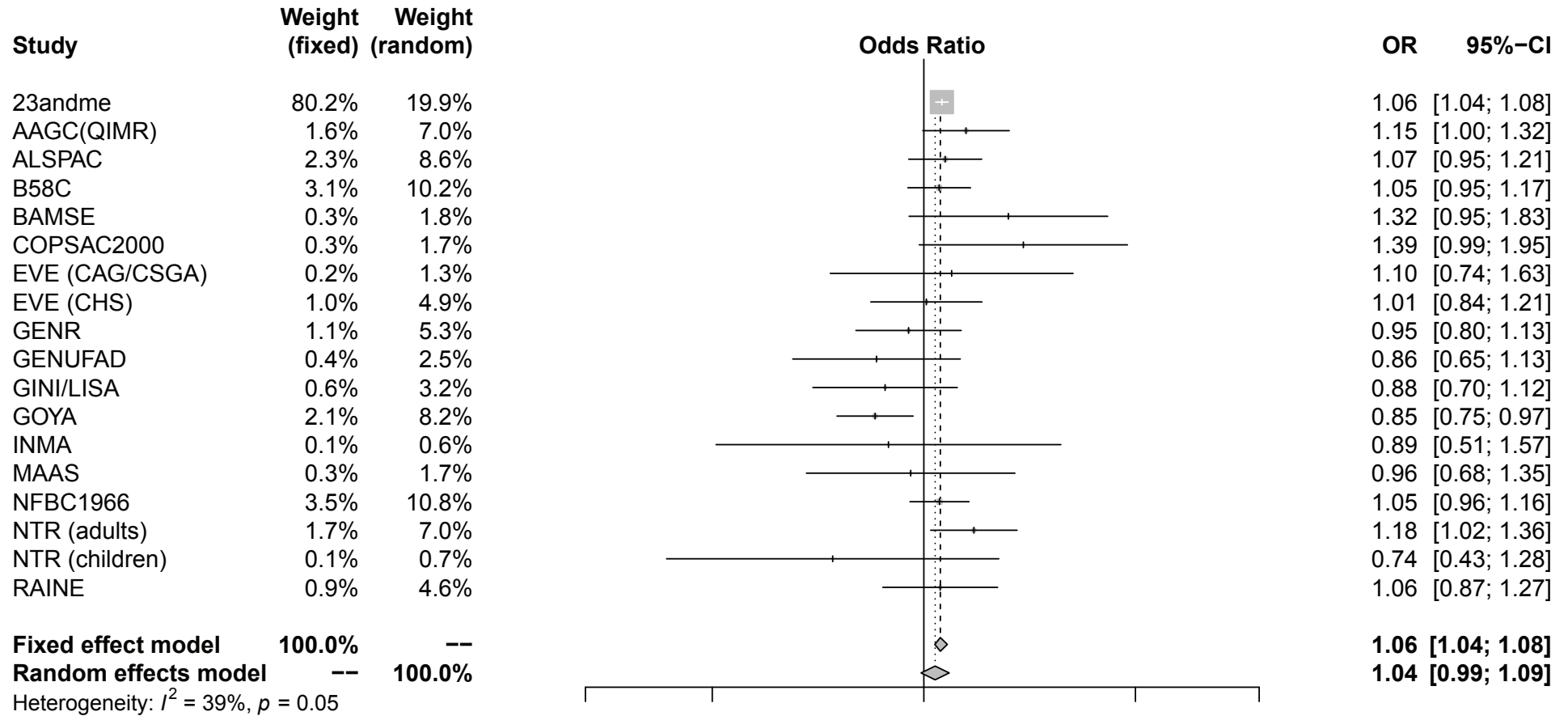
rs9282864[C]



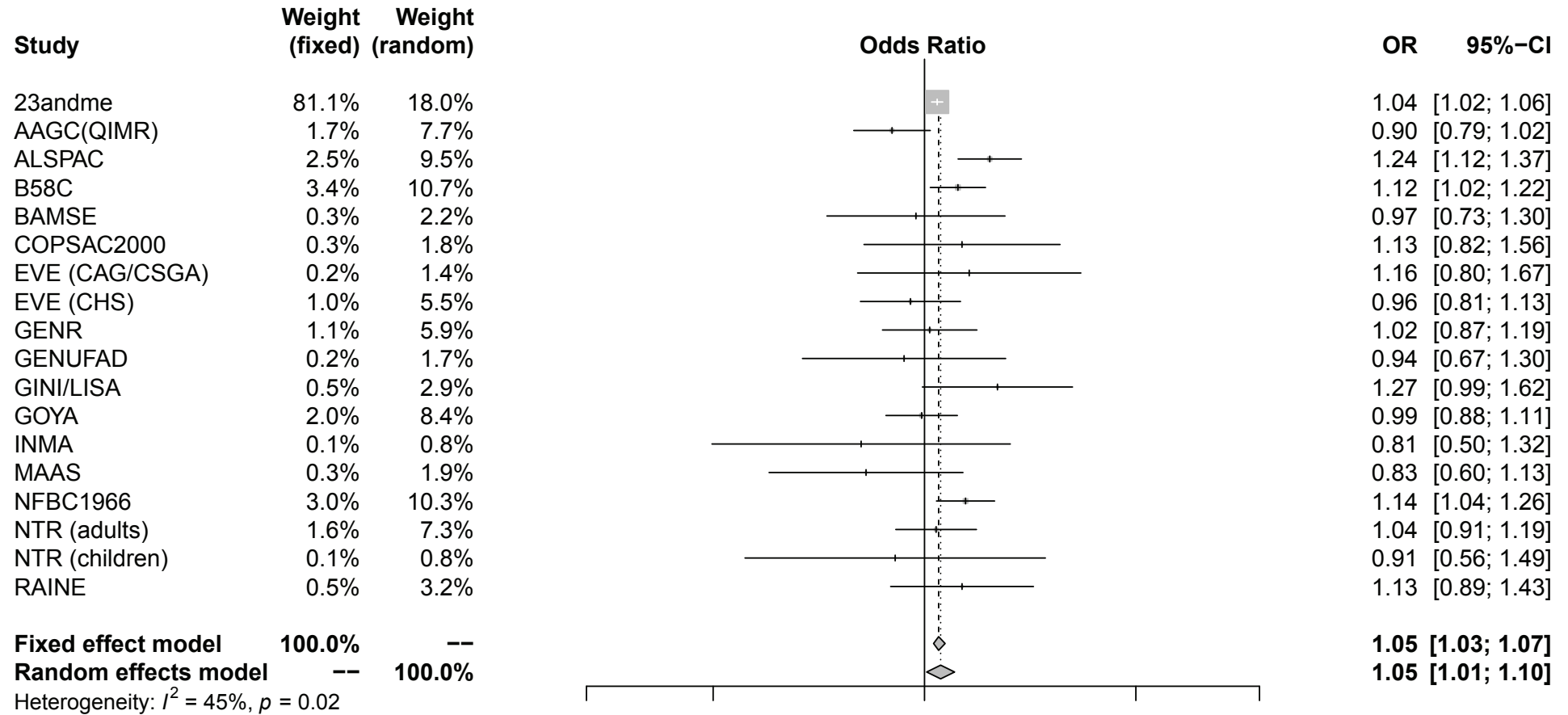
rs9687749[T]



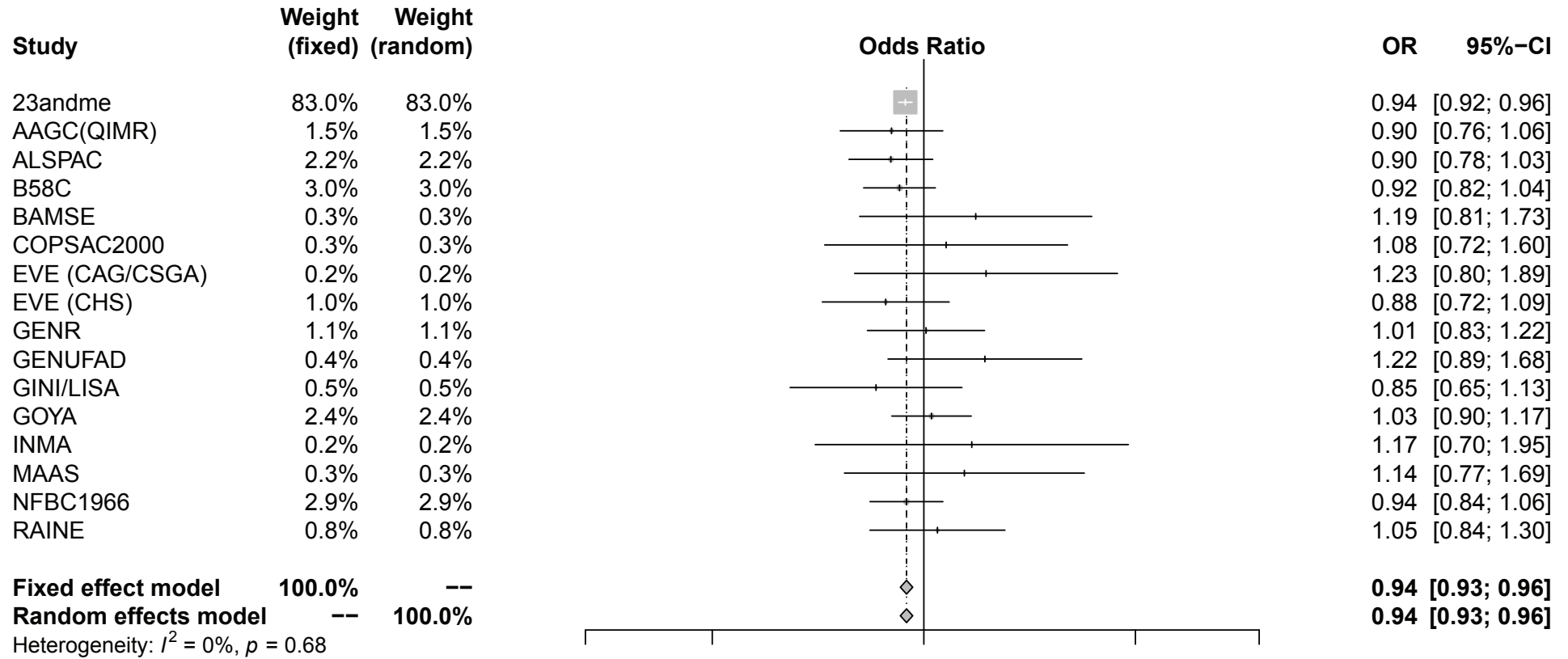
rs61977073[G]



rs6470578[T]



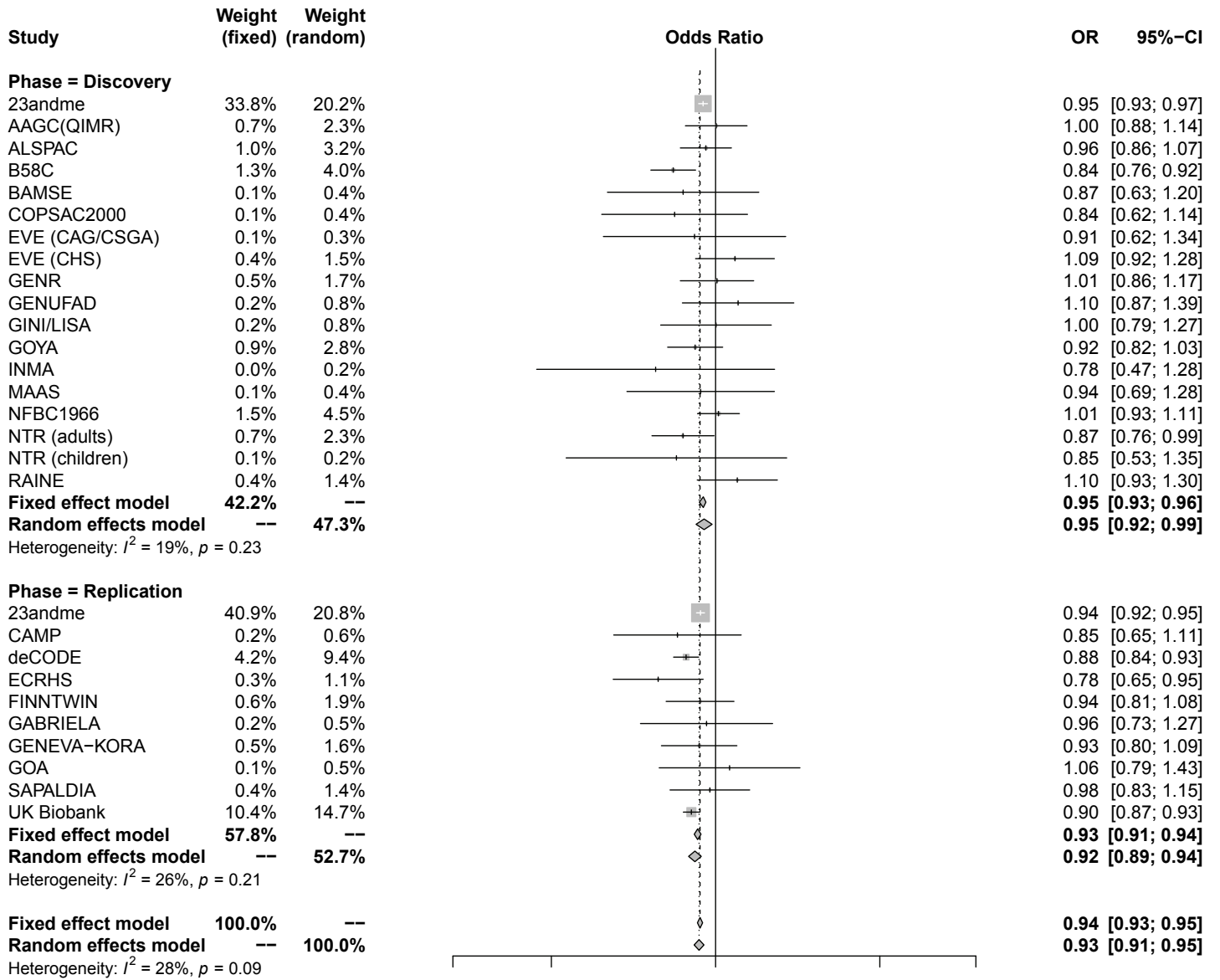
rs3787184[G]



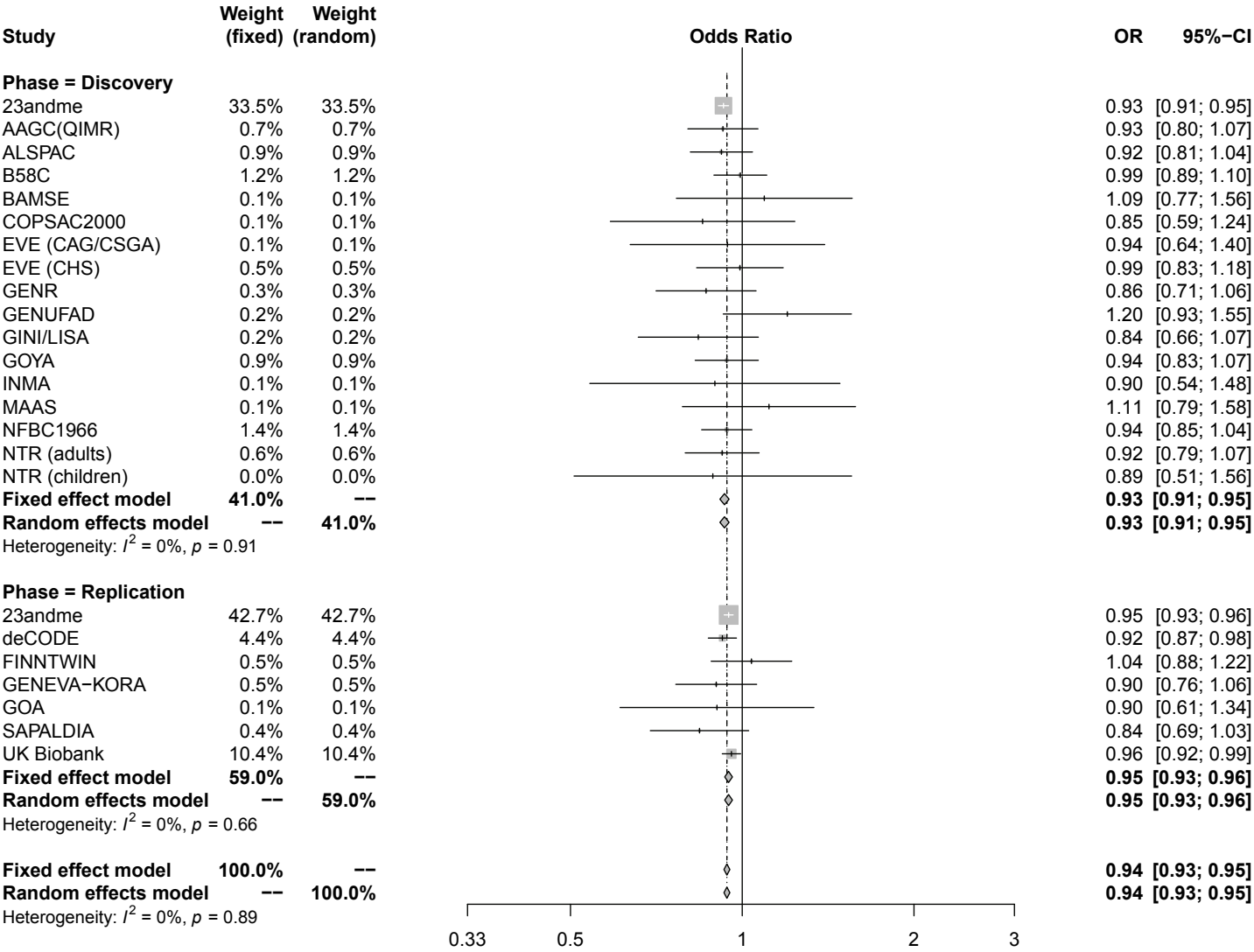
Supplementary Figure 3

Forest plots from the discovery and replication phase for markers not previously associated to allergic rhinitis, including weights, estimates and 95% confidence intervals for inverse variance weighted fixed and random effects meta-analyses from genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals.

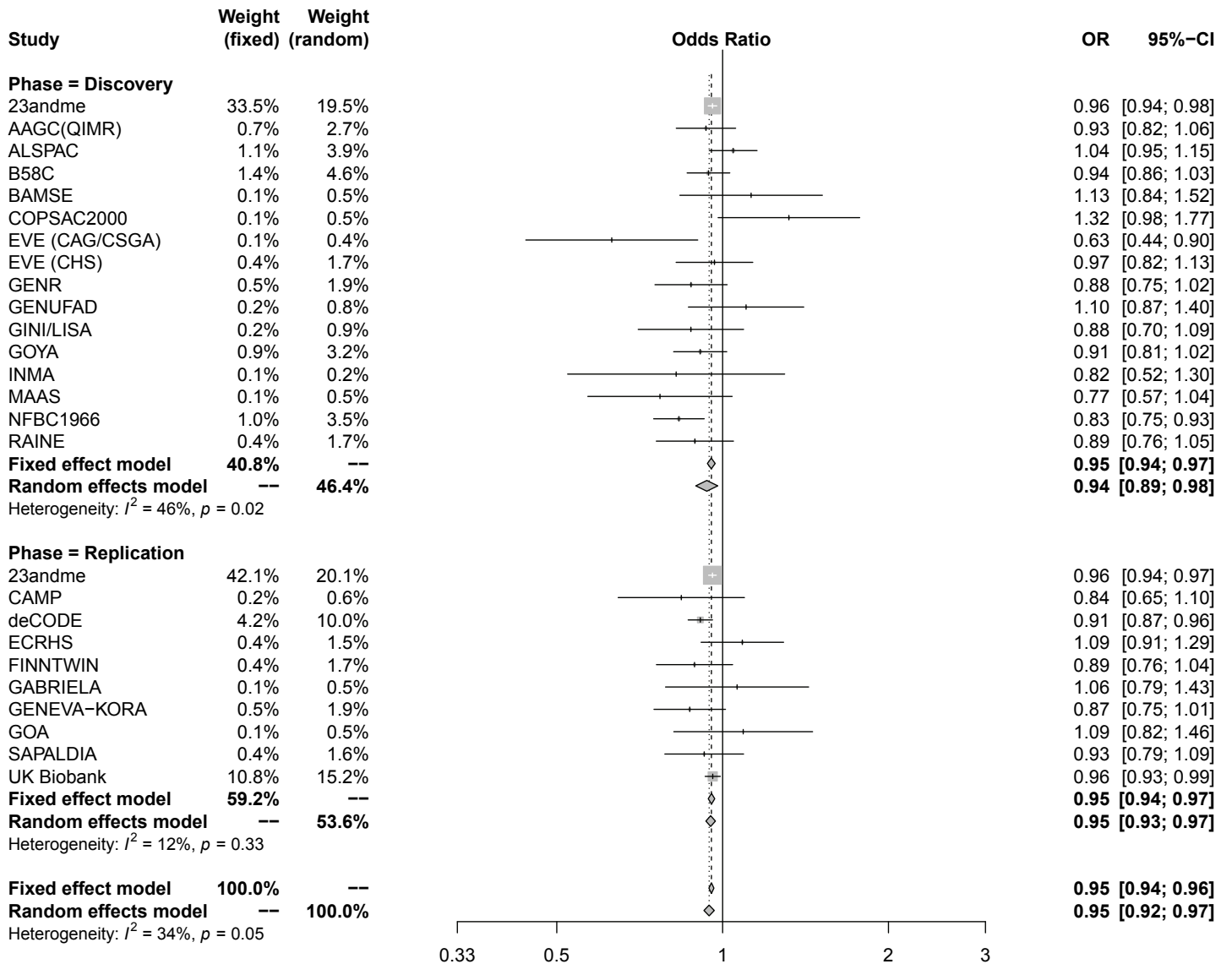
rs7717955[T]



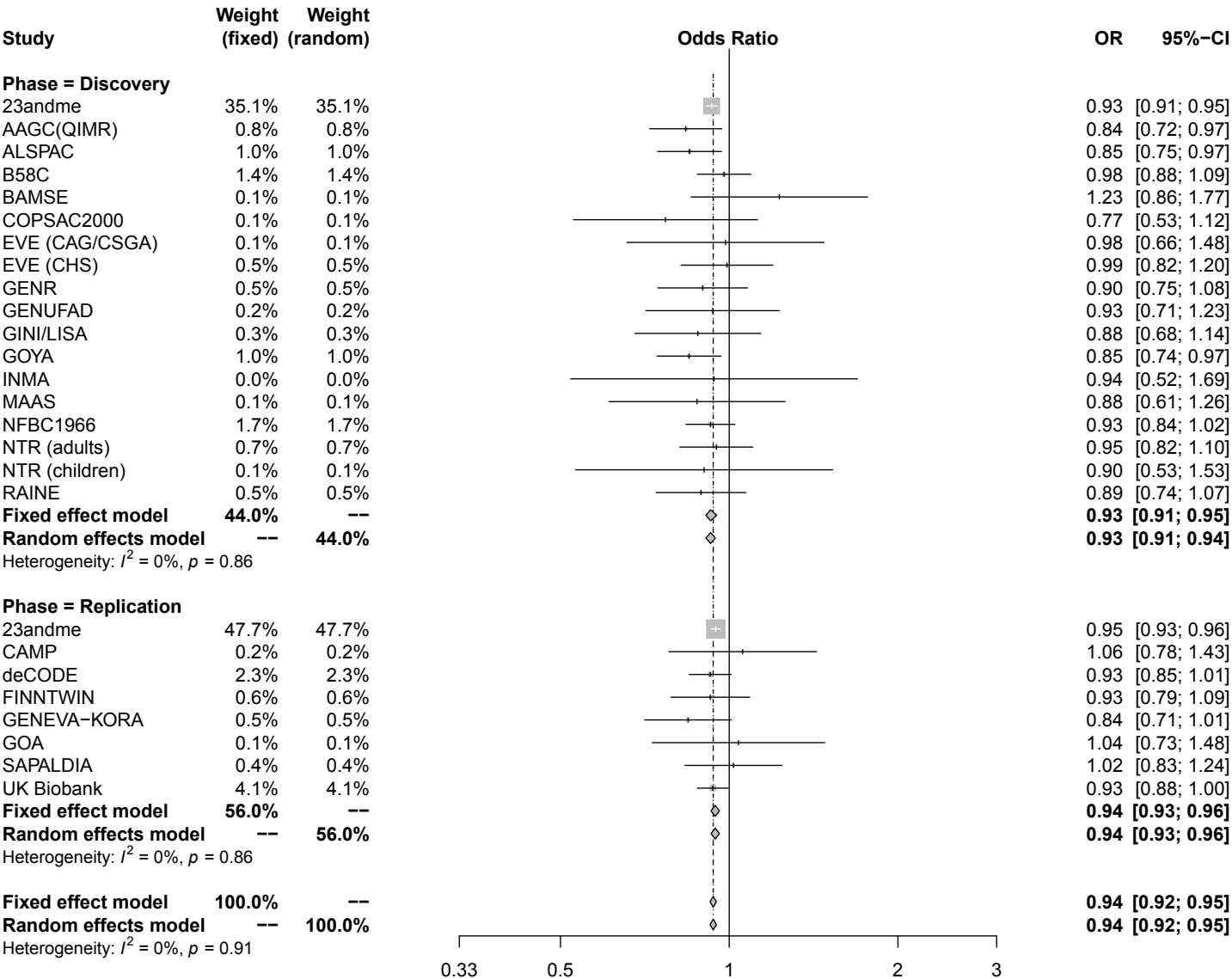
rs63406760[G]



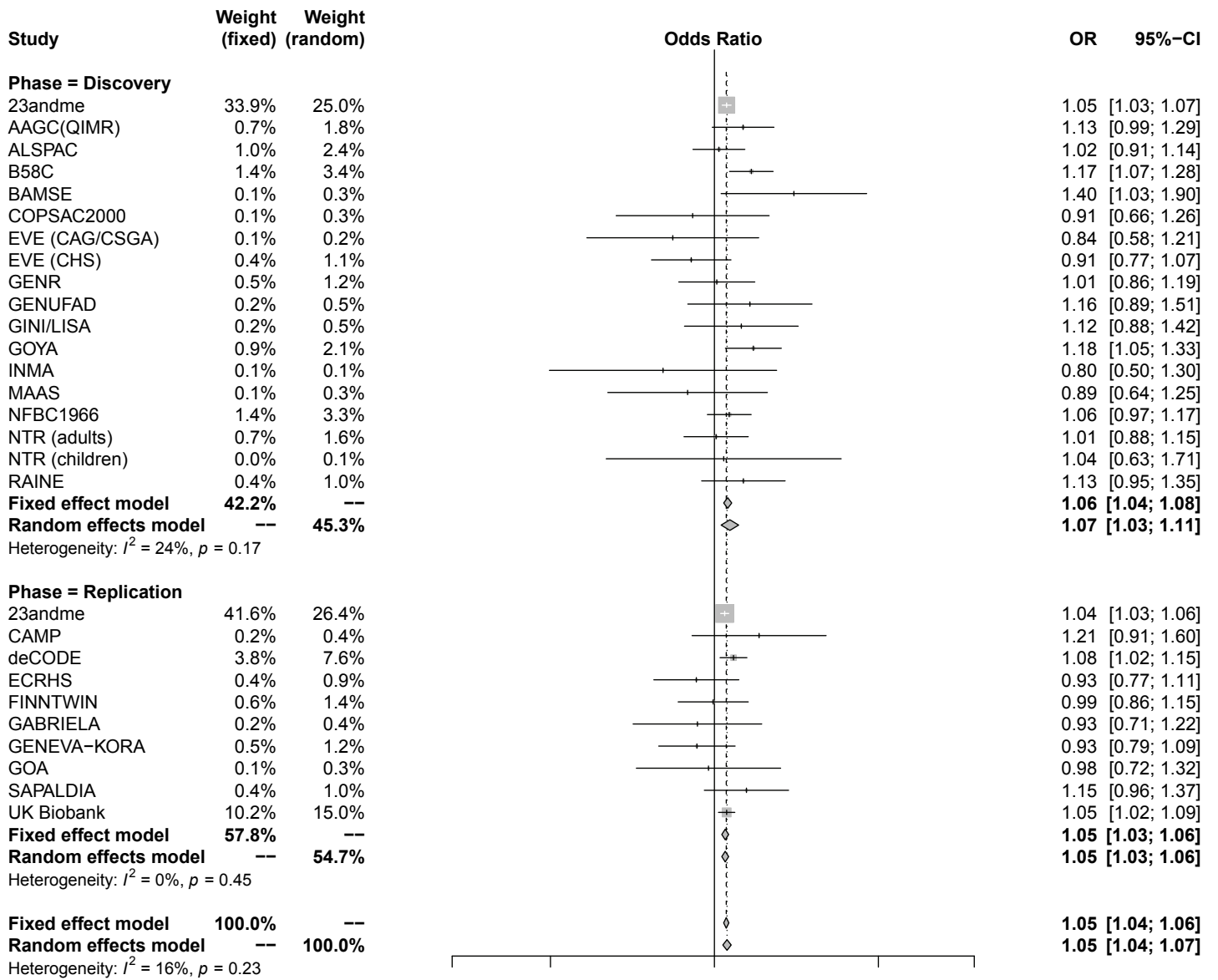
rs1504215[A]



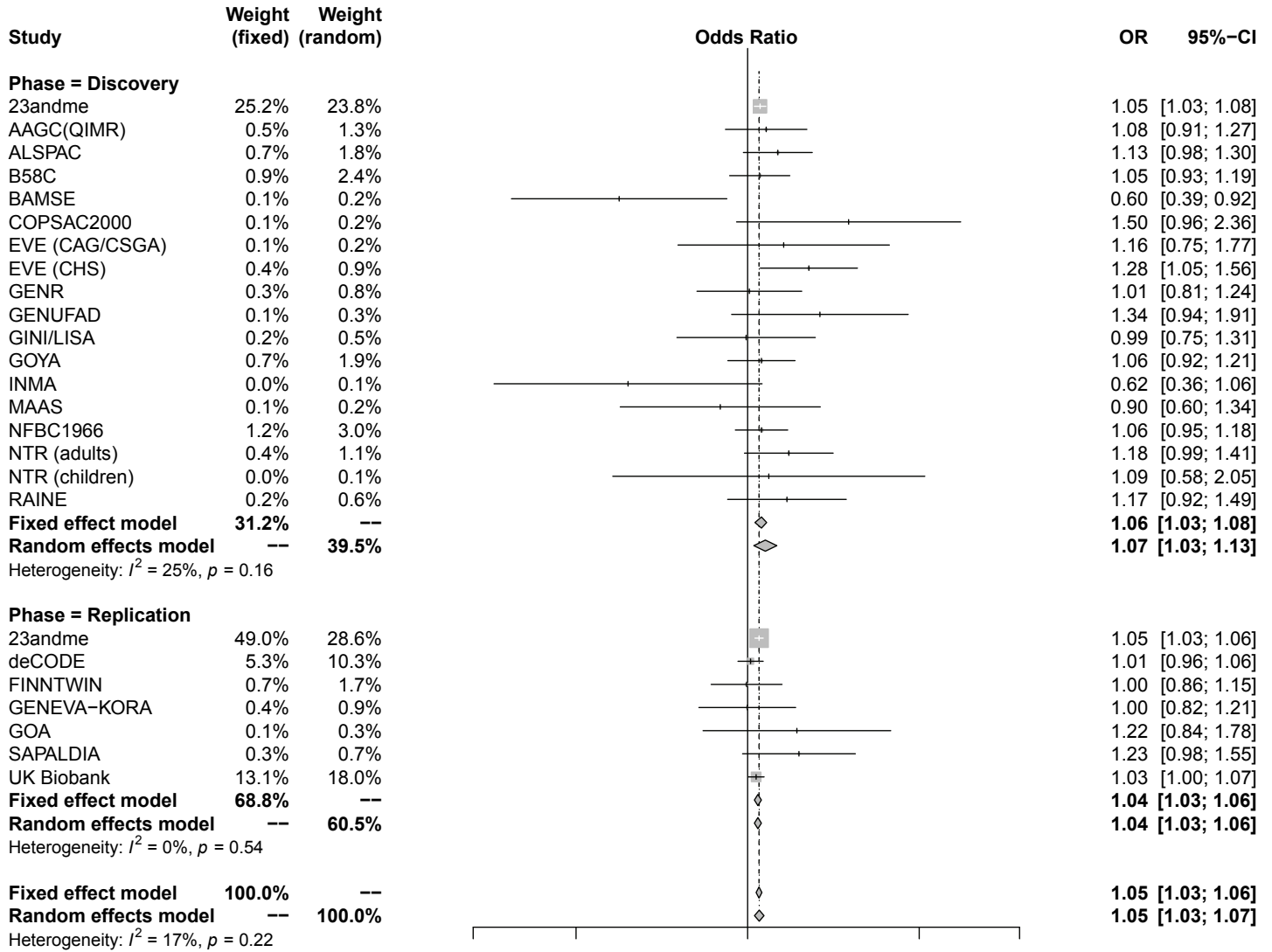
rs28361986[A]



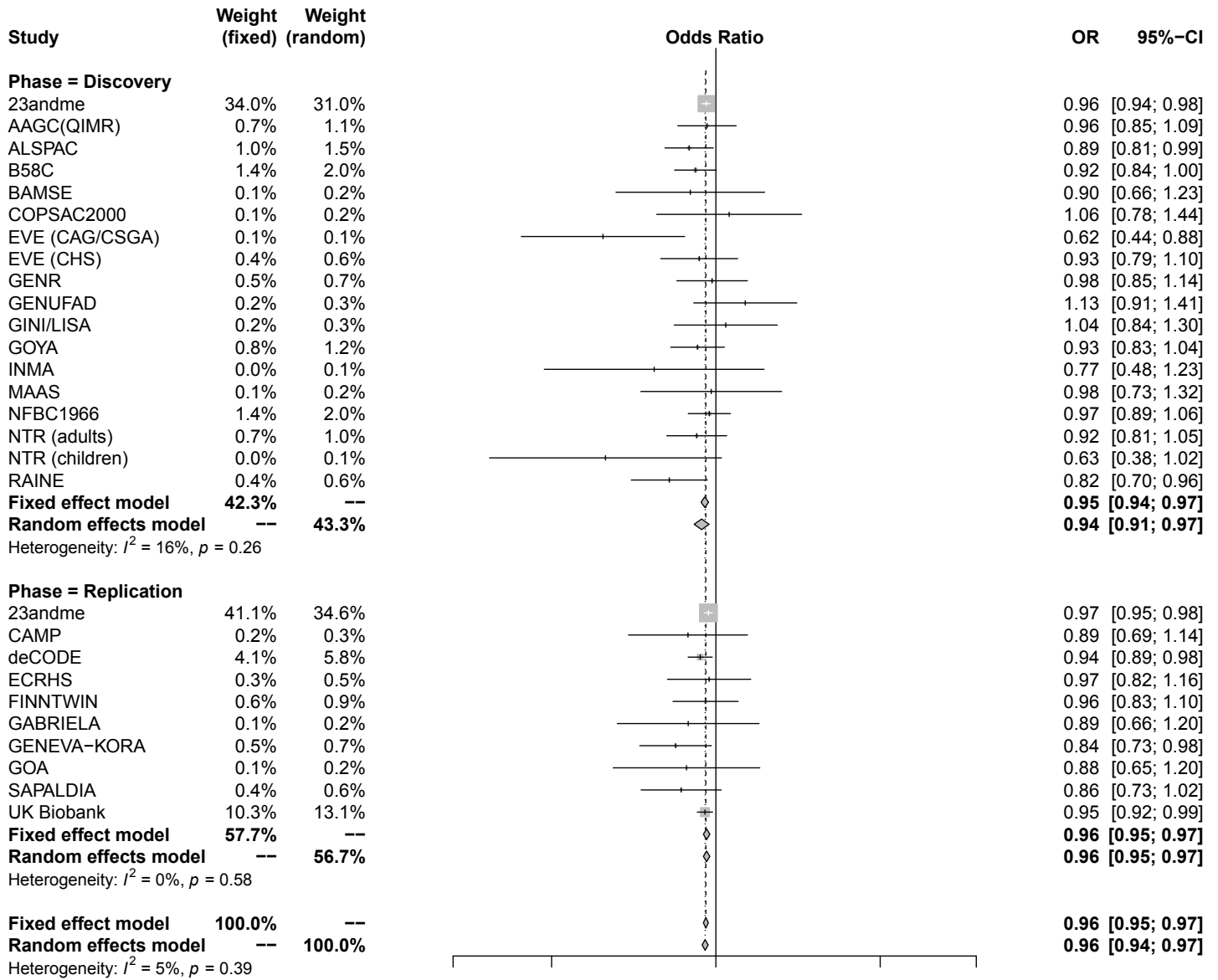
rs2070902[T]



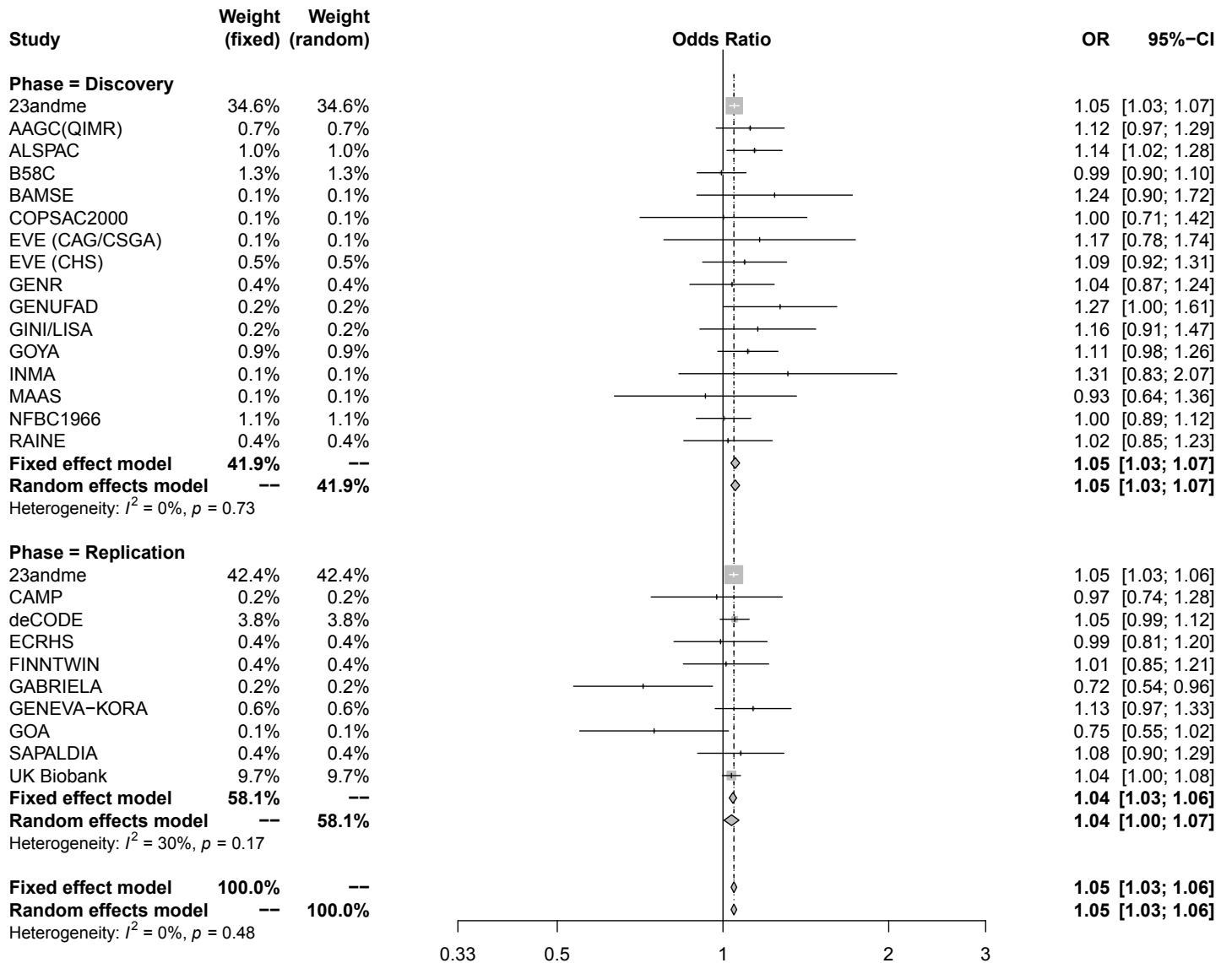
rs111371454[G]



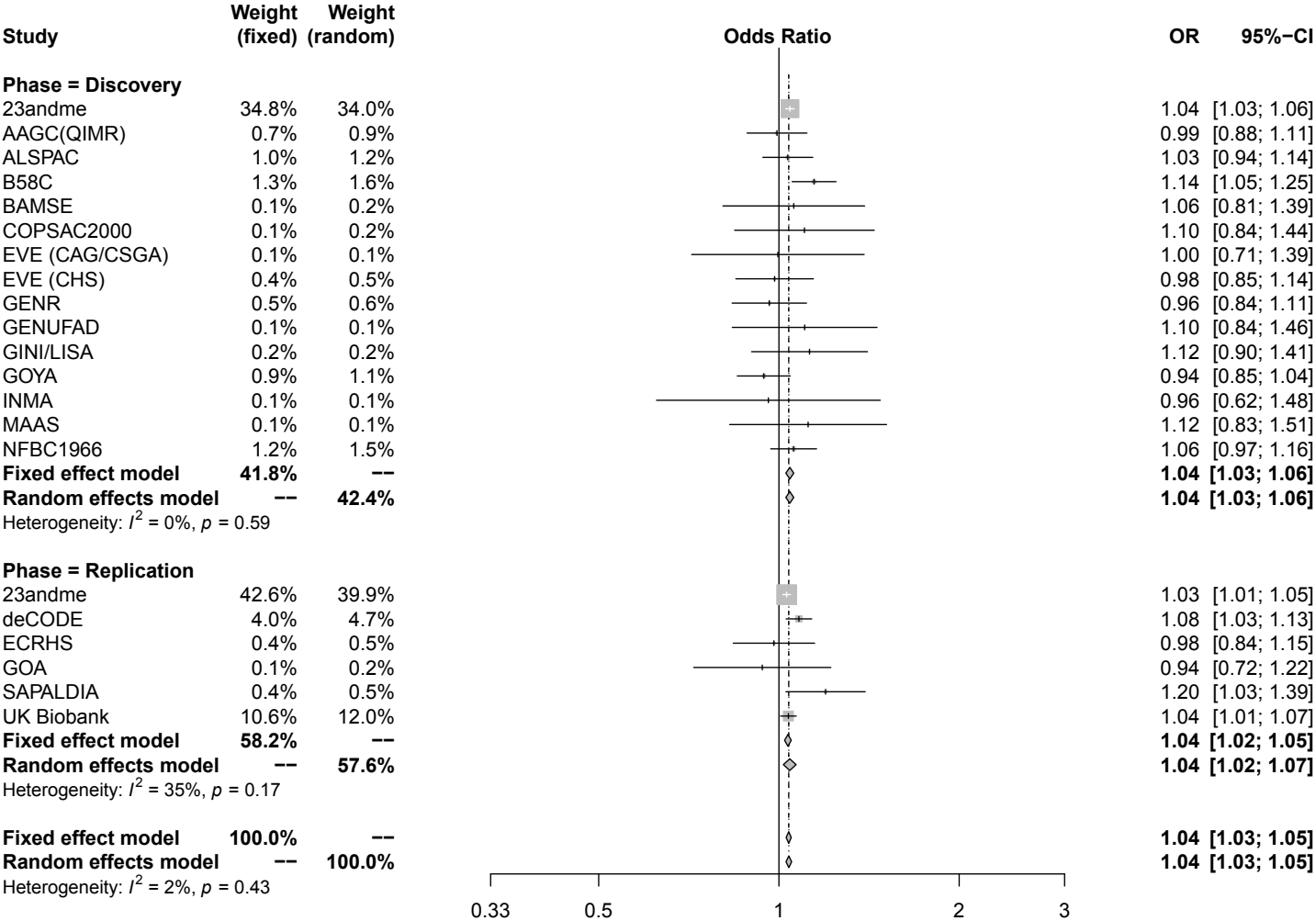
rs12509403[T]



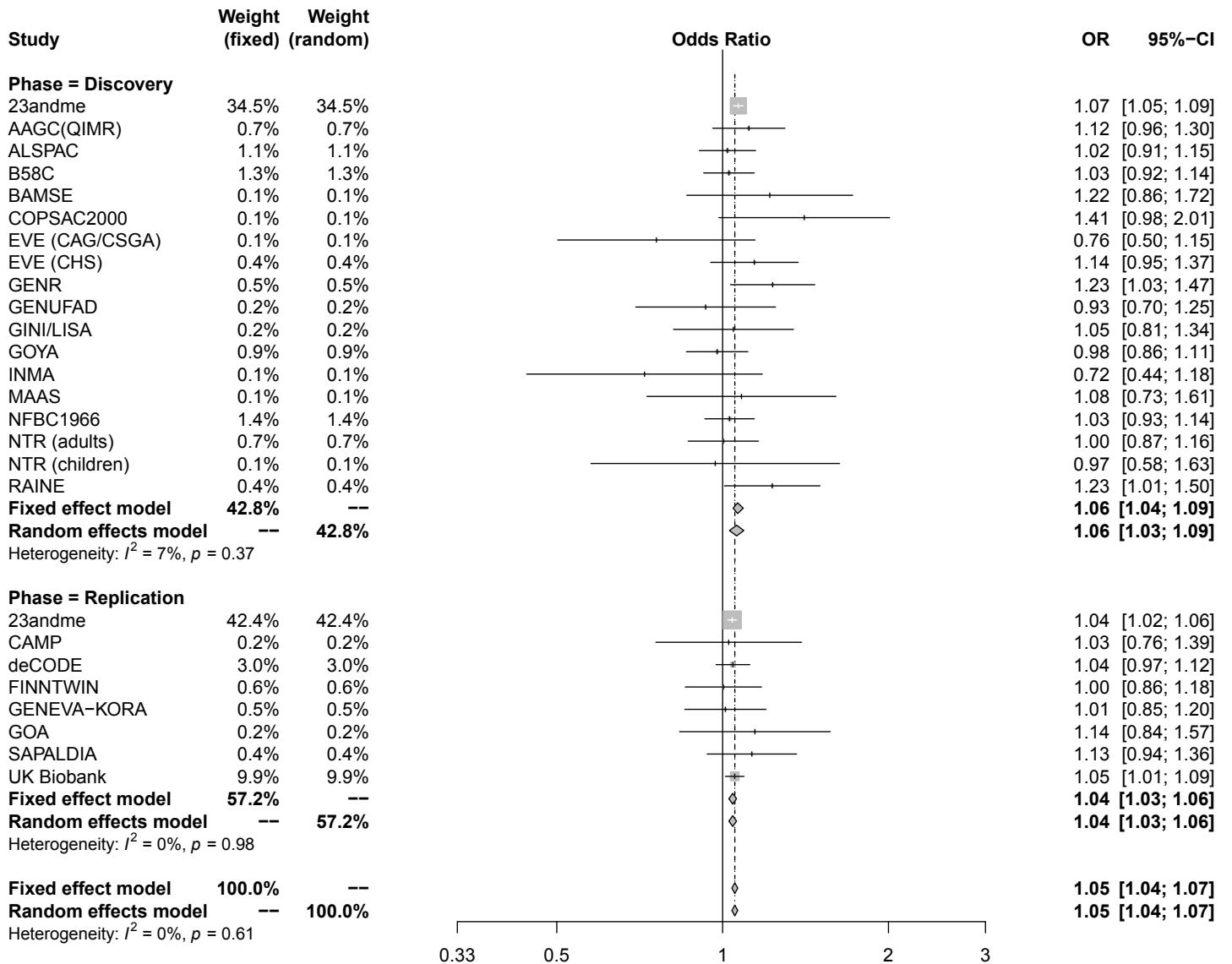
rs9648346[G]



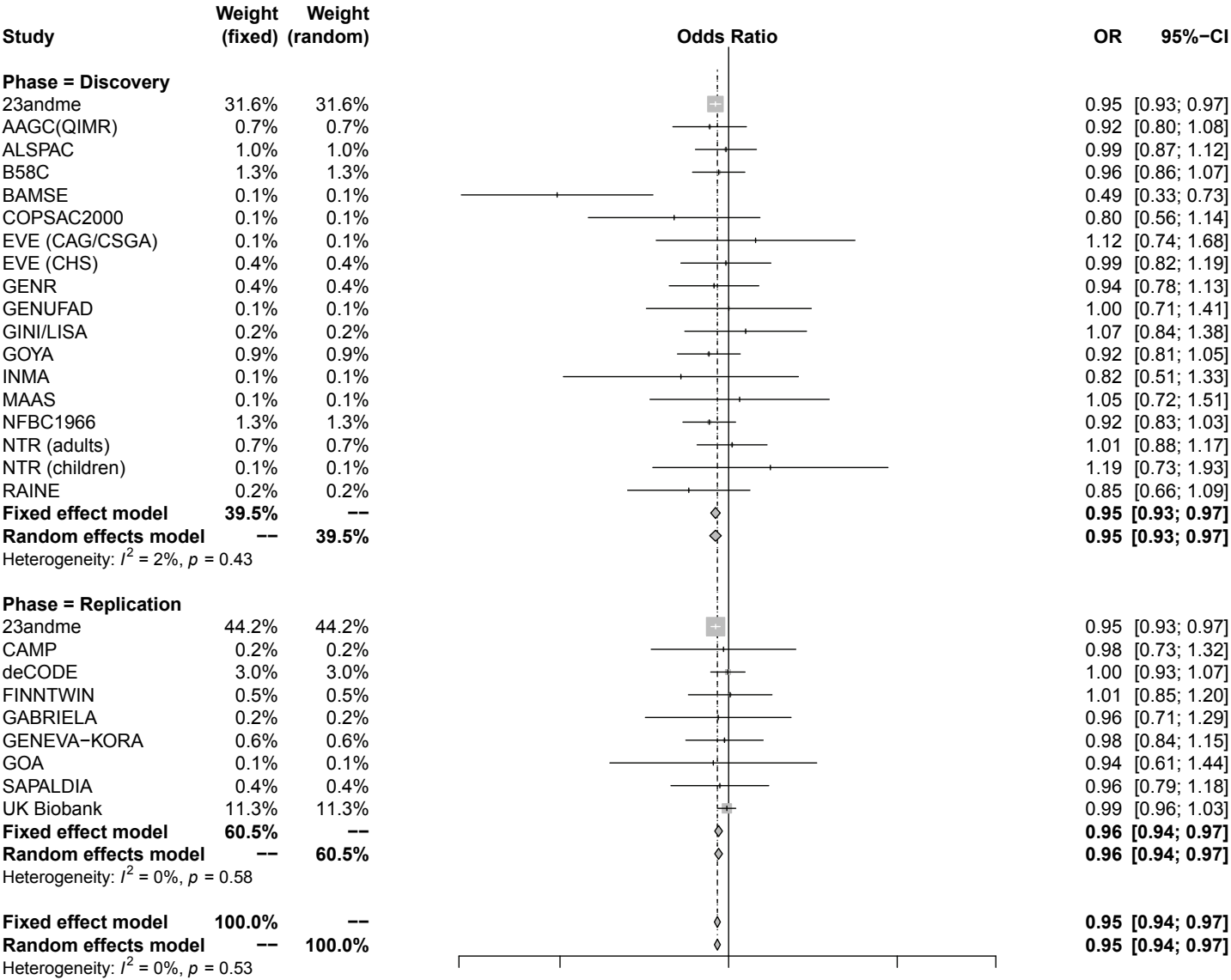
rs35350651[C]



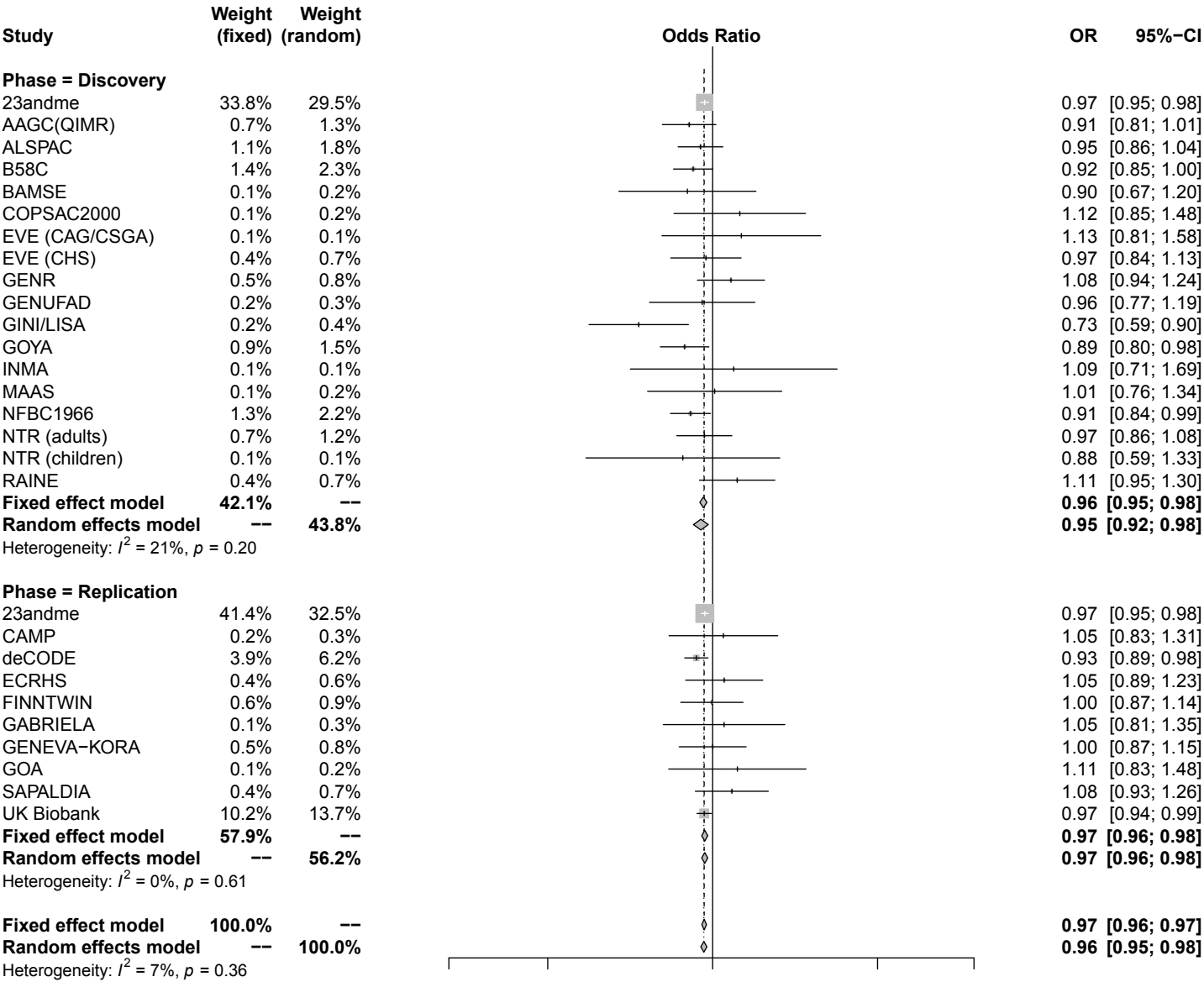
rs2519093[T]



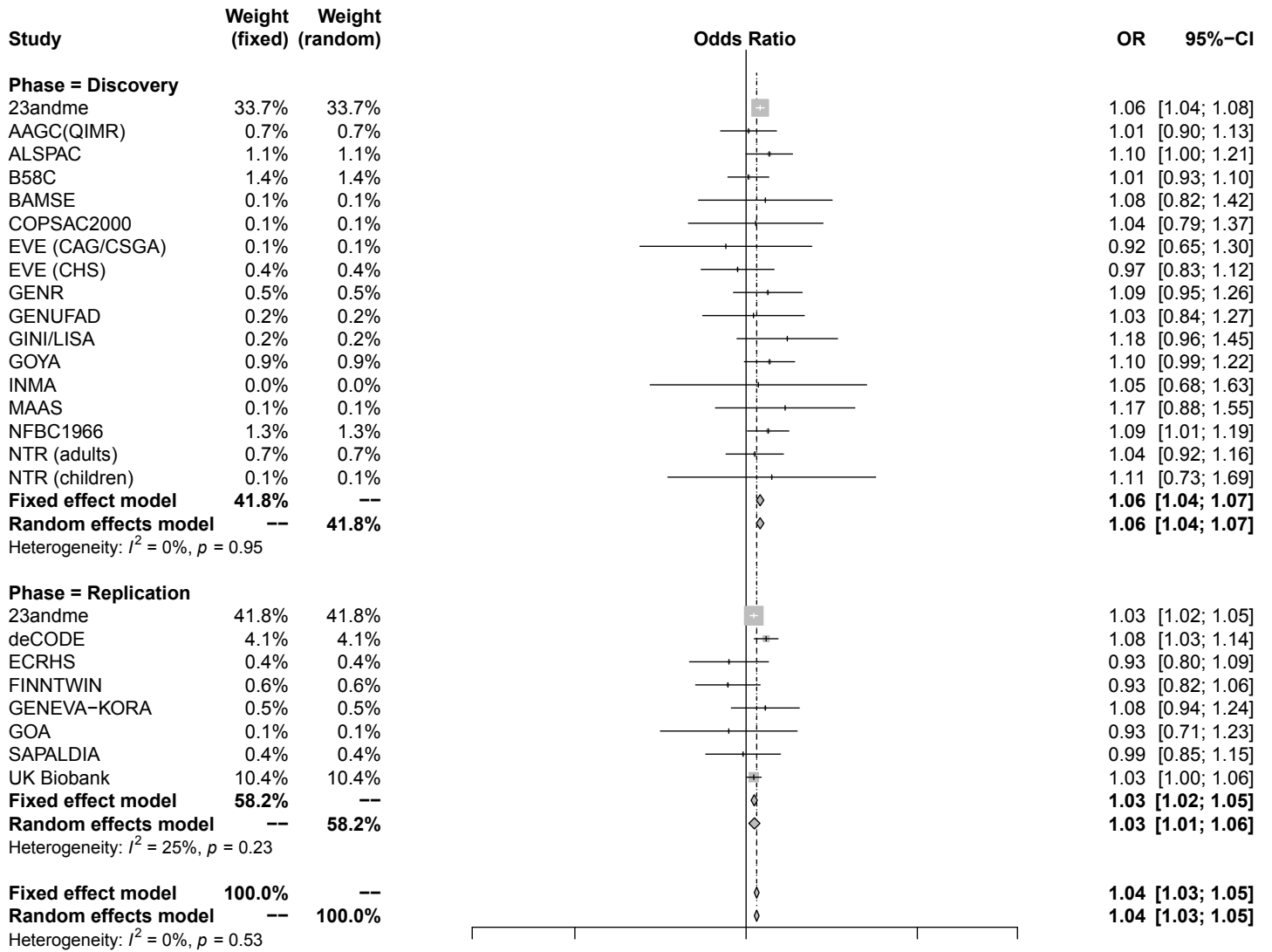
rs62257549[A]



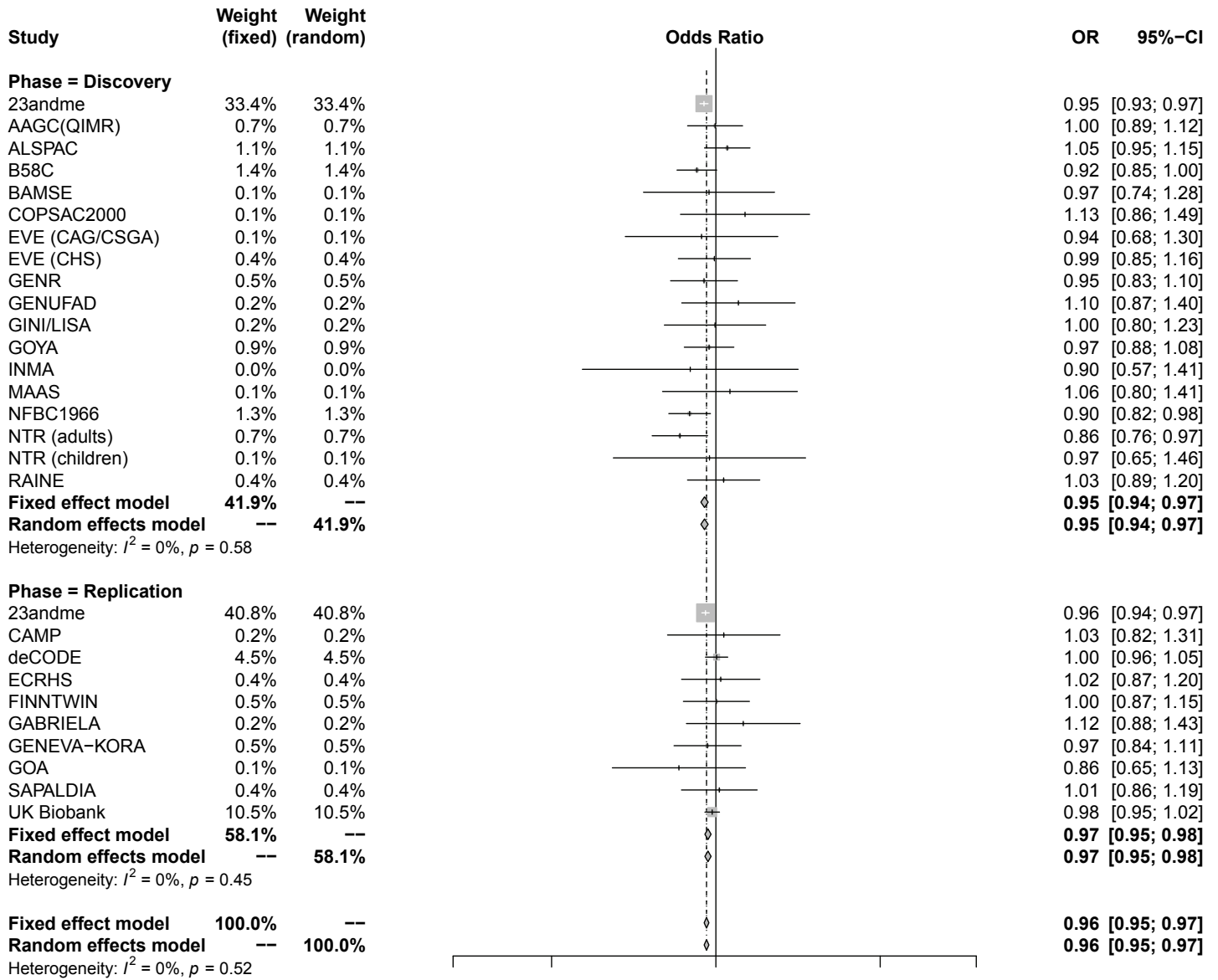
rs11677002[C]



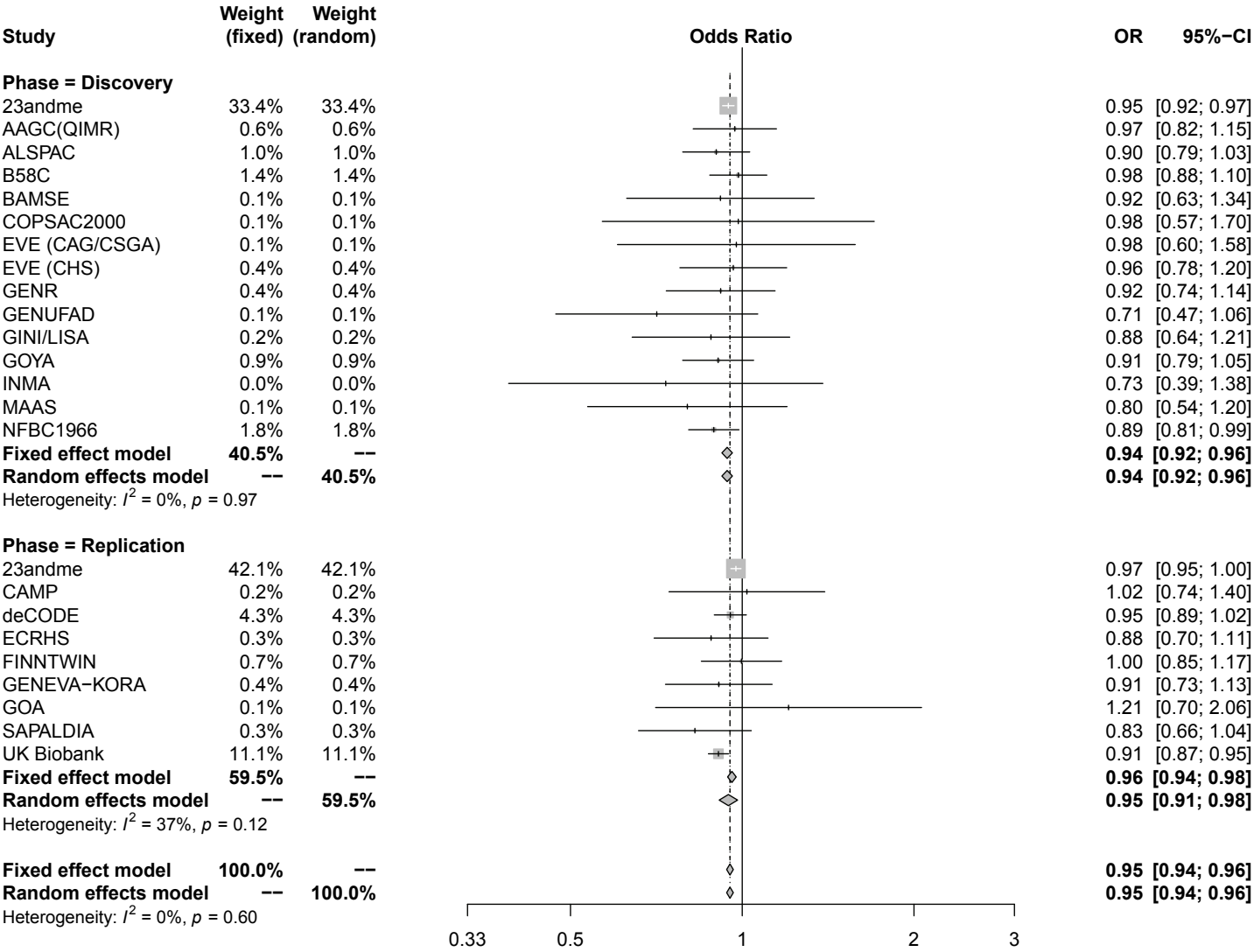
rs35597970[-]



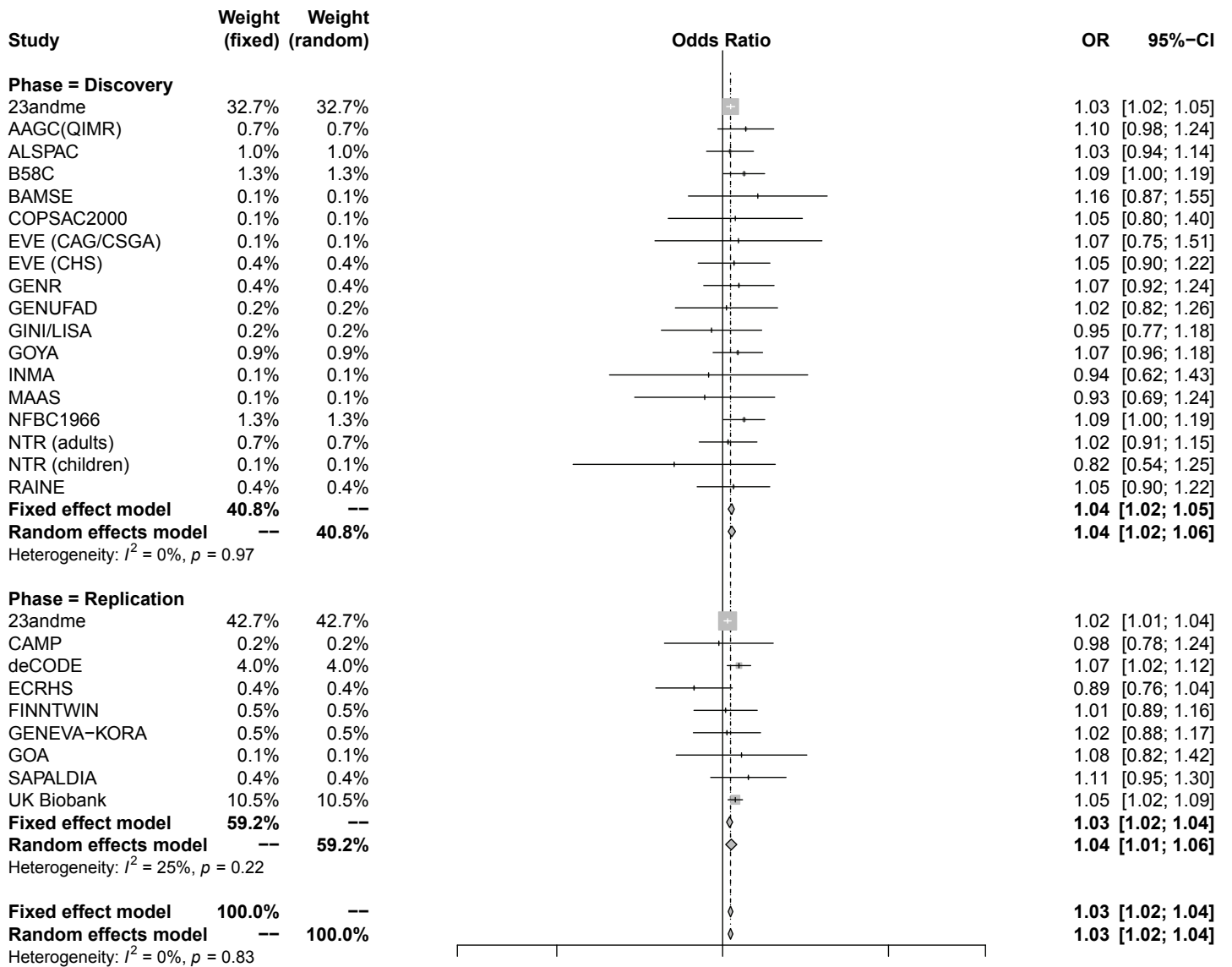
rs2815765[T]



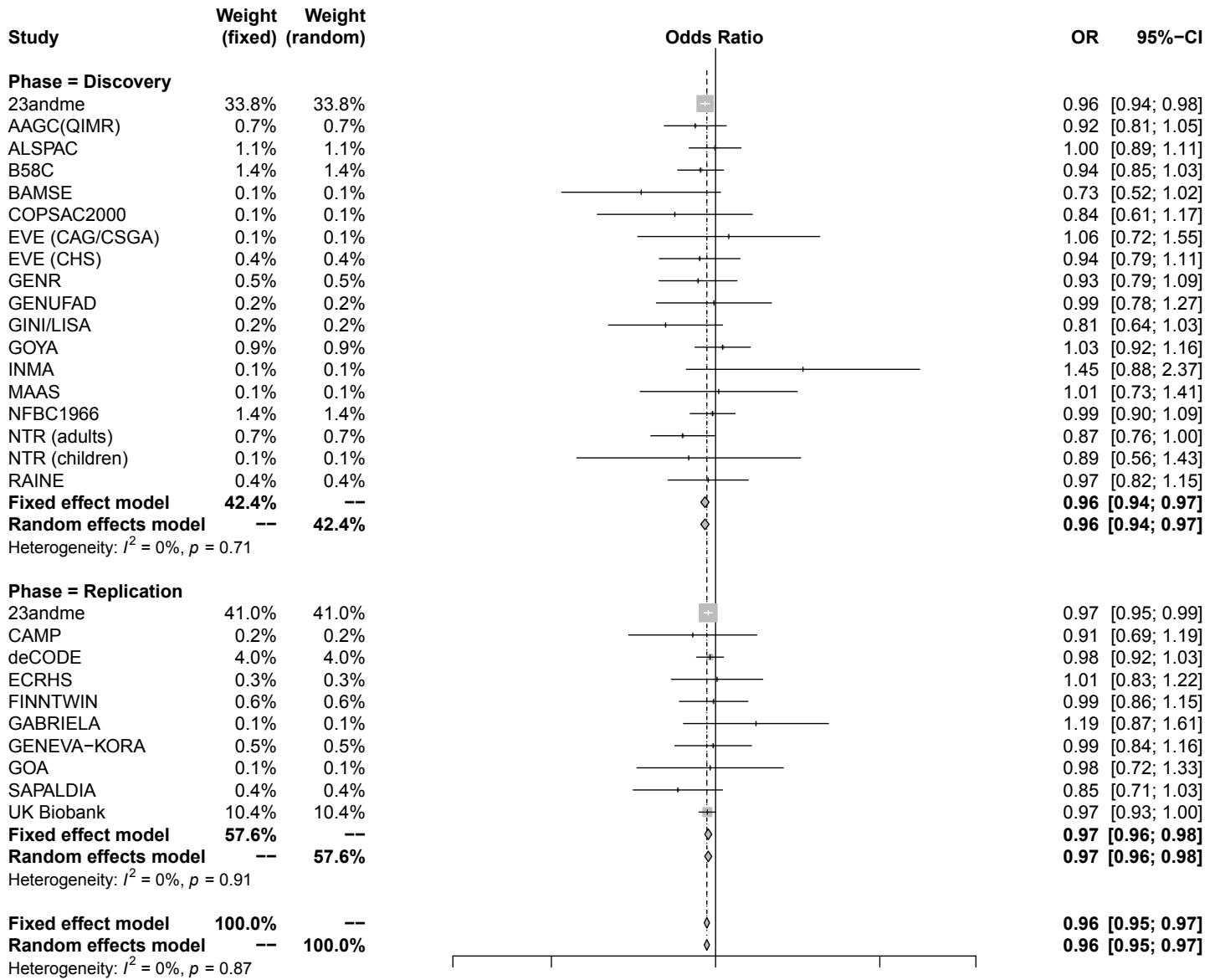
rs11671925[A]



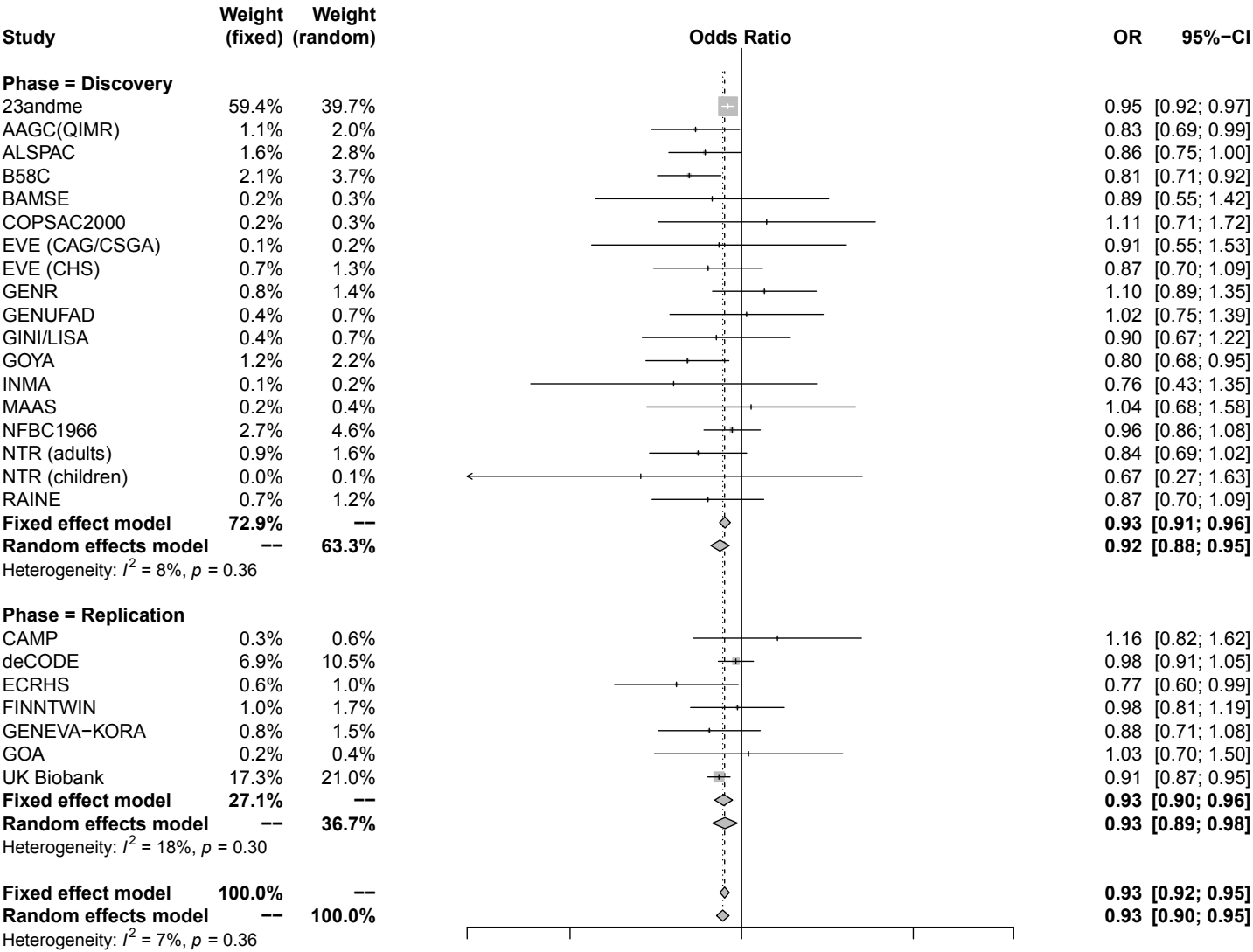
rs2461475[C]



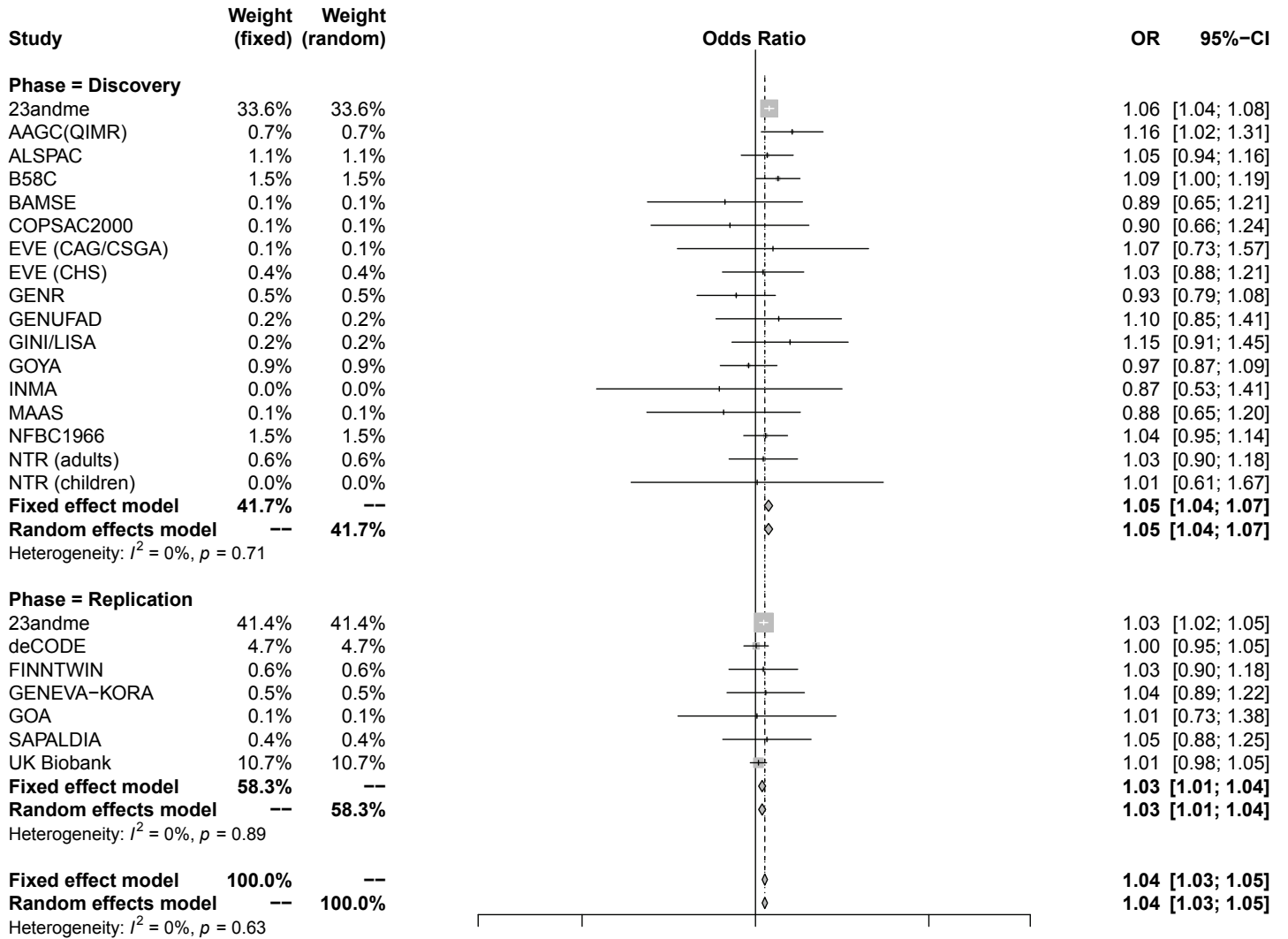
rs6738964[G]



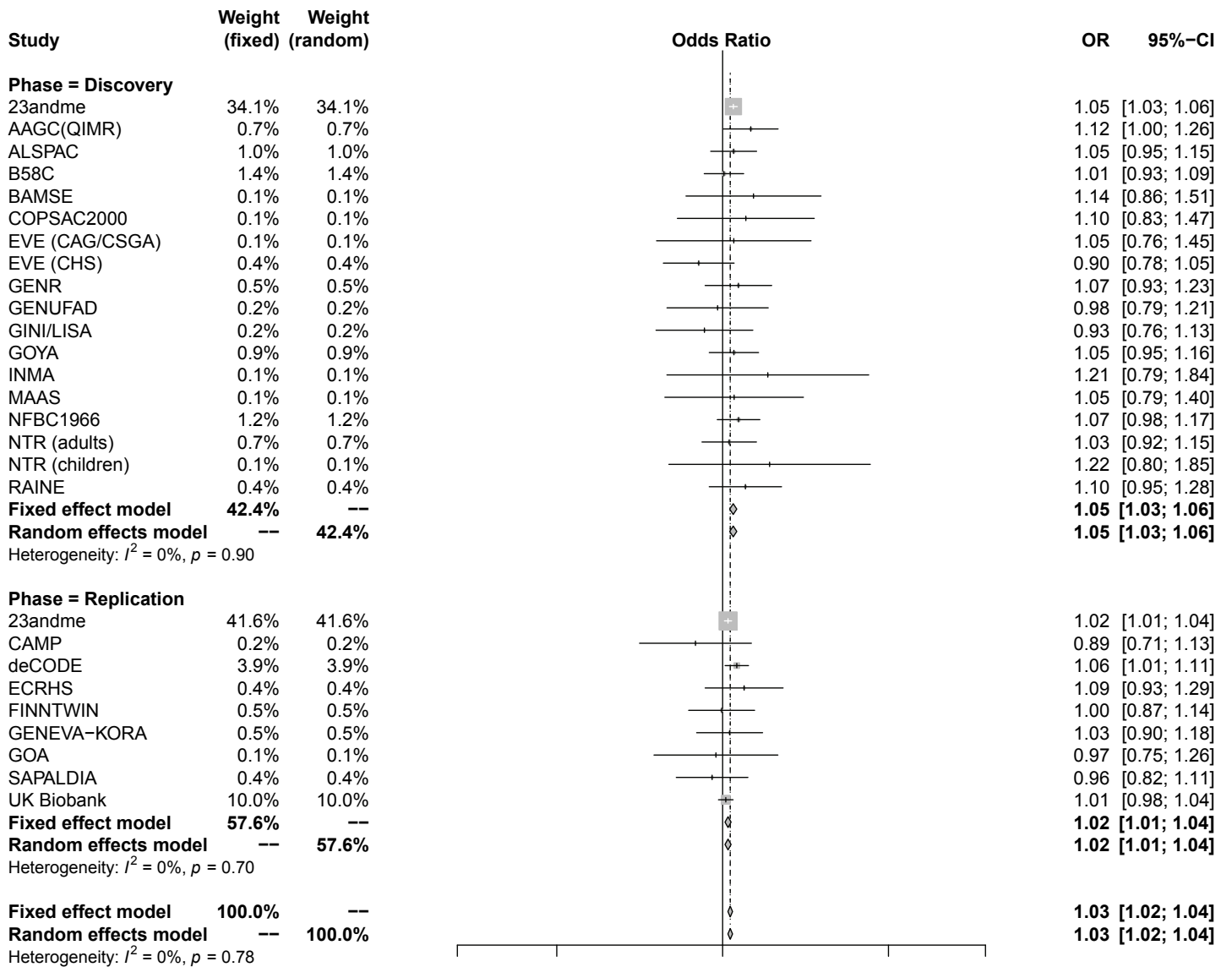
rs10519067[A]



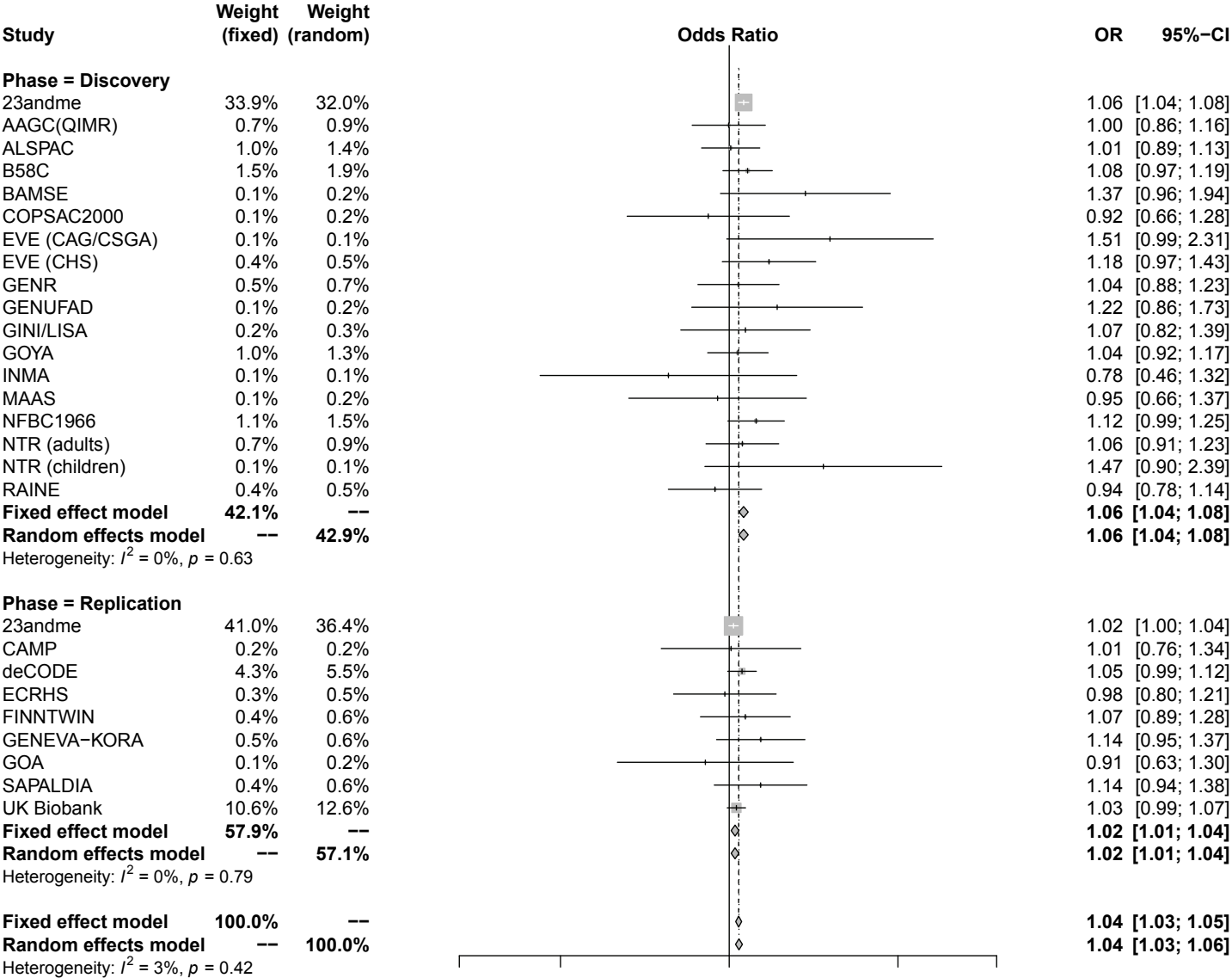
rs138050288[-]



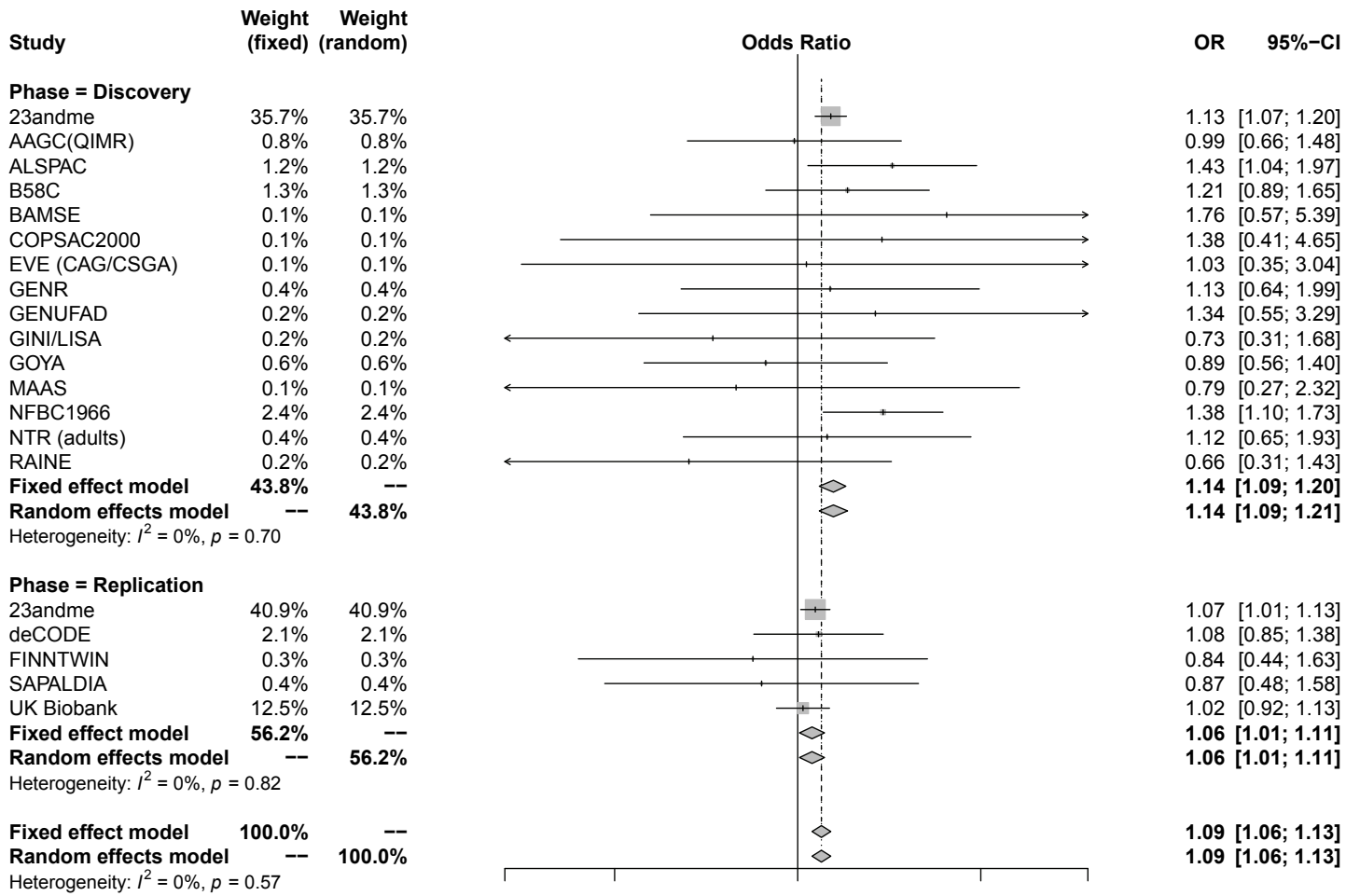
rs7328203[G]



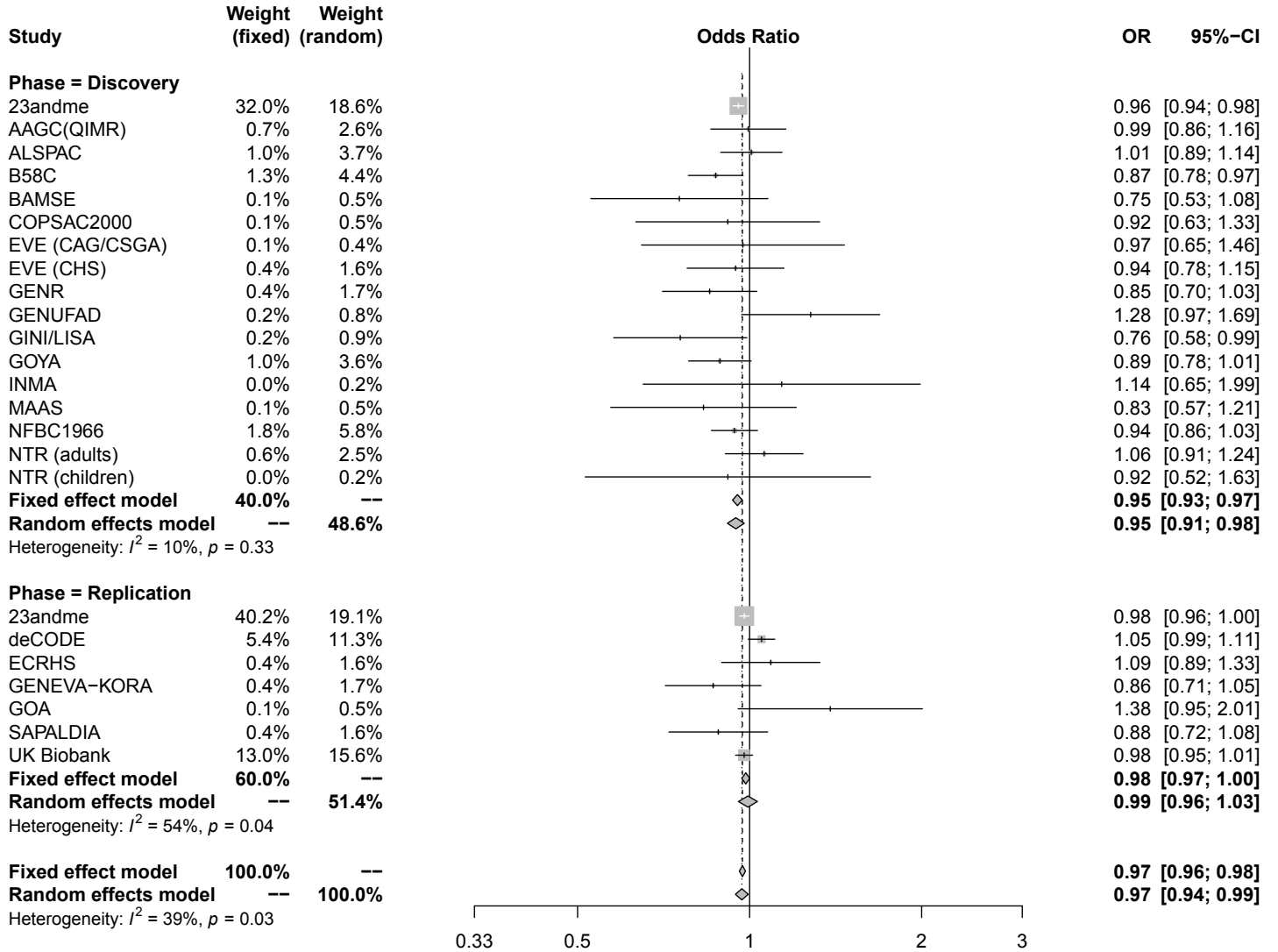
rs11169225[A]



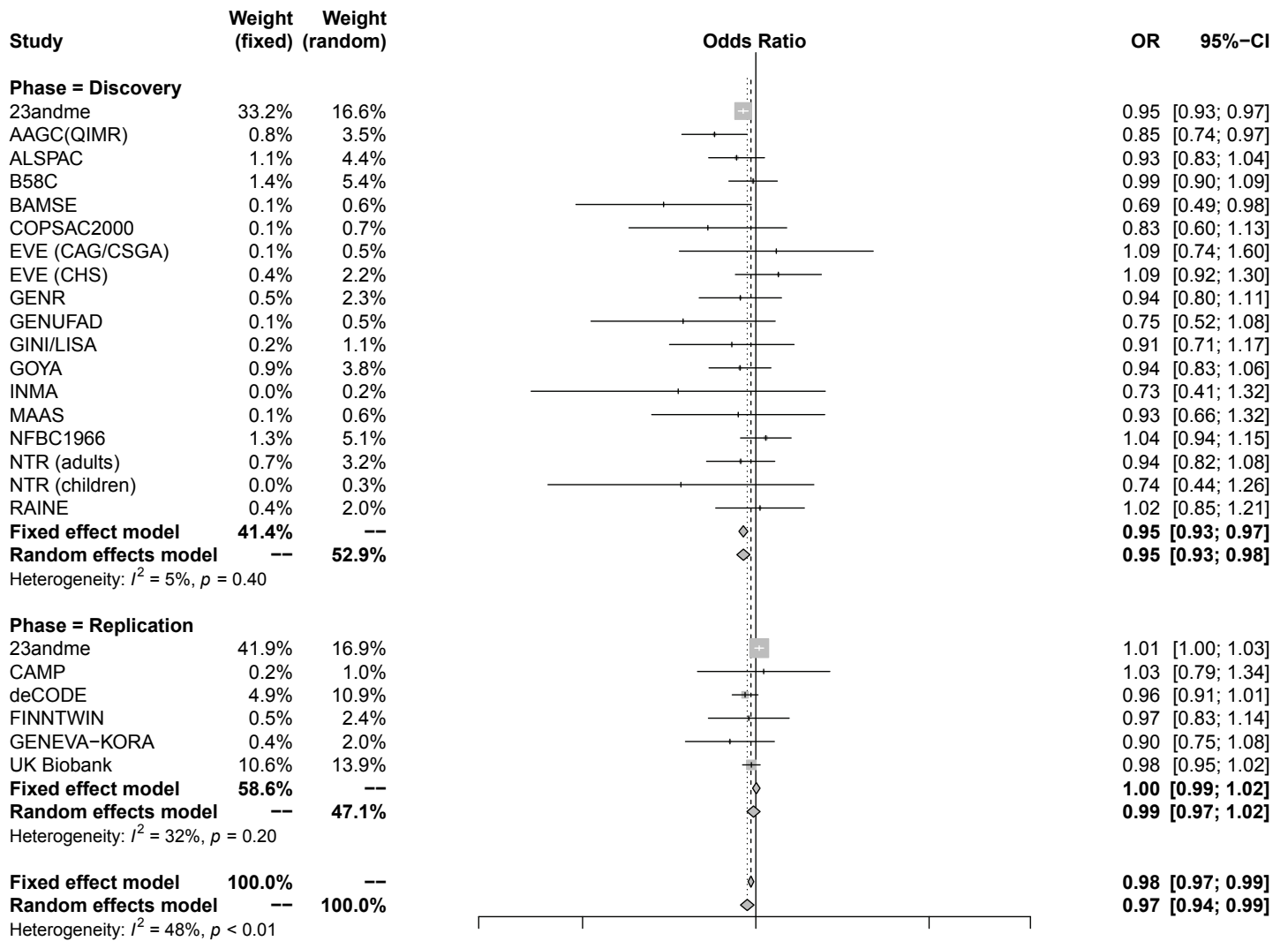
rs141023293[A]



rs201428899[G]

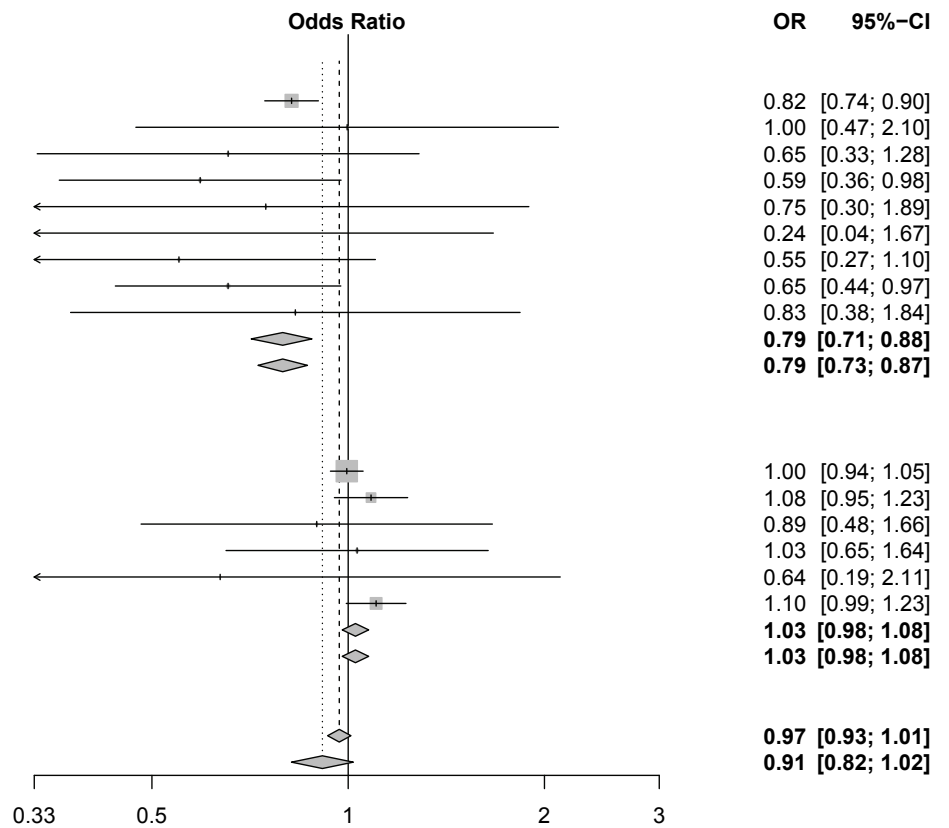


rs6011016[C]

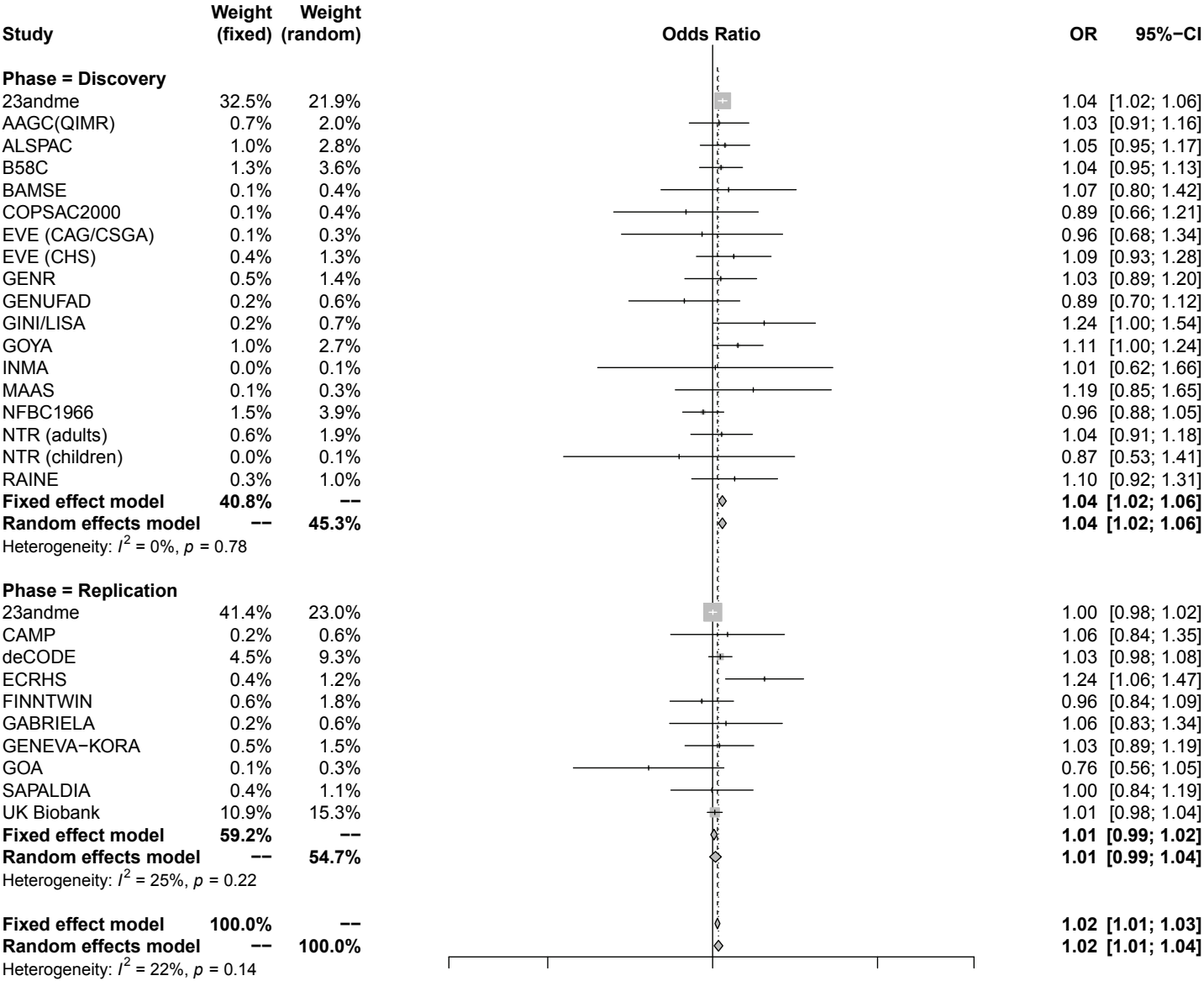


rs149341190[C]

Study	Weight (fixed)	Weight (random)
Phase = Discovery		
23andme	18.9%	18.4%
AAGC(QIMR)	0.3%	1.9%
ALSPAC	0.4%	2.3%
B58C	0.7%	3.9%
GENR	0.2%	1.3%
GENUFAD	0.0%	0.3%
GOYA	0.4%	2.2%
NFBC1966	1.1%	5.6%
RAINE	0.3%	1.7%
Fixed effect model	22.2%	--
Random effects model	--	37.7%
Heterogeneity: $I^2 = 0\%$, $p = 0.66$		
Phase = Replication		
23andme	51.0%	20.1%
deCODE	10.1%	16.4%
FINNTWIN	0.4%	2.7%
GENEVA-KORA	0.8%	4.4%
SAPALDIA	0.1%	0.8%
UK Biobank	15.3%	17.8%
Fixed effect model	77.8%	--
Random effects model	--	62.3%
Heterogeneity: $I^2 = 0\%$, $p = 0.49$		
Fixed effect model	100.0%	--
Random effects model	--	100.0%
Heterogeneity: $I^2 = 61\%$, $p < 0.01$		

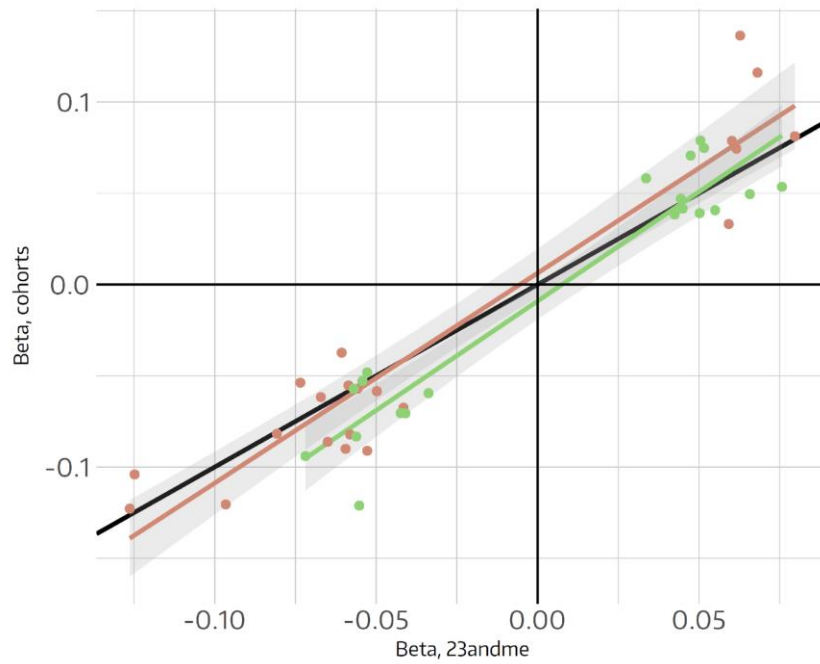


rs10760123[T]



Supplementary Figure 4

Scatter plot of beta coefficients between 23andme and non-23andme cohorts from the discovery phase of the allergic rhinitis GWAS, from genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals. Only known and replicating novel markers are shown. Red indicates known loci, green indicates novel, replicating loci. The black line represents identity.

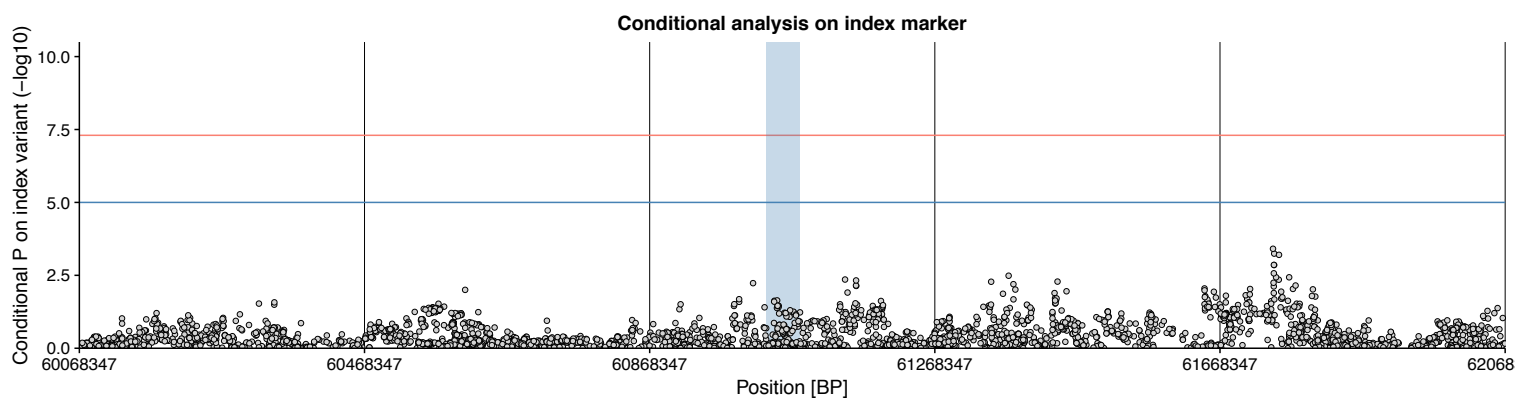
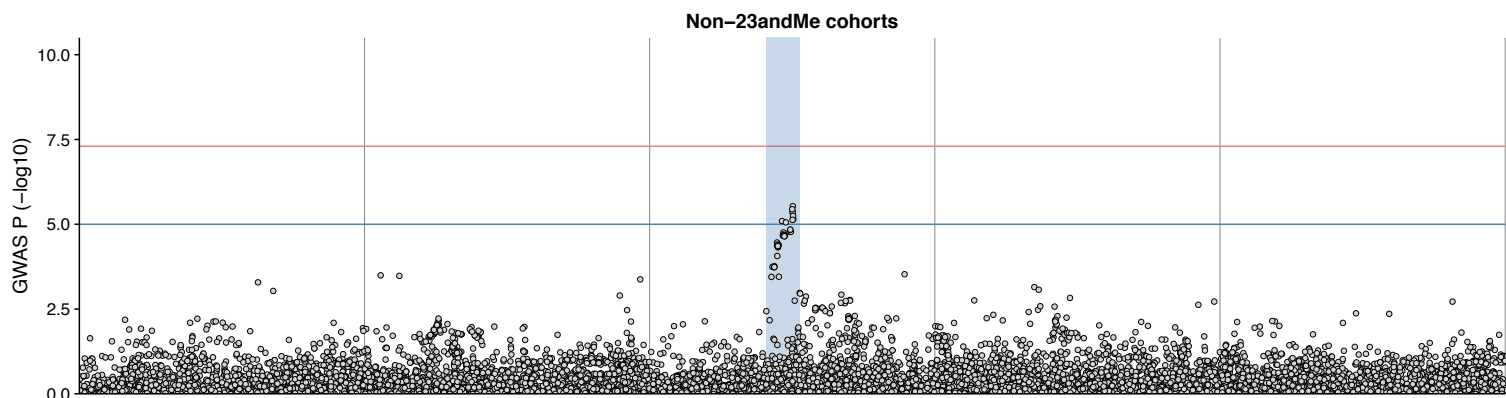
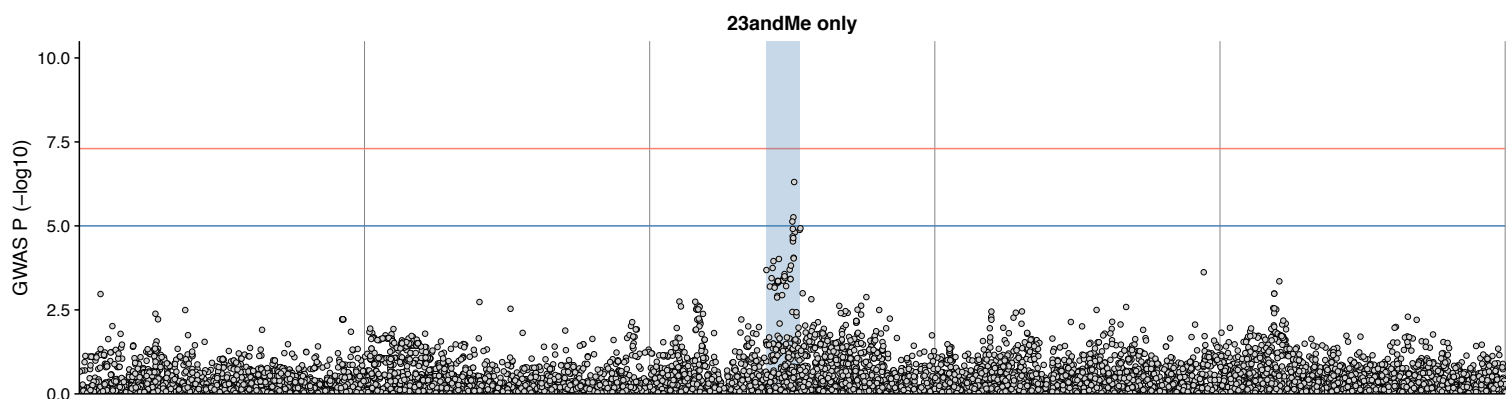
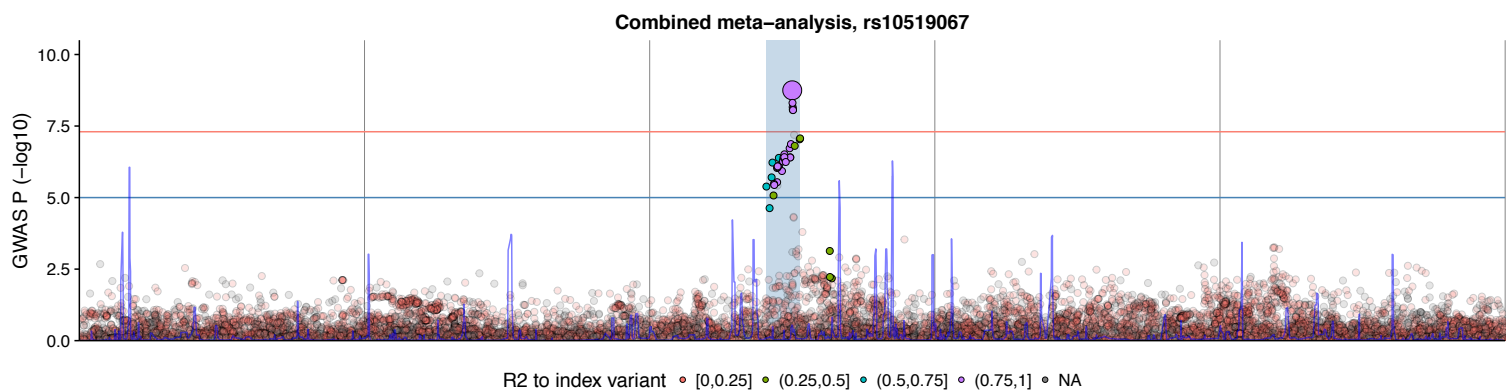
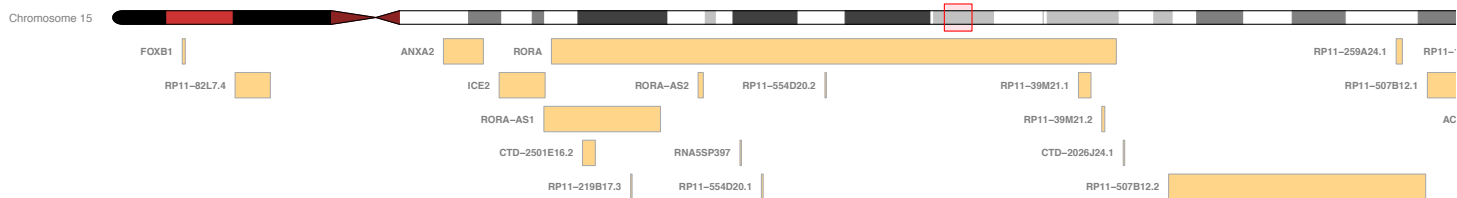


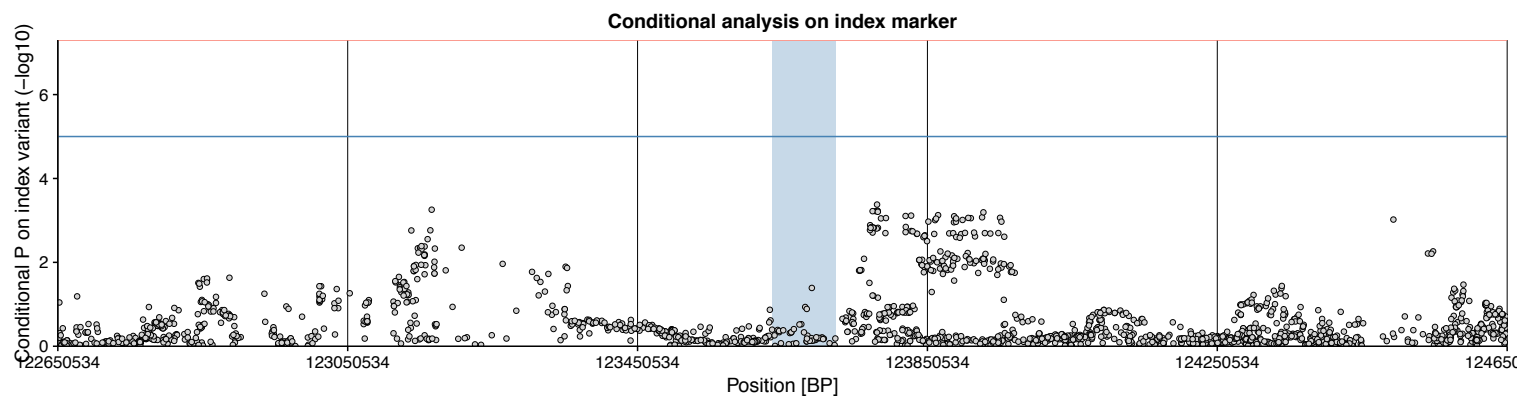
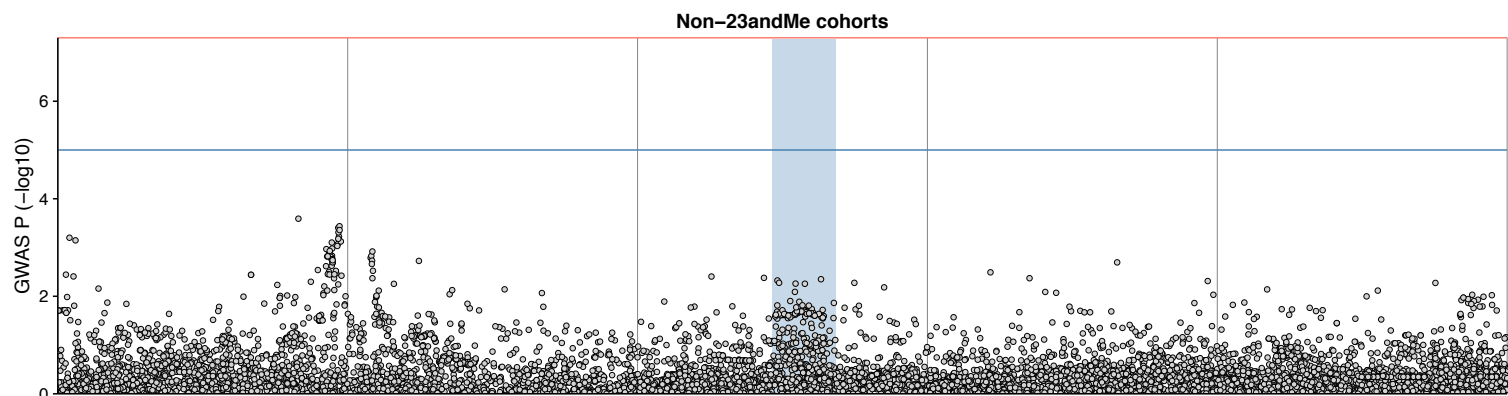
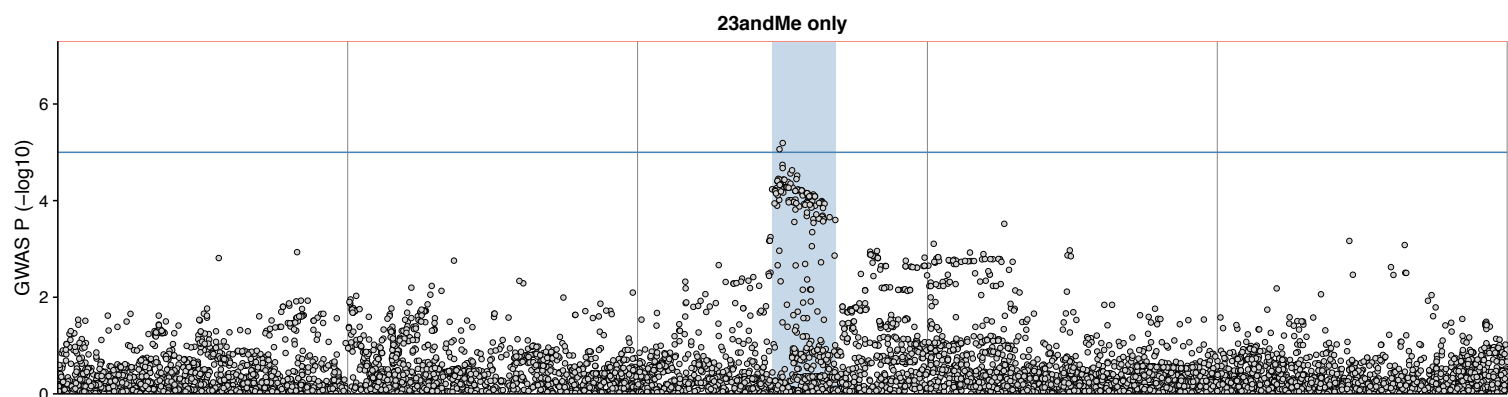
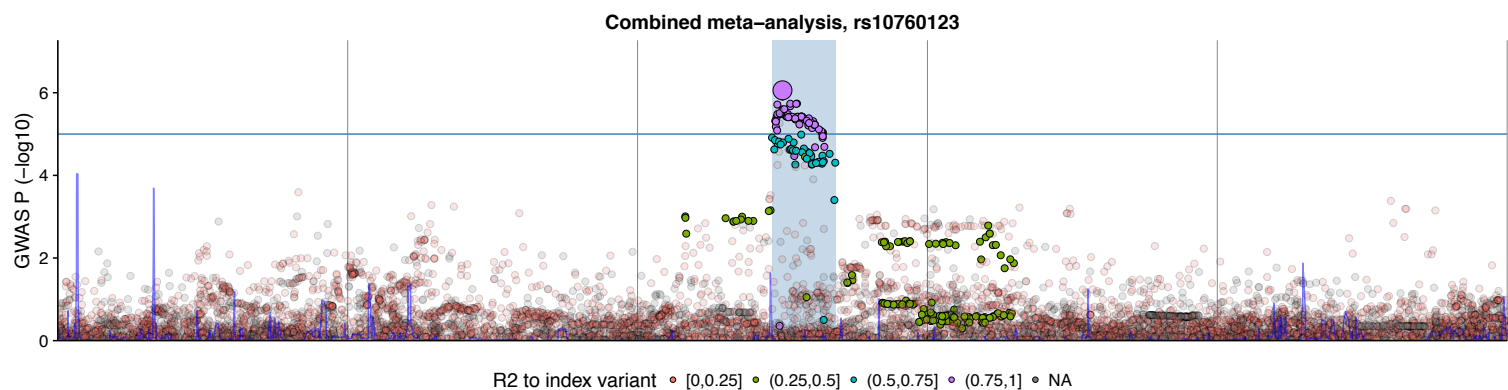
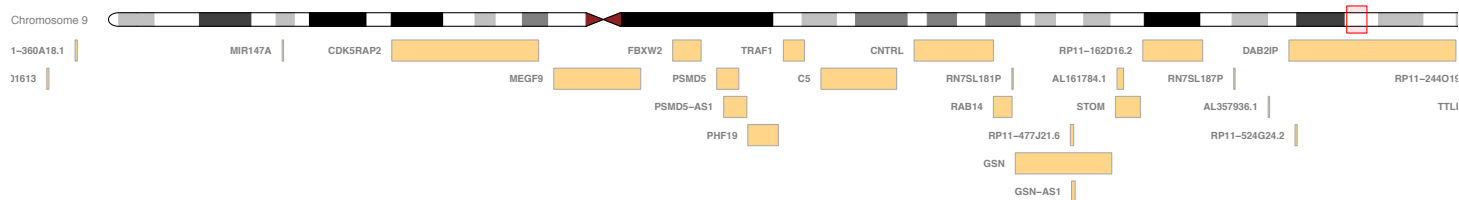
Supplementary Figure 5

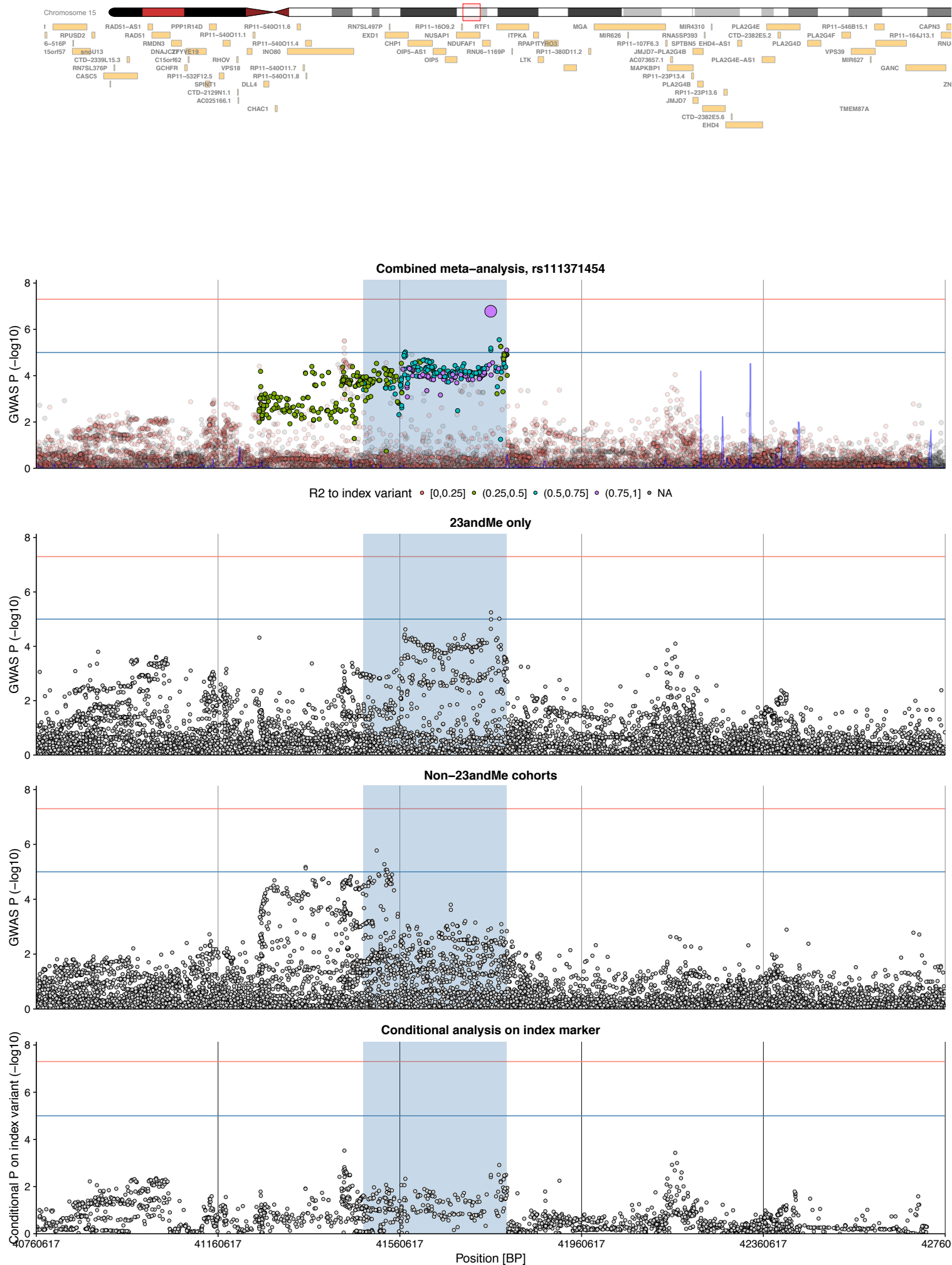
Regional locus plots of allergic rhinitis discovery phase markers. Top panel shows results from the discovery meta-analysis, color coded for 1000g R^2 values. AR discovery markers not present in 1000g are colored in grey ("NA"). Index marker has increased circle size. Locus definition (tag $R^2 > 0.5$ to index) is indicated by blue region. Red line indicates $p = 5e-8$, blue line indicates $p = 1e-5$.

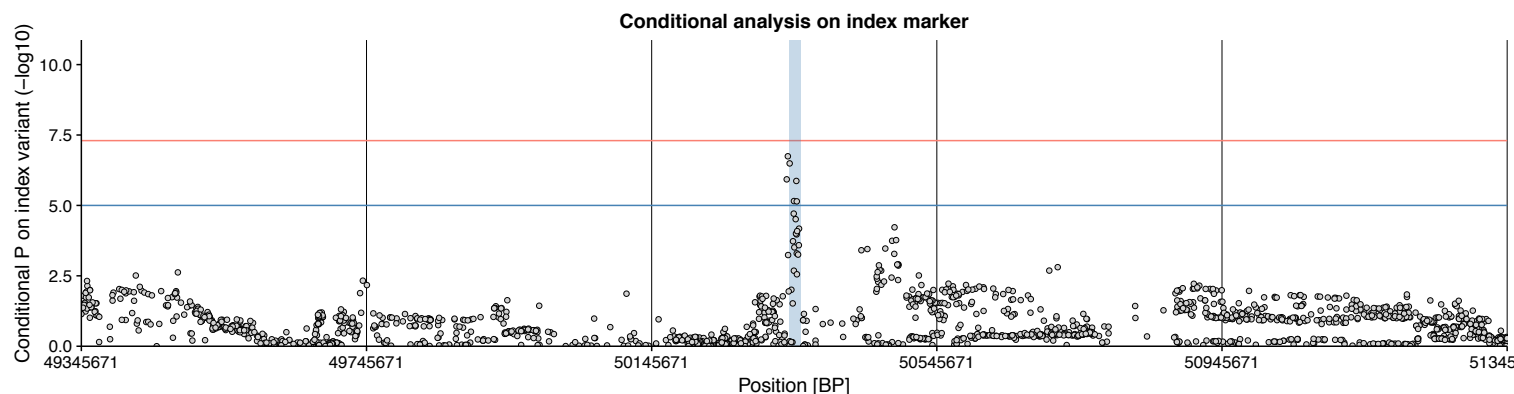
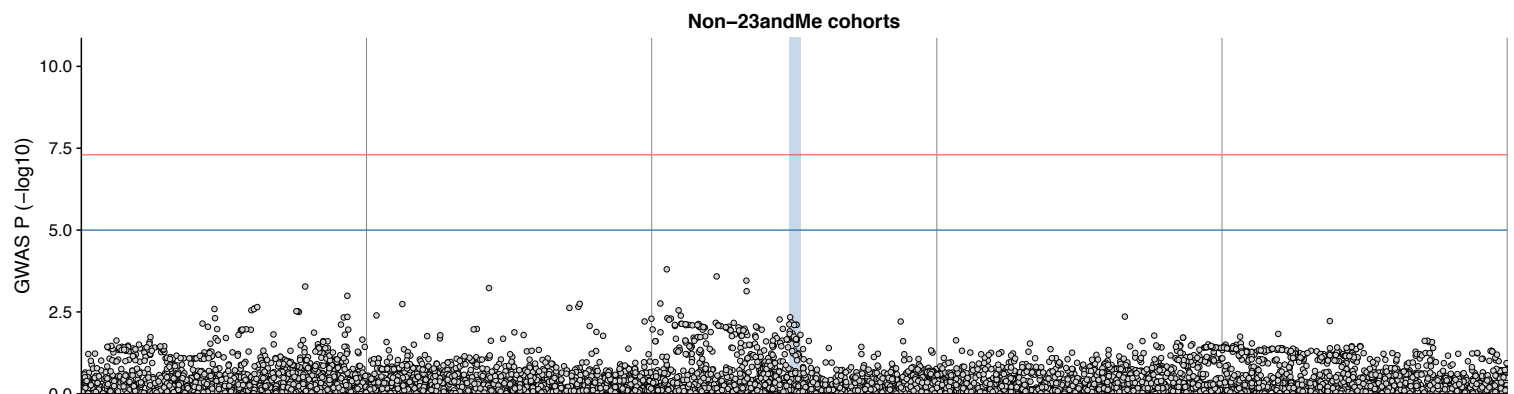
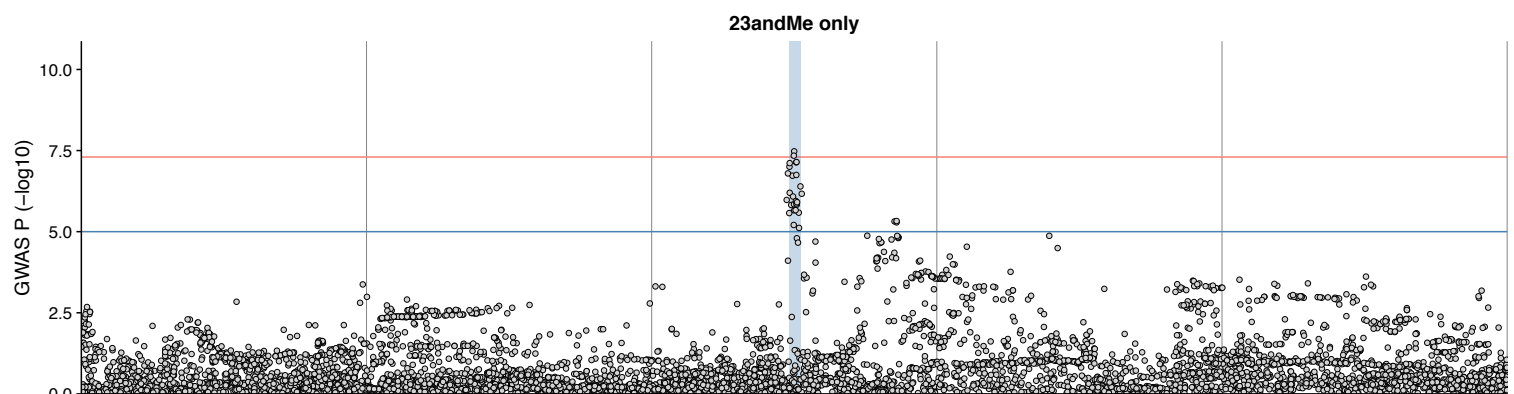
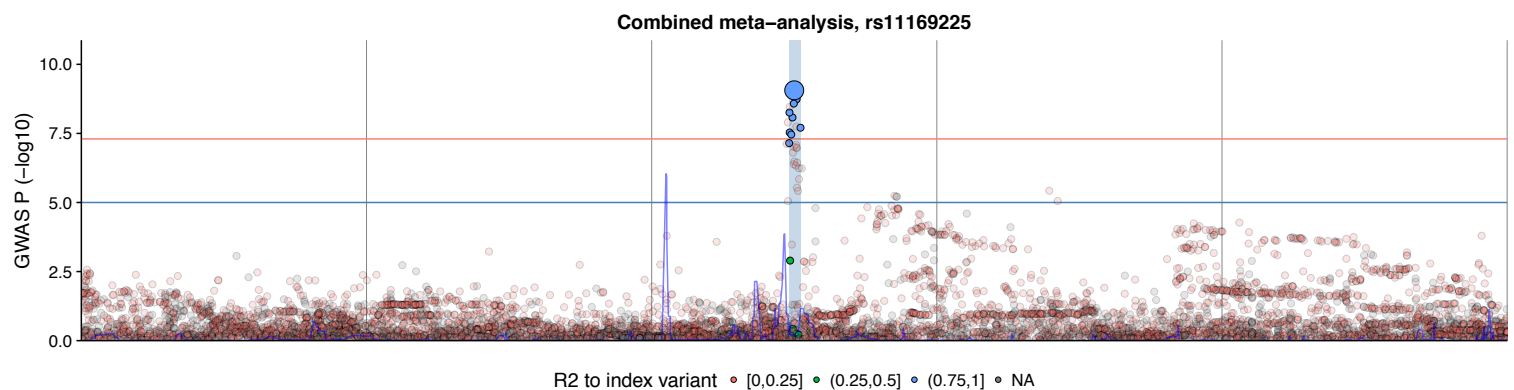
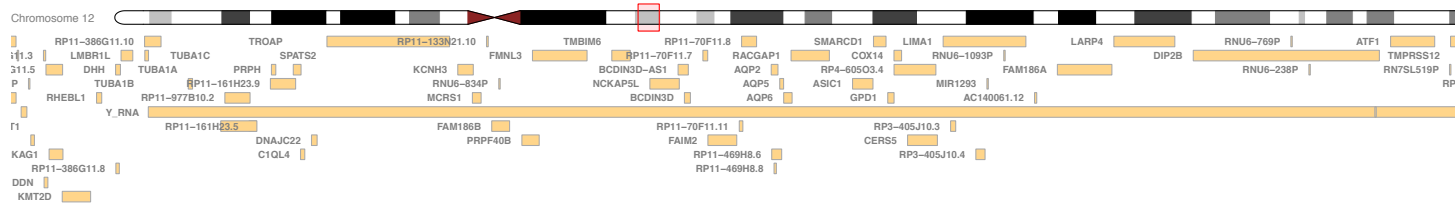
Middle upper panel shows association results for 23andme, and middle lower panel shows association values for non-23andme cohorts.

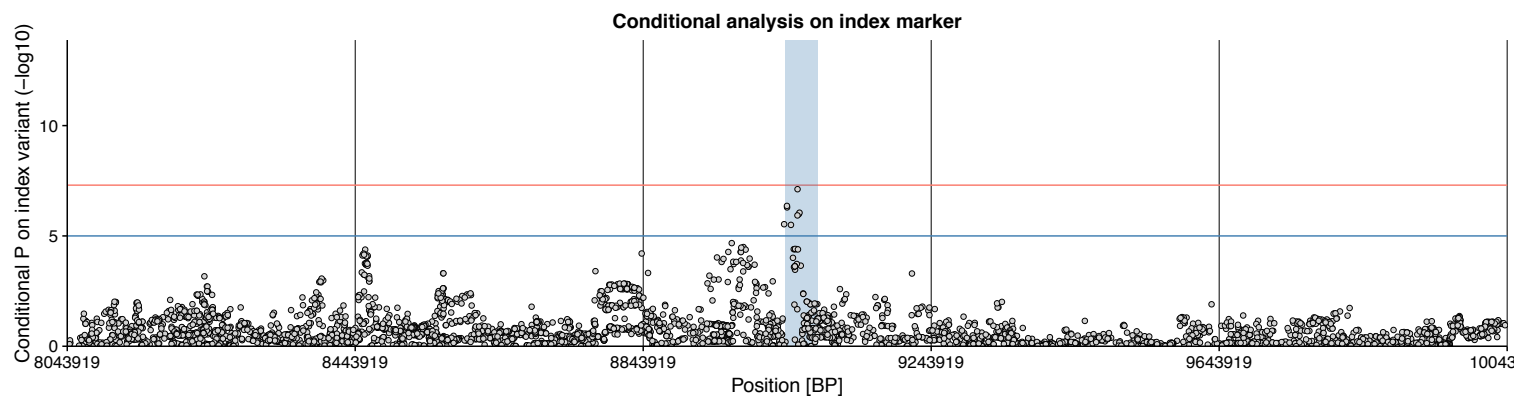
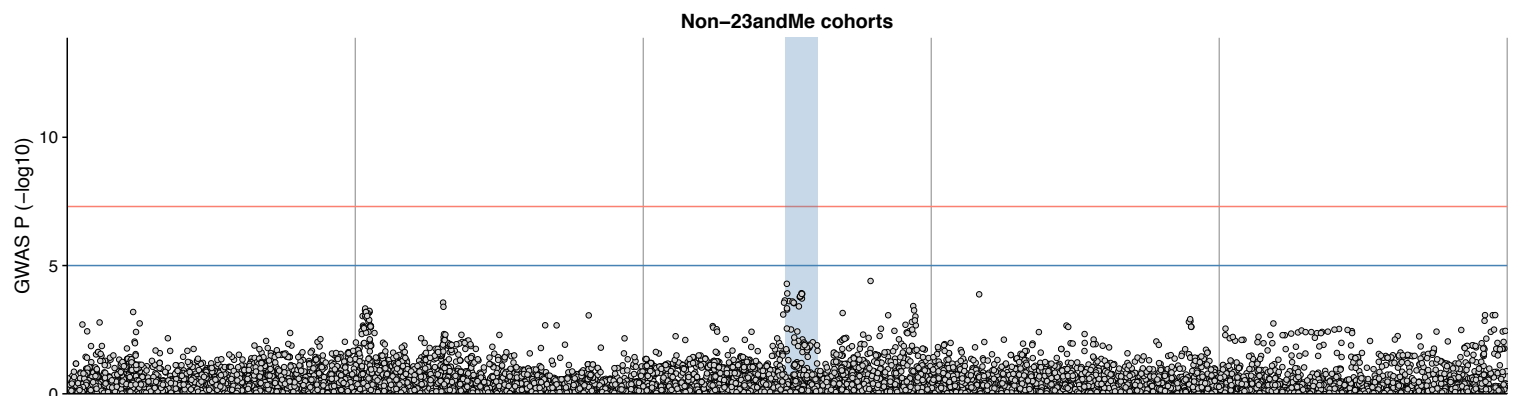
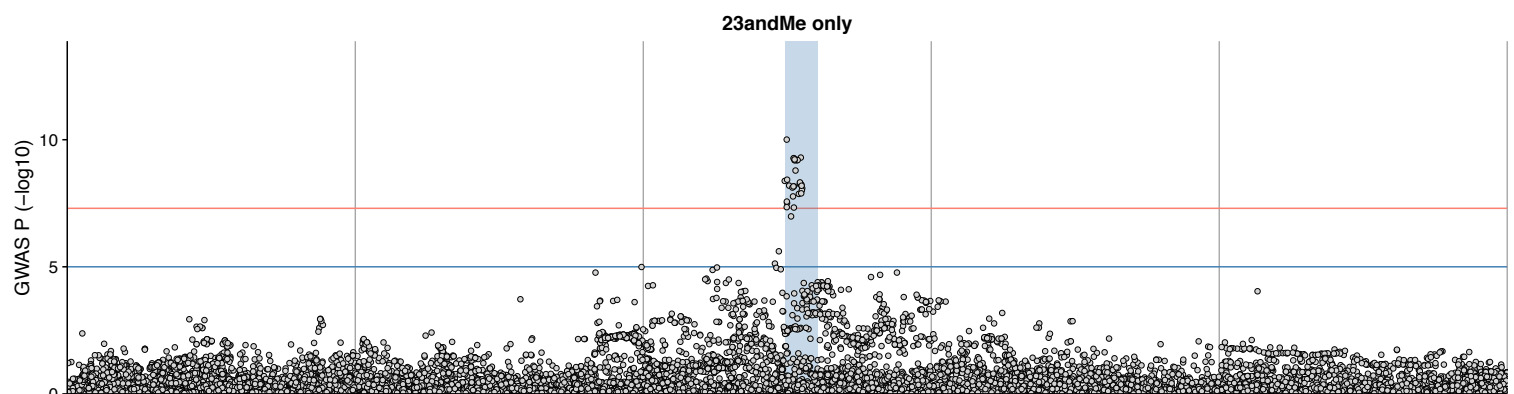
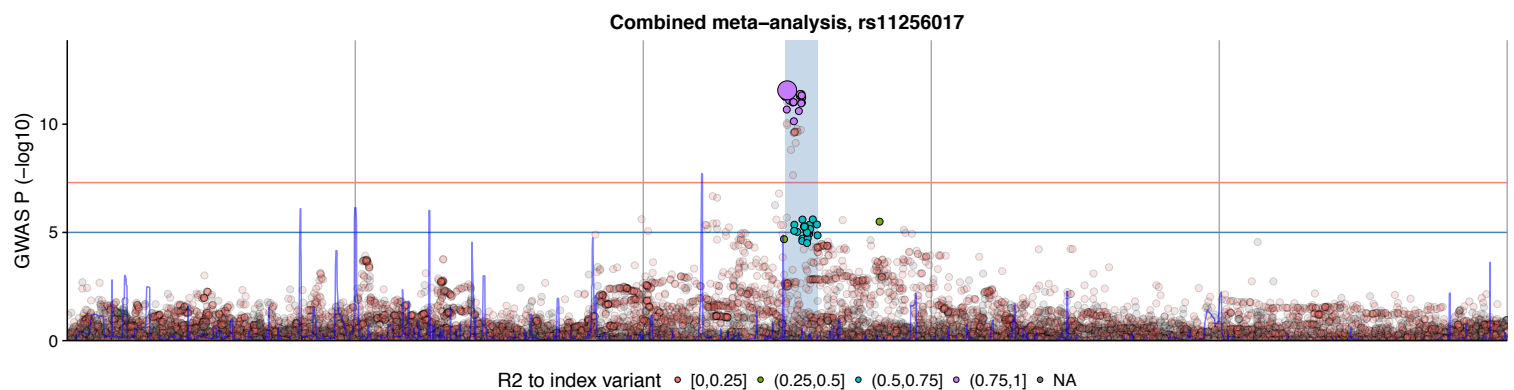
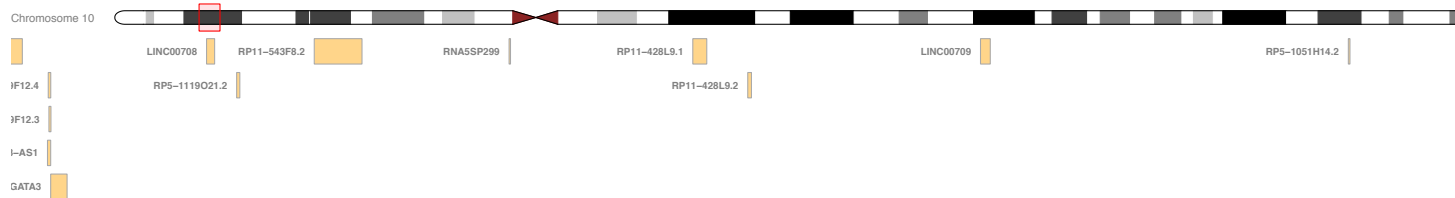
Bottom panel shows results for the conditional analysis on index marker.

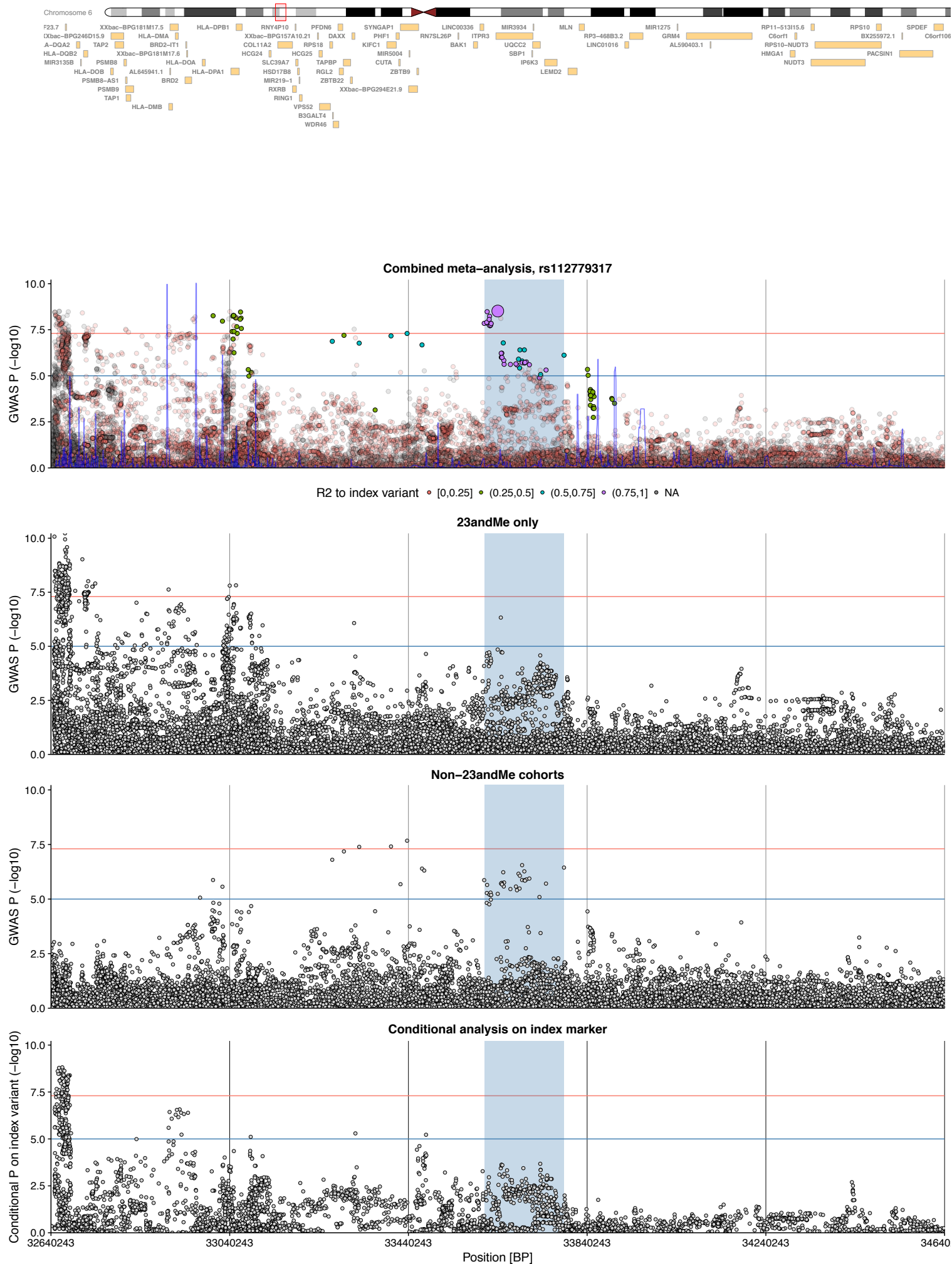


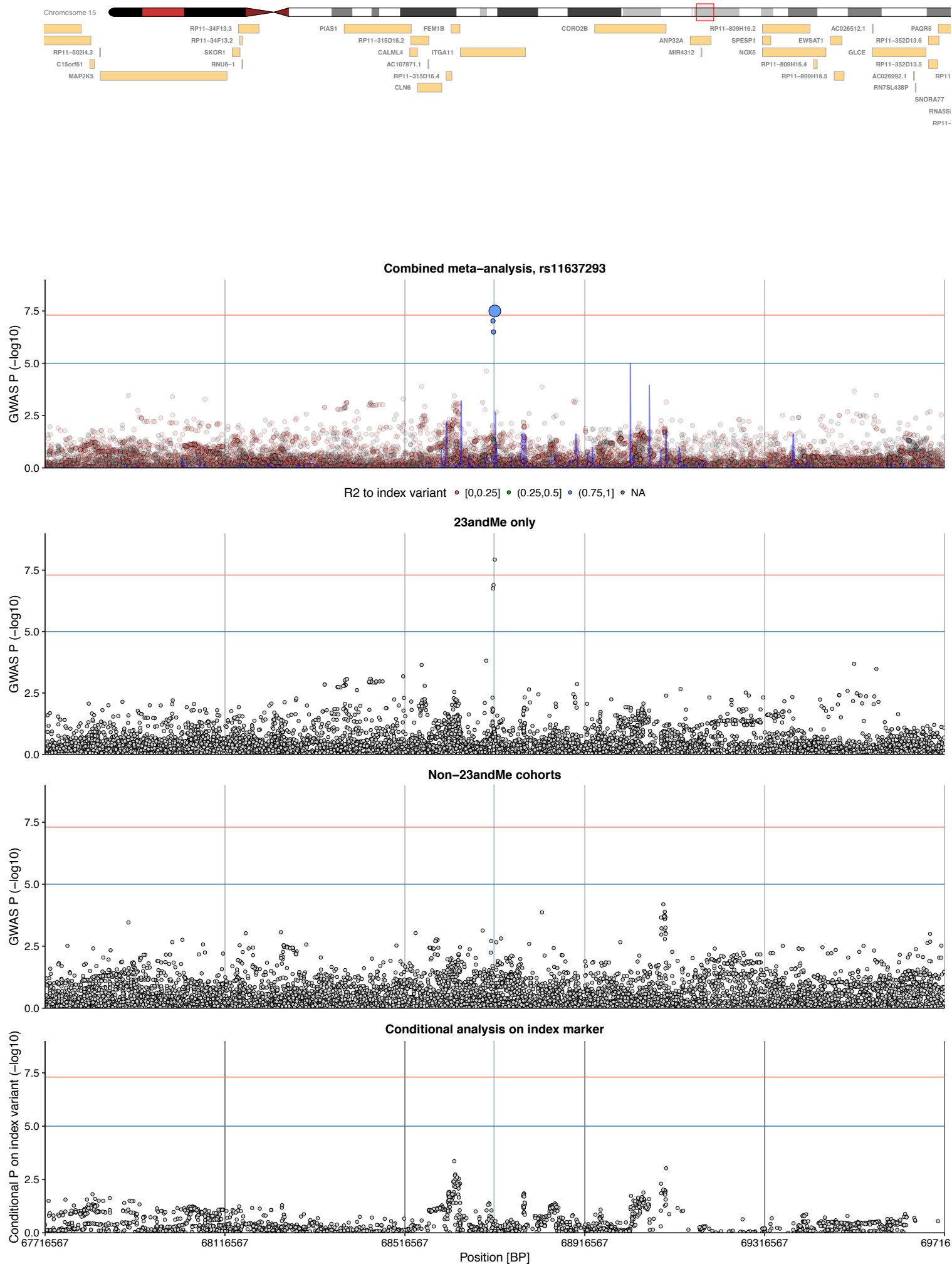


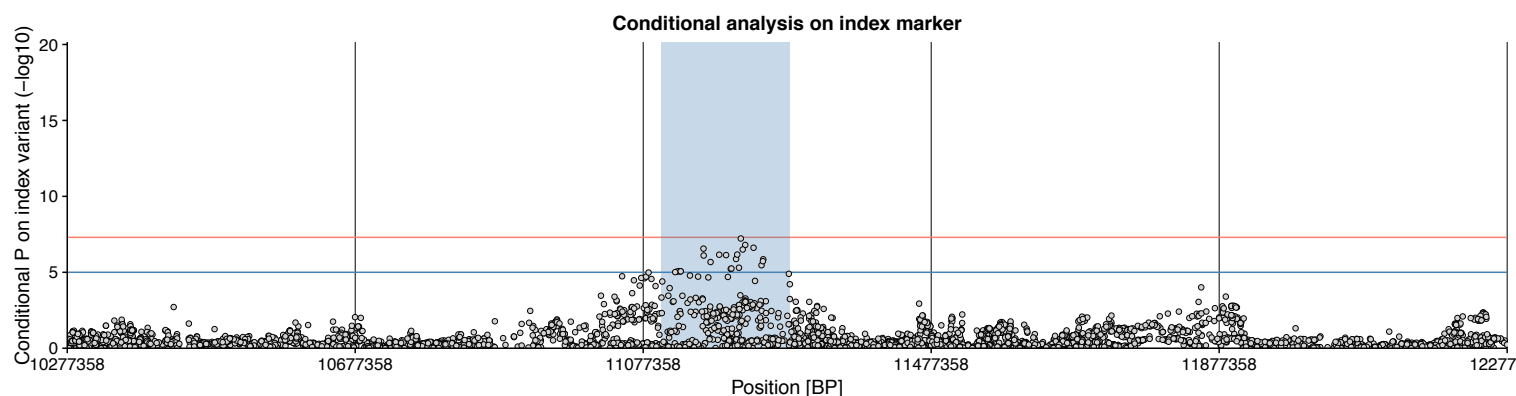
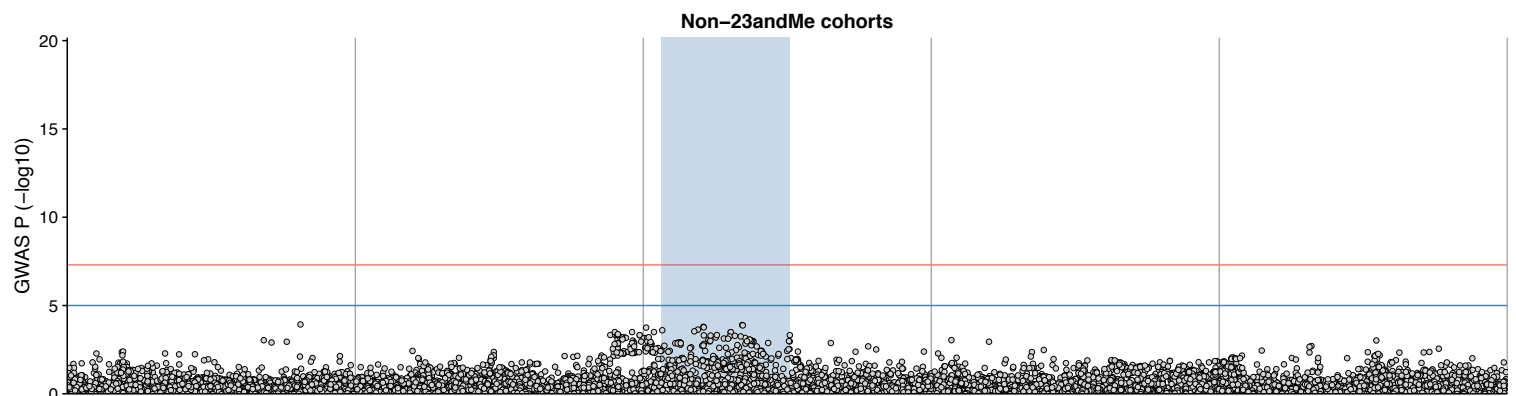
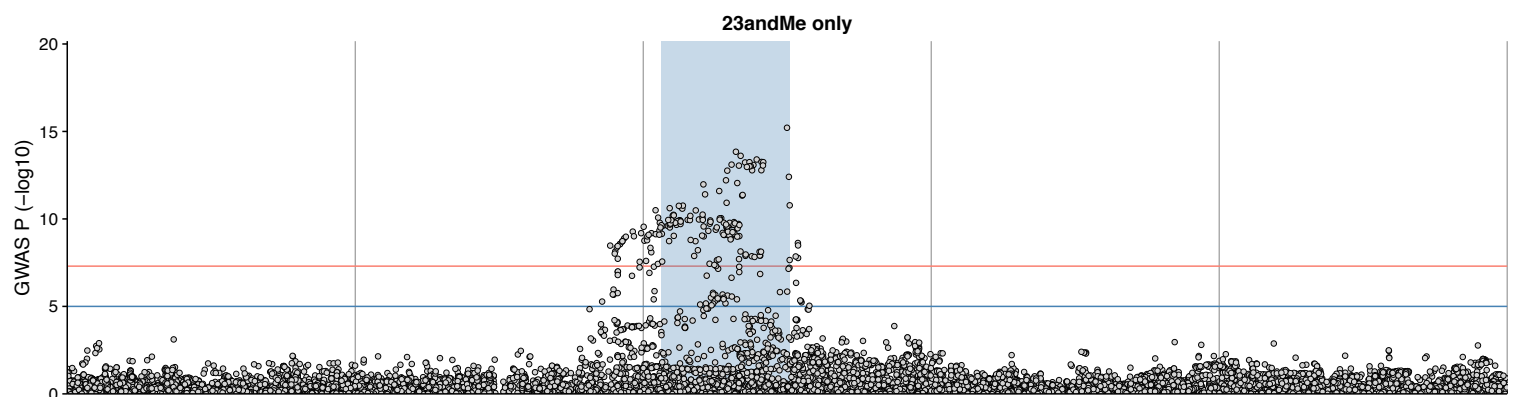
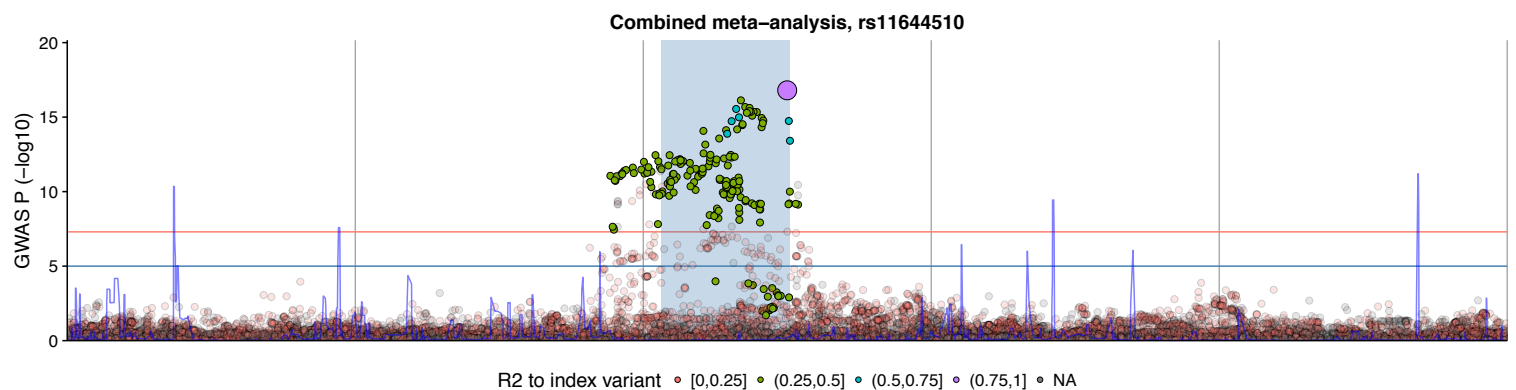
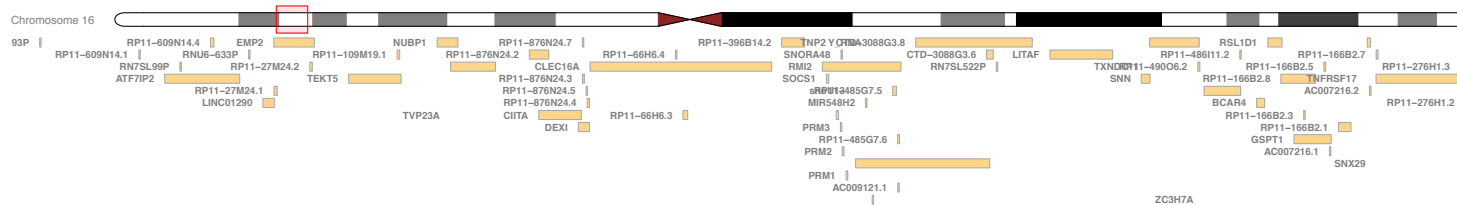


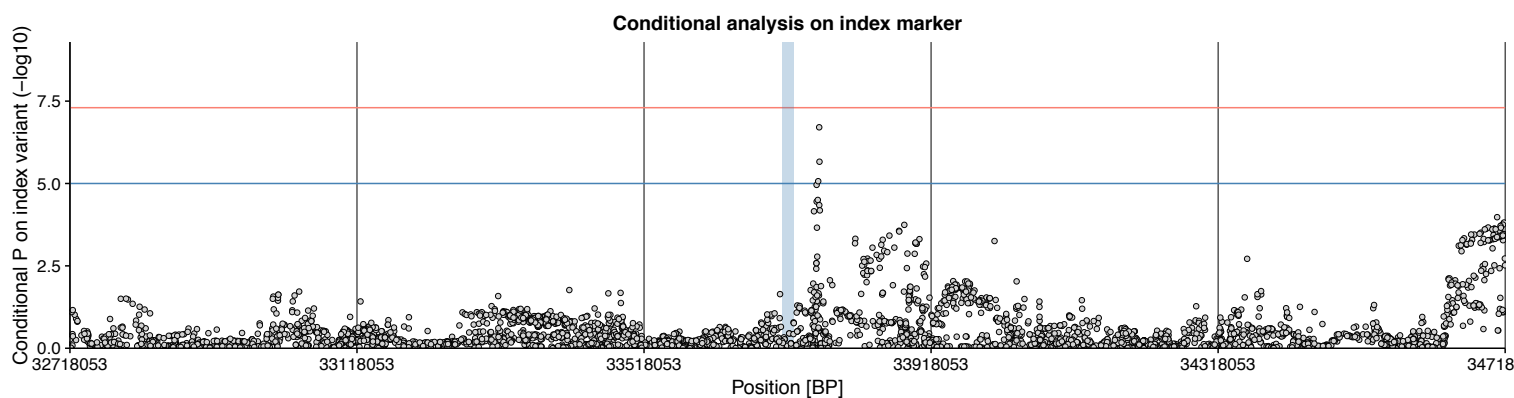
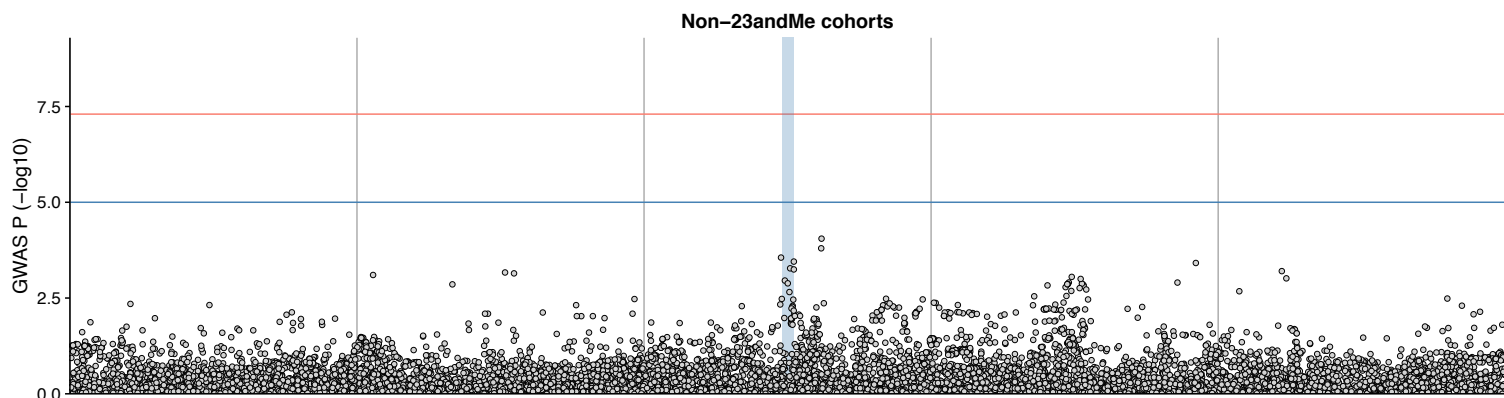
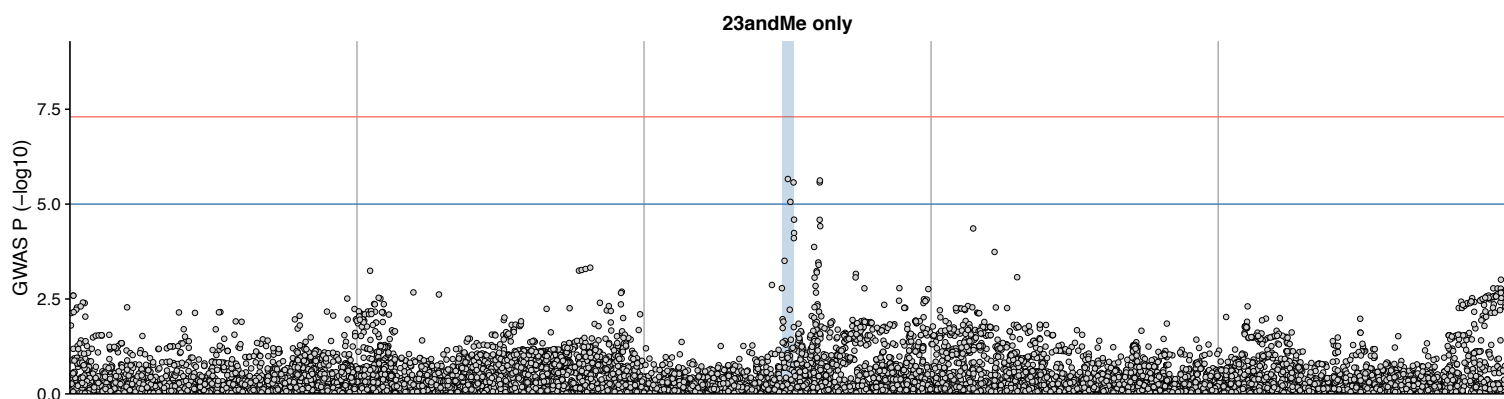
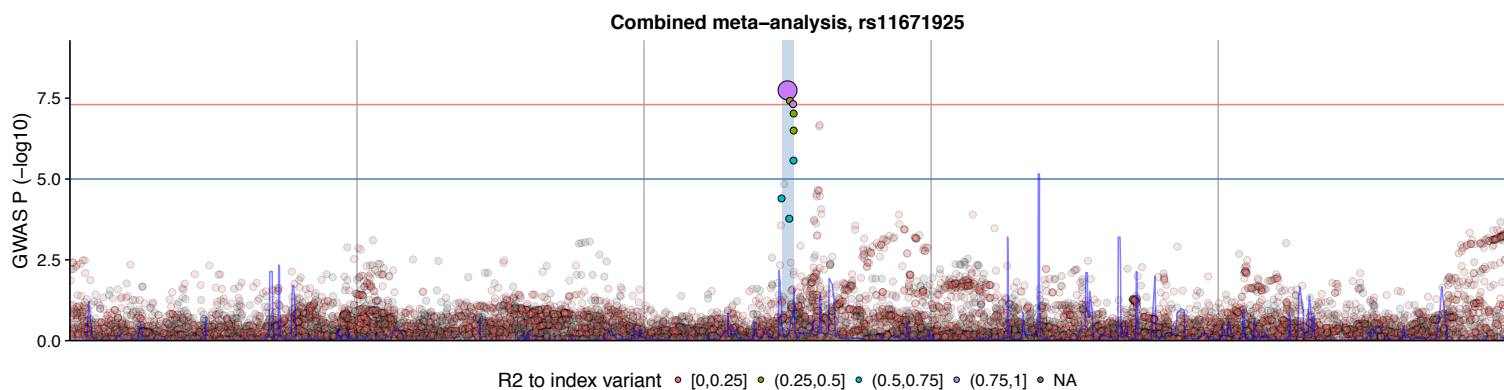
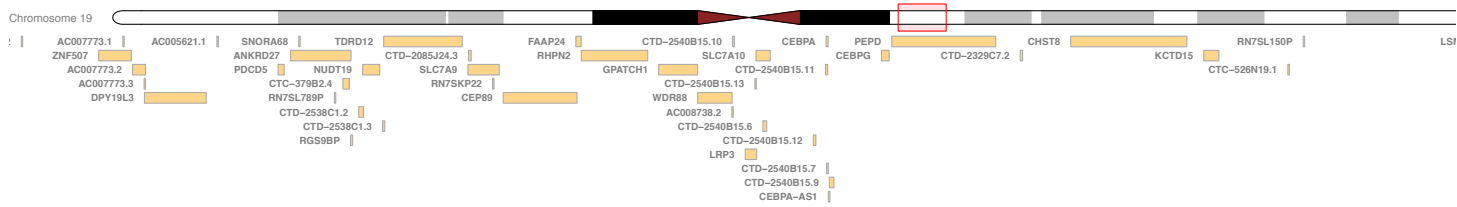


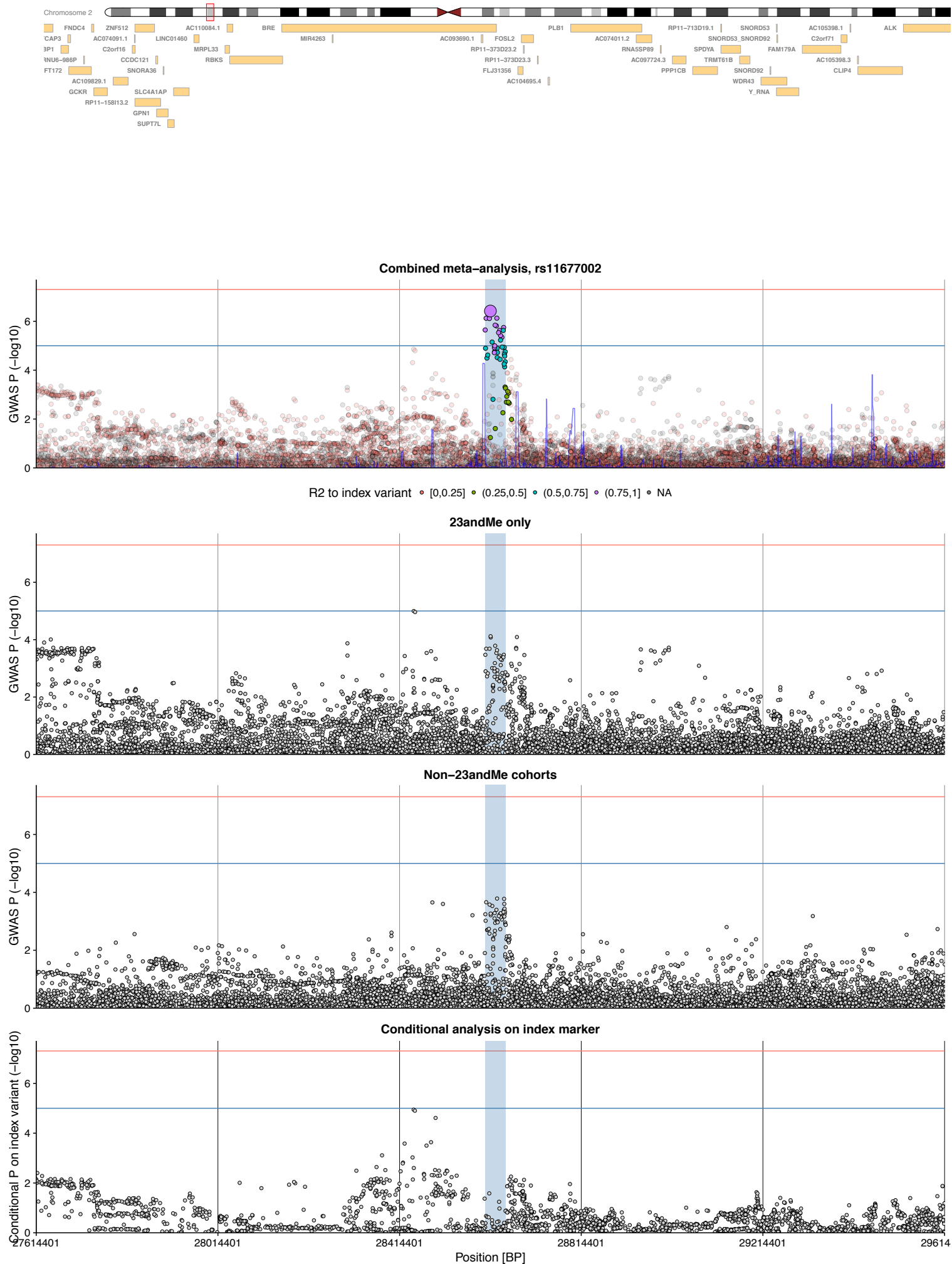


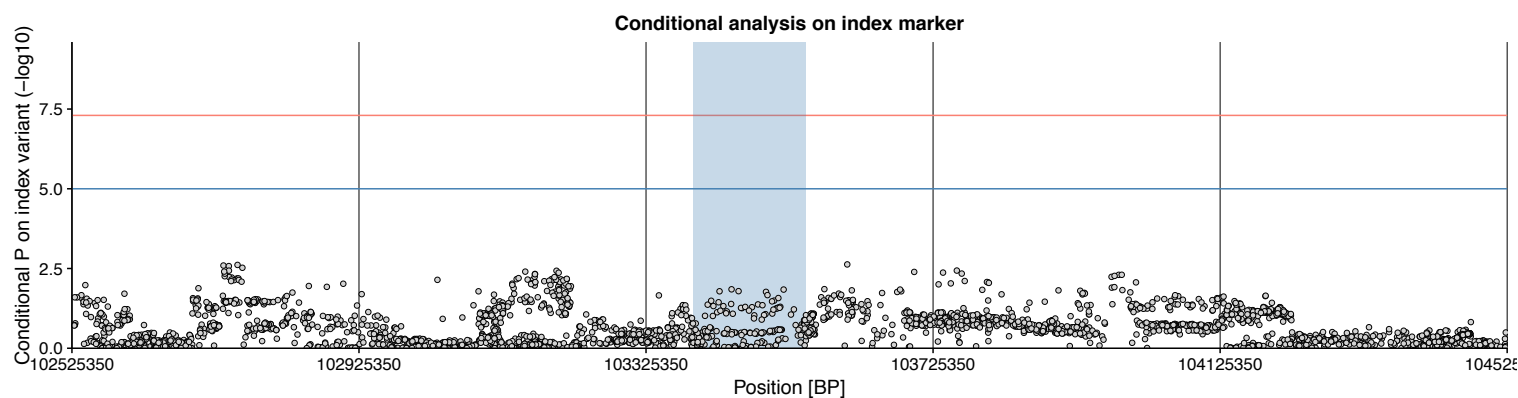
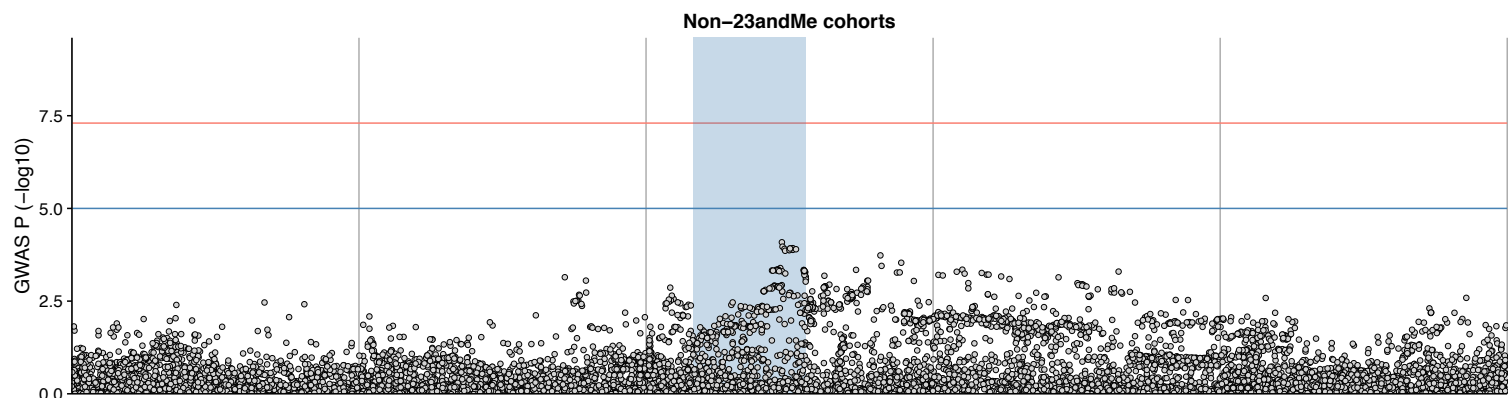
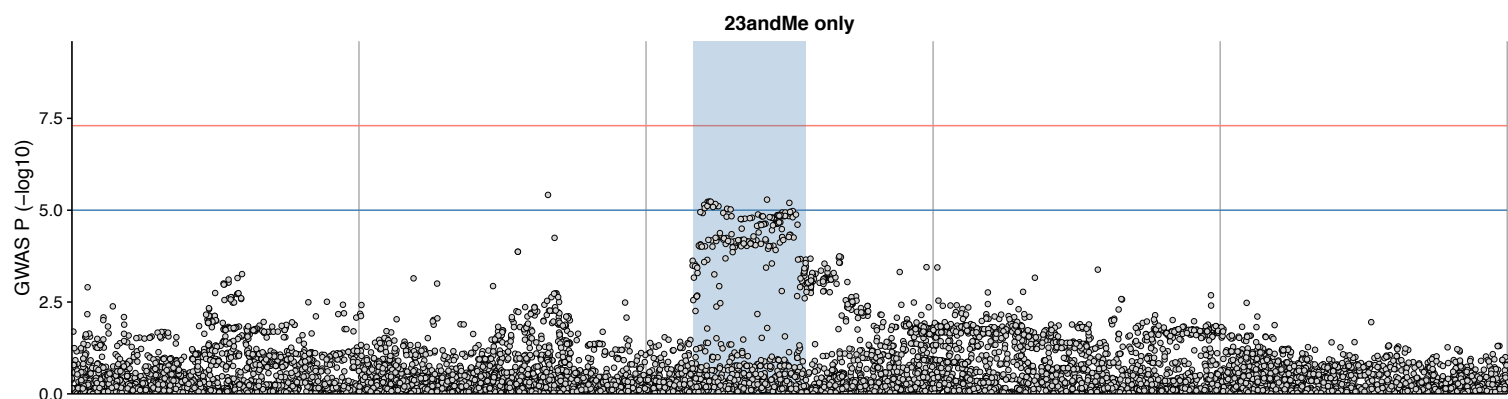
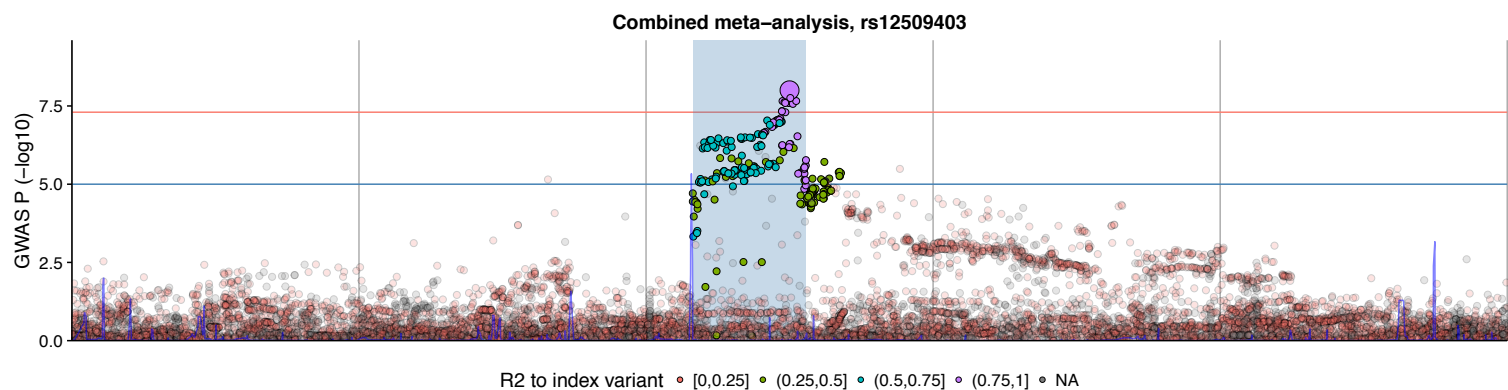
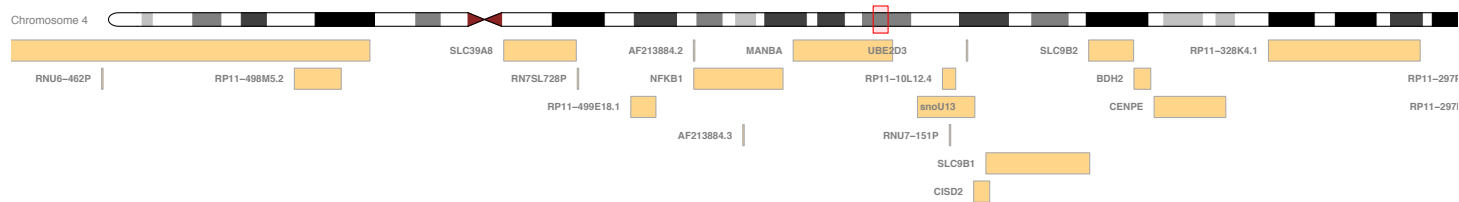


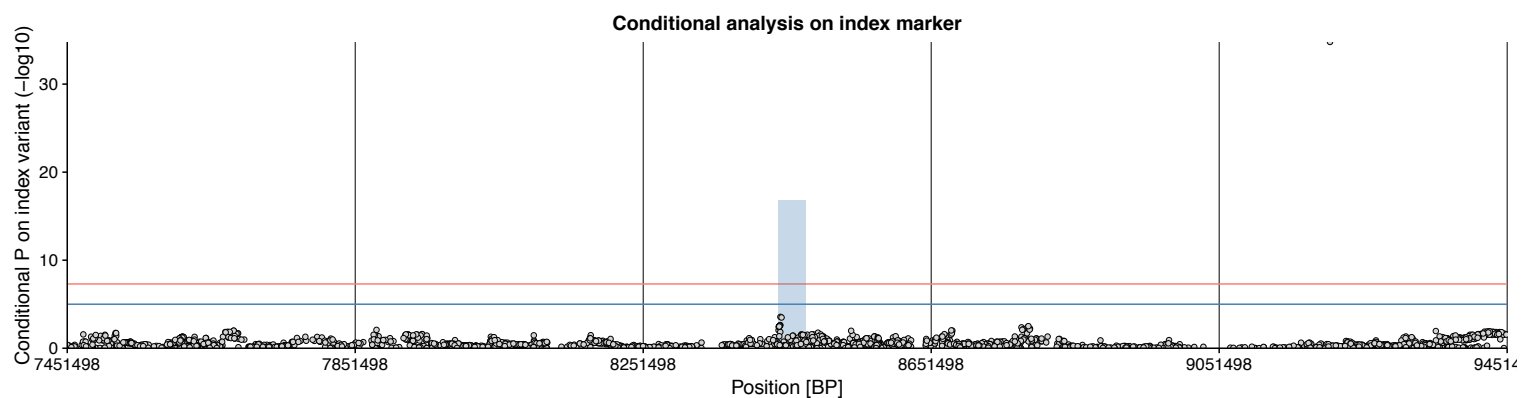
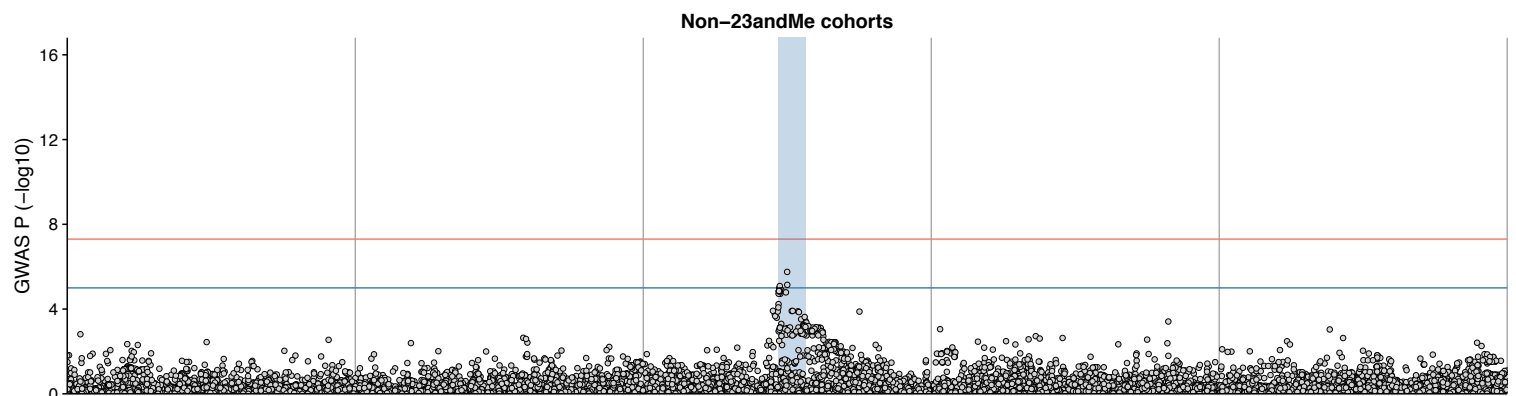
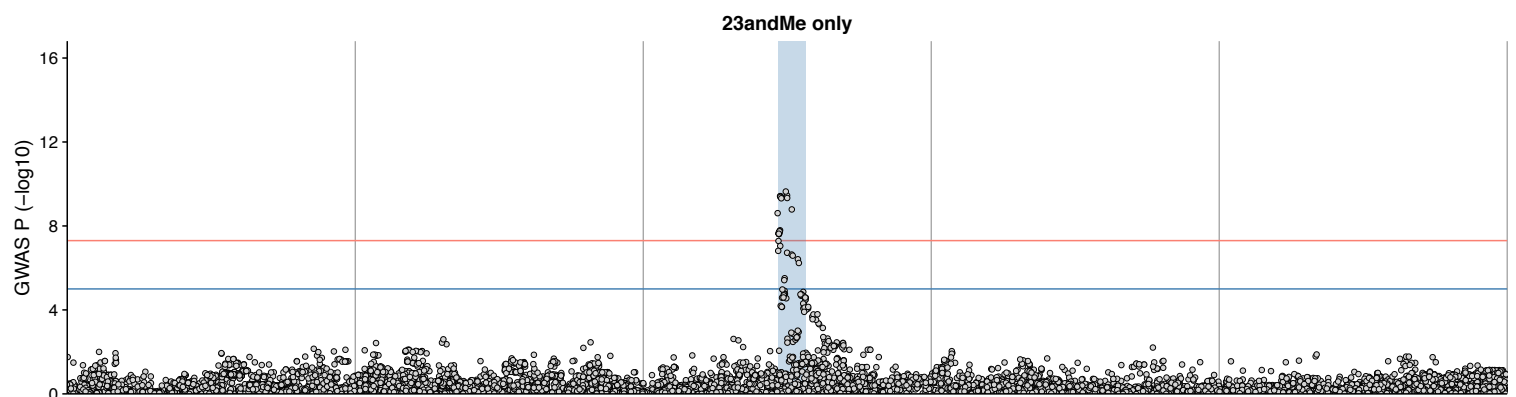
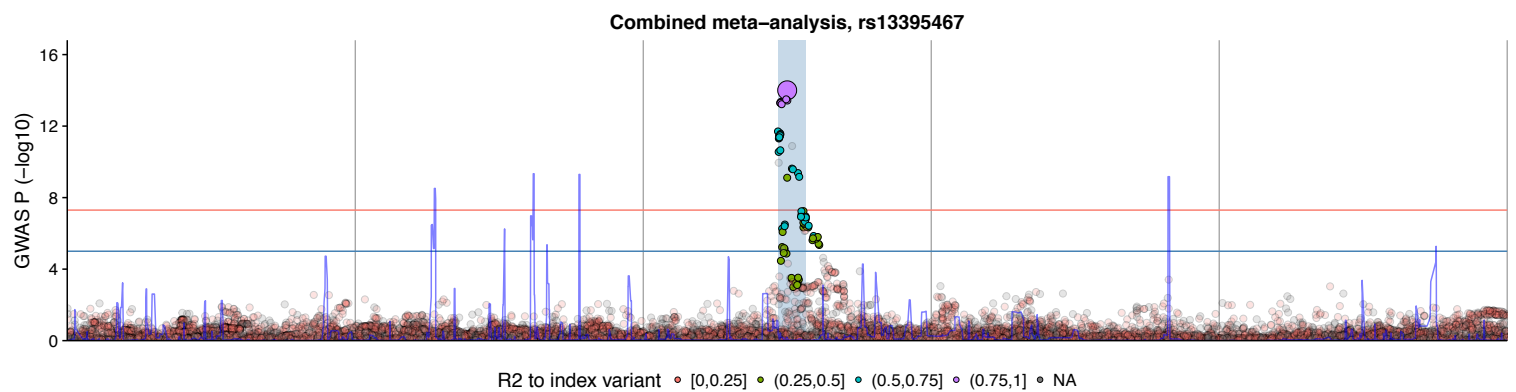
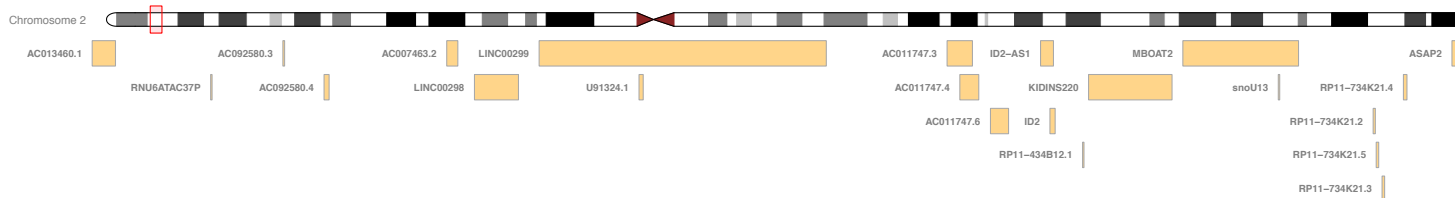


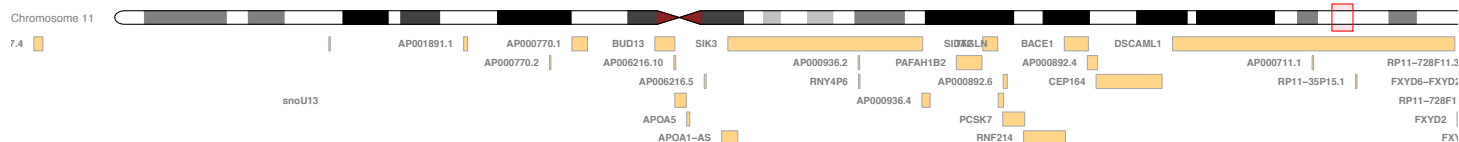




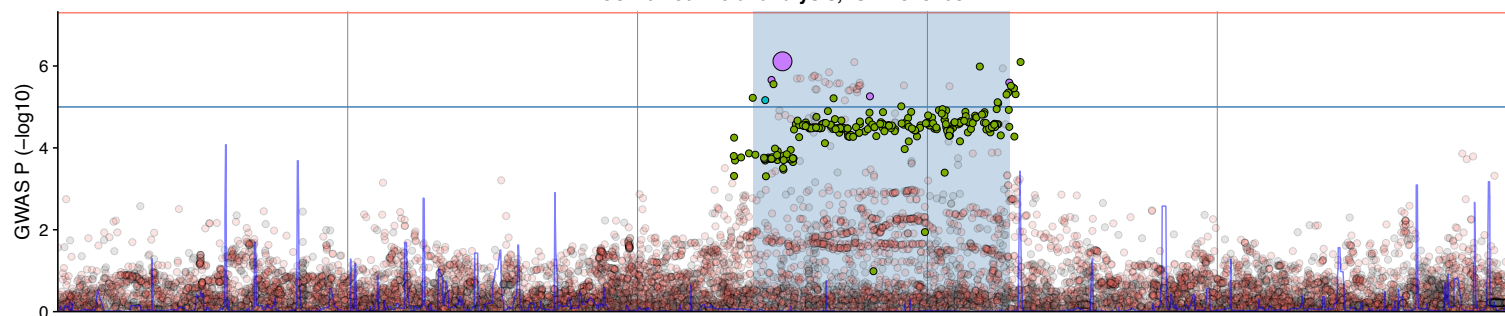






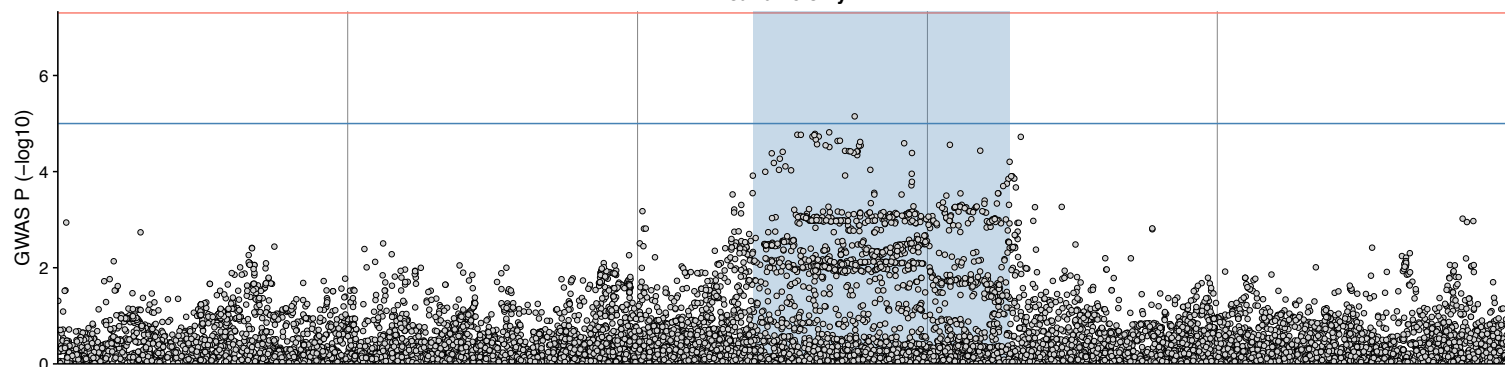


Combined meta-analysis, rs141023293

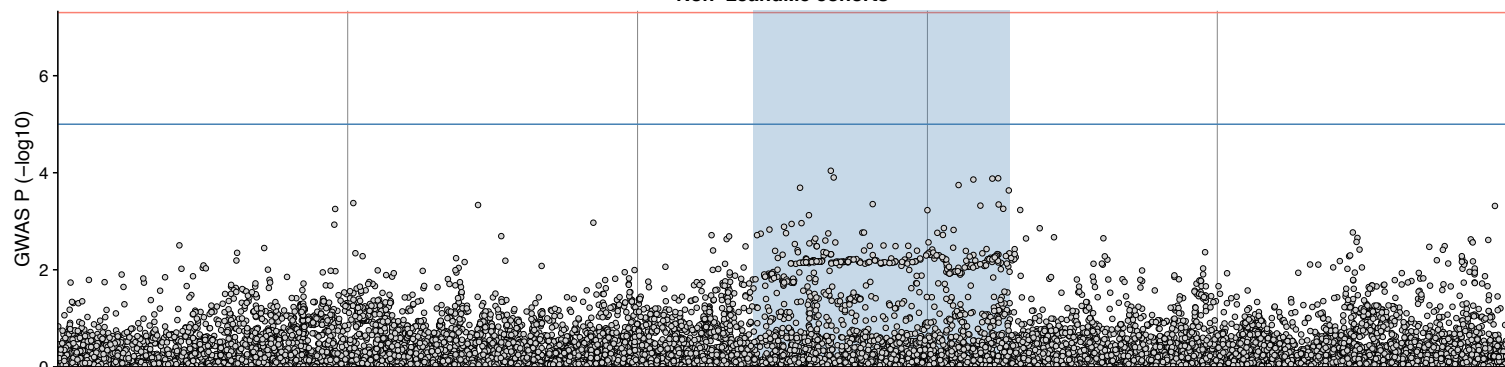


R2 to index variant • [0,0.25] • (0.25,0.5] • (0.5,0.75] • (0.75,1] • NA

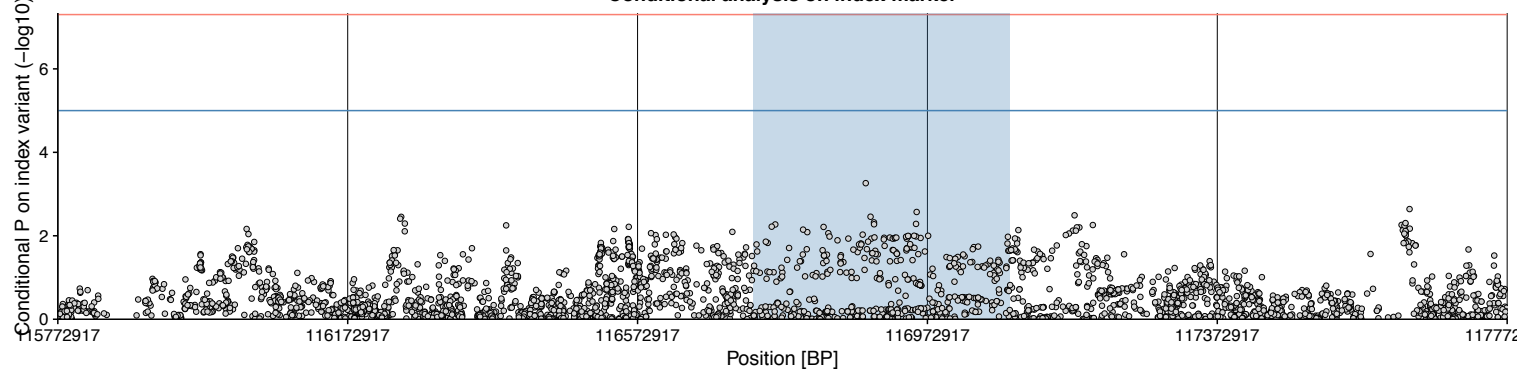
23andMe only

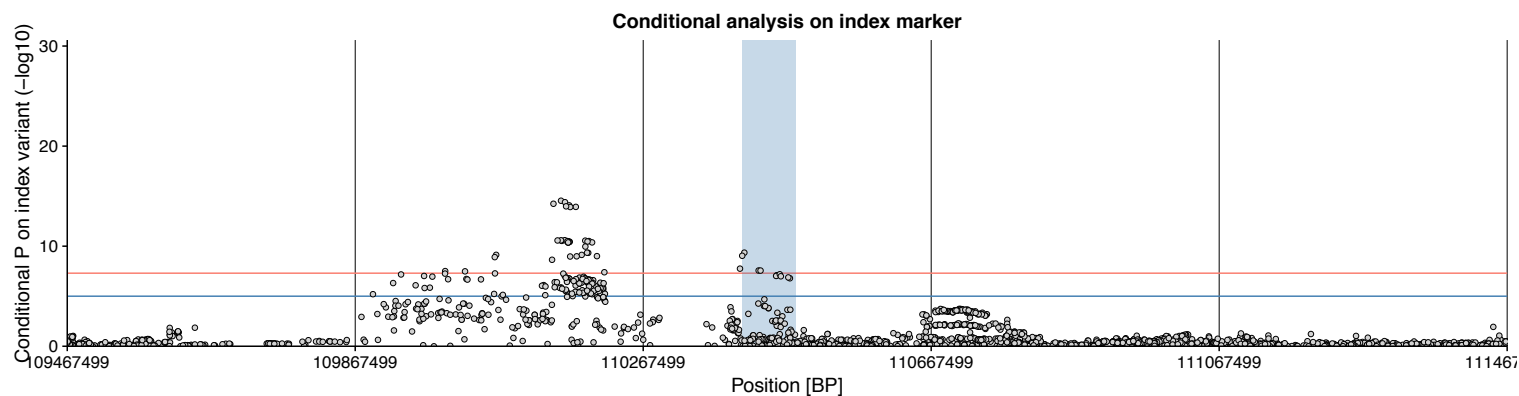
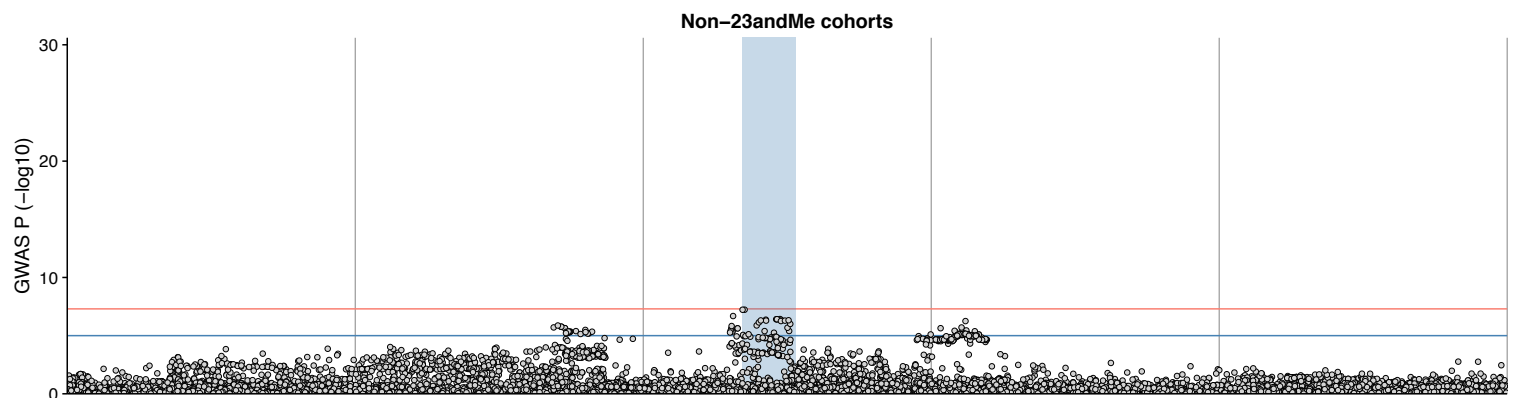
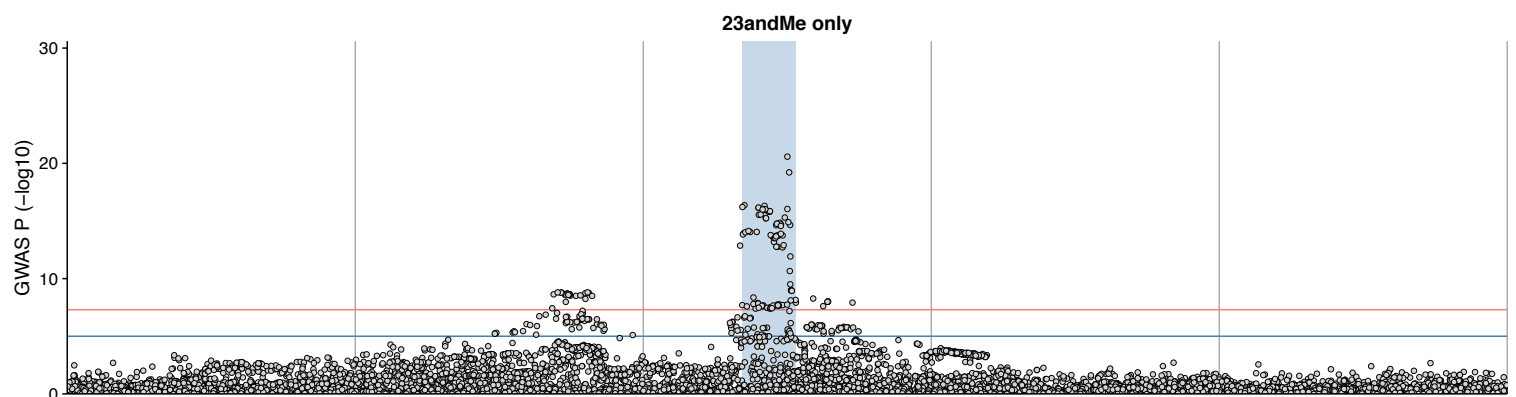
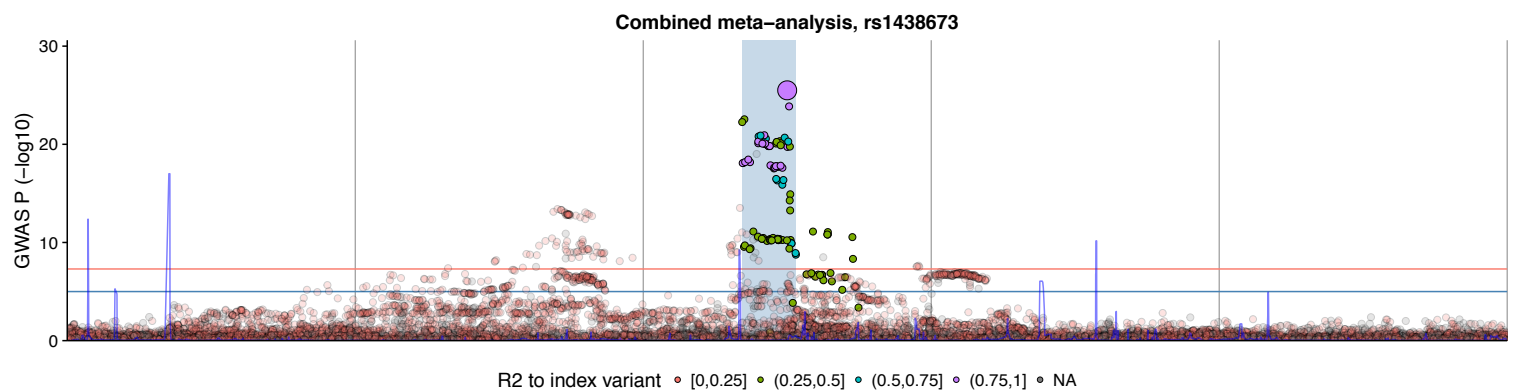
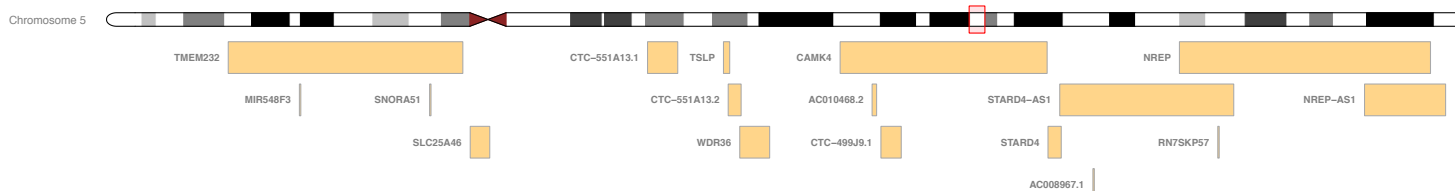


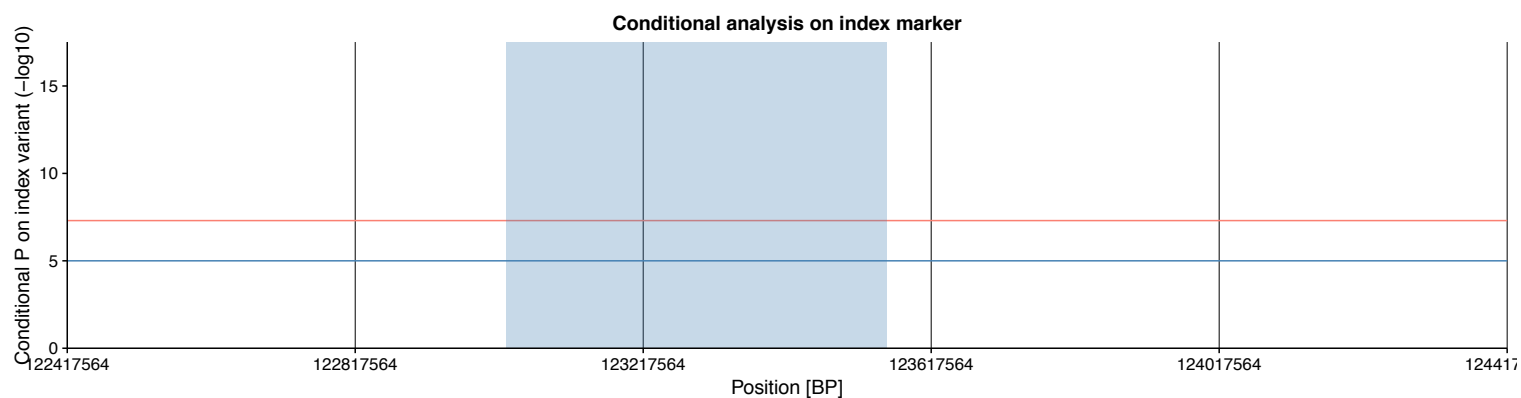
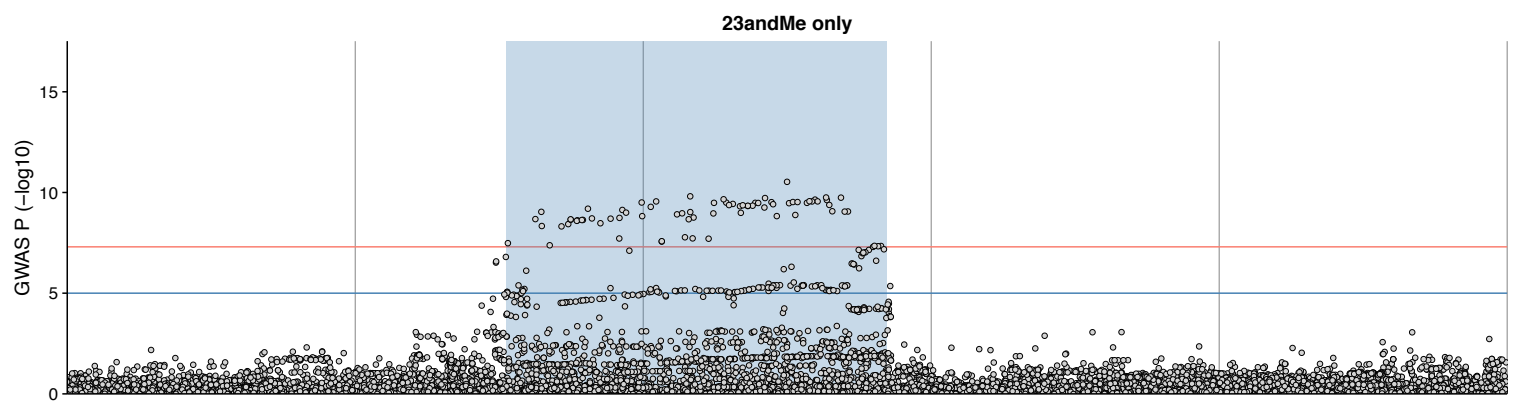
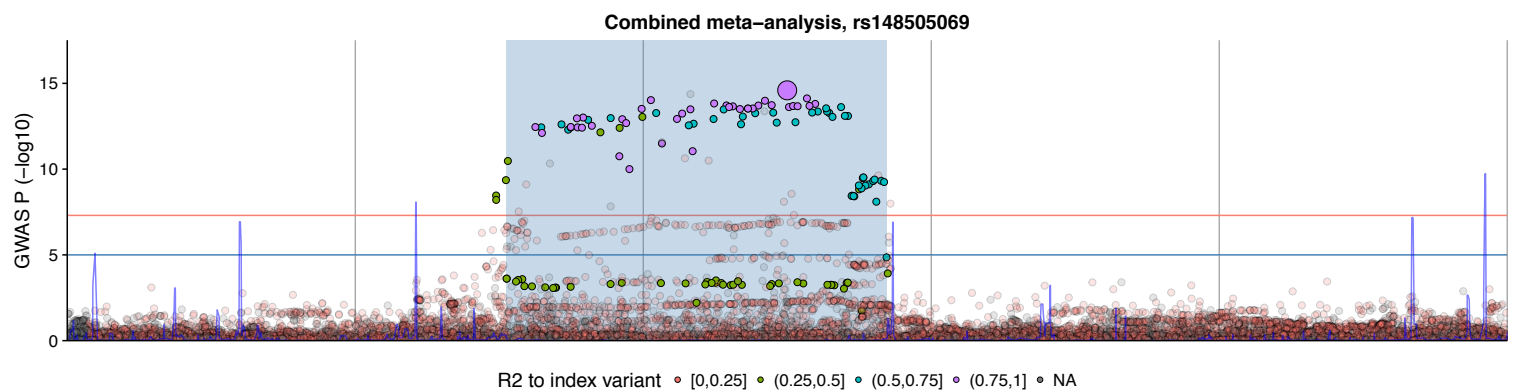
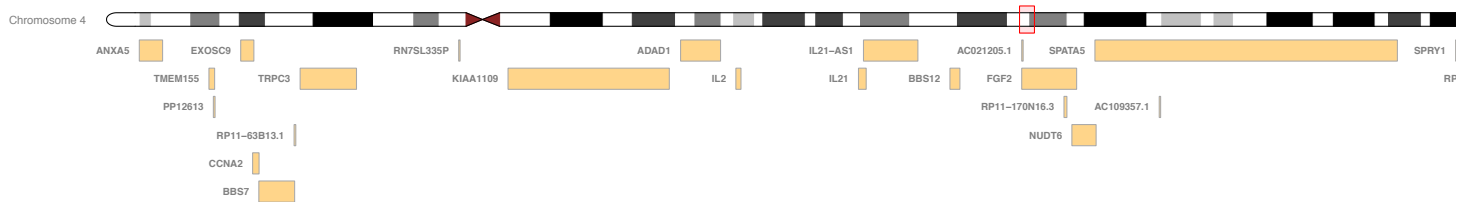
Non-23andMe cohorts

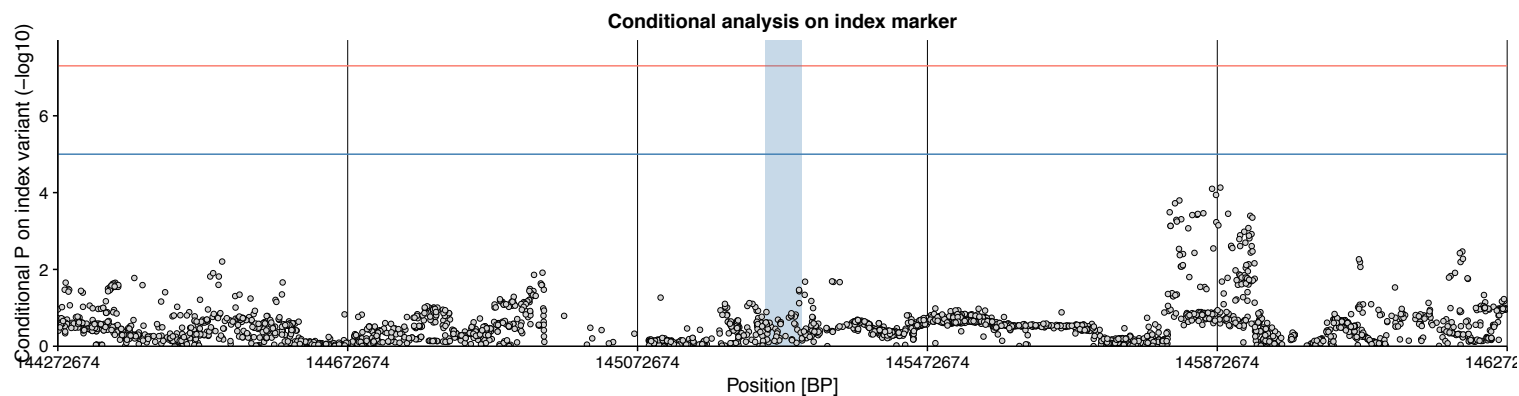
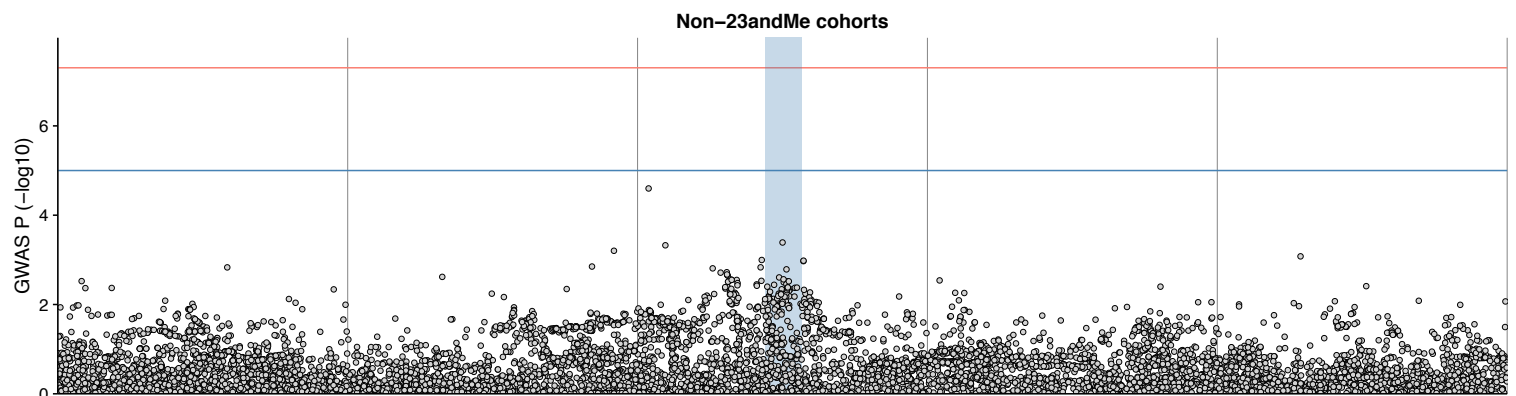
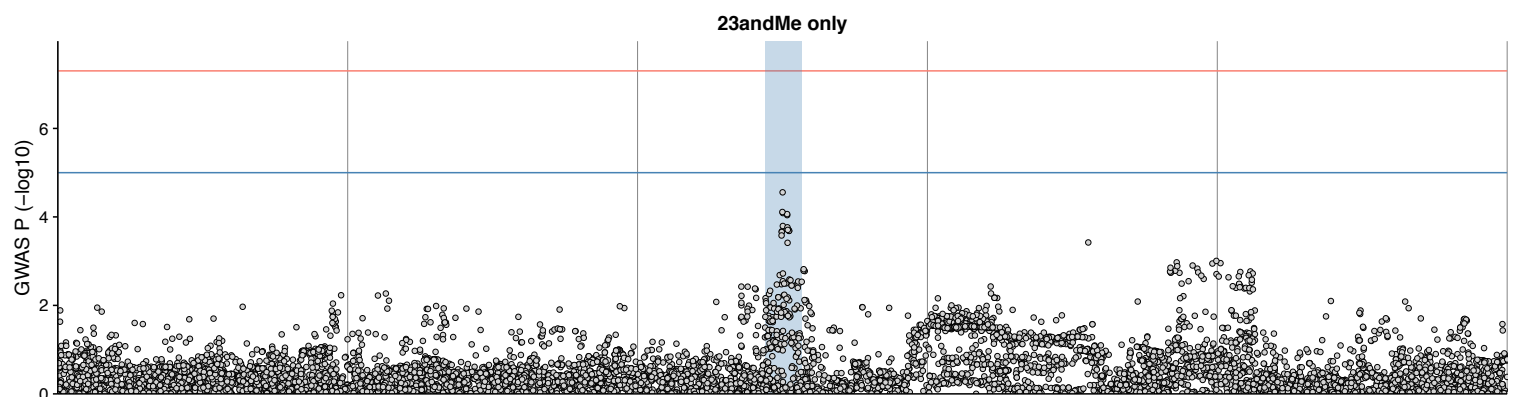
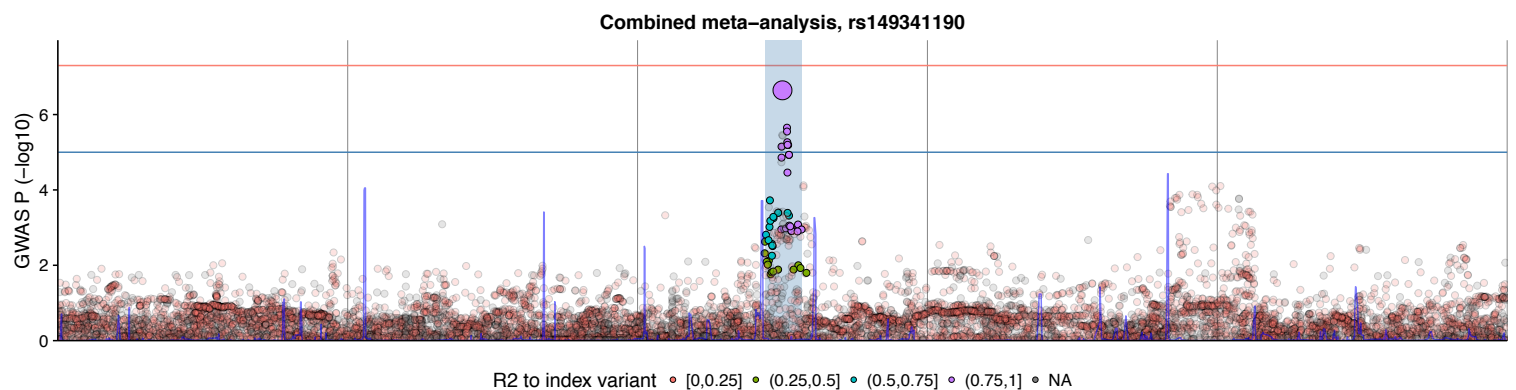
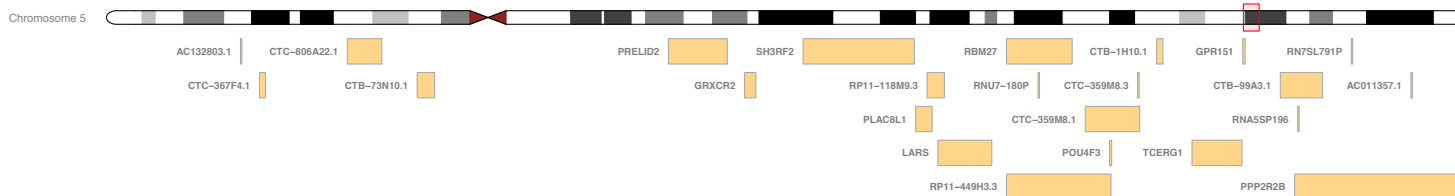


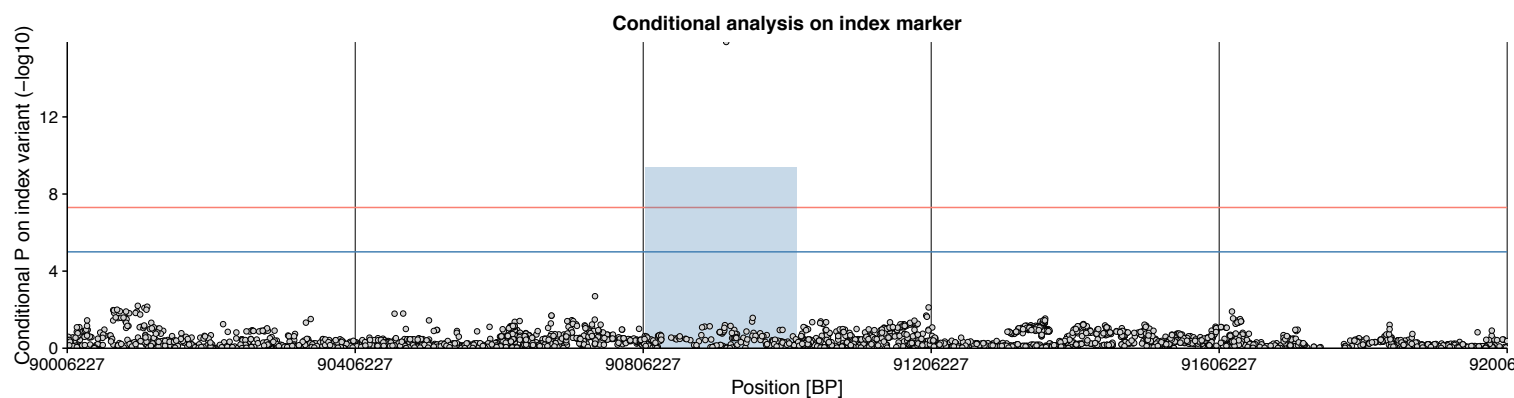
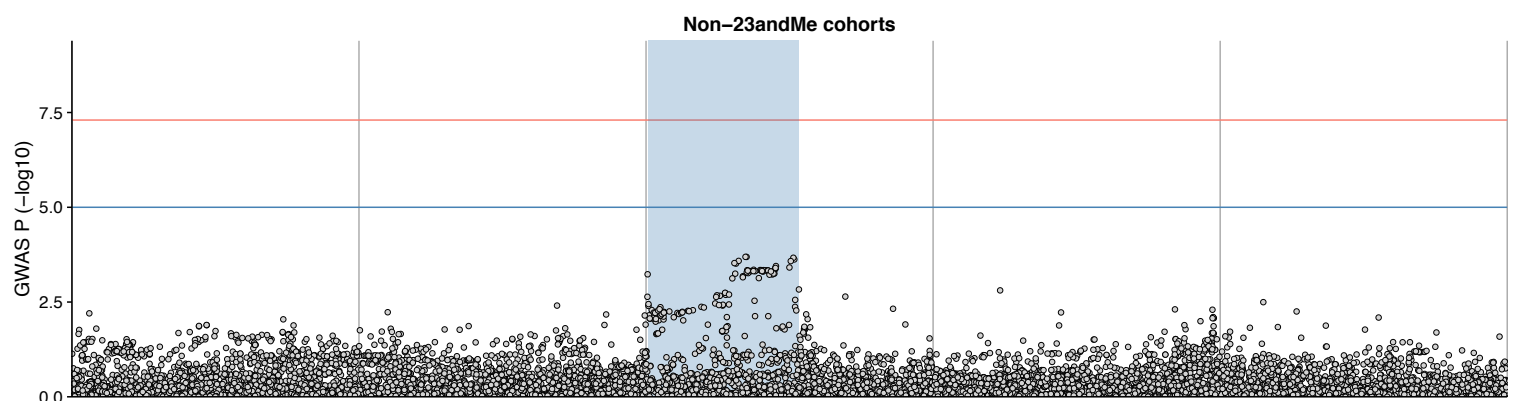
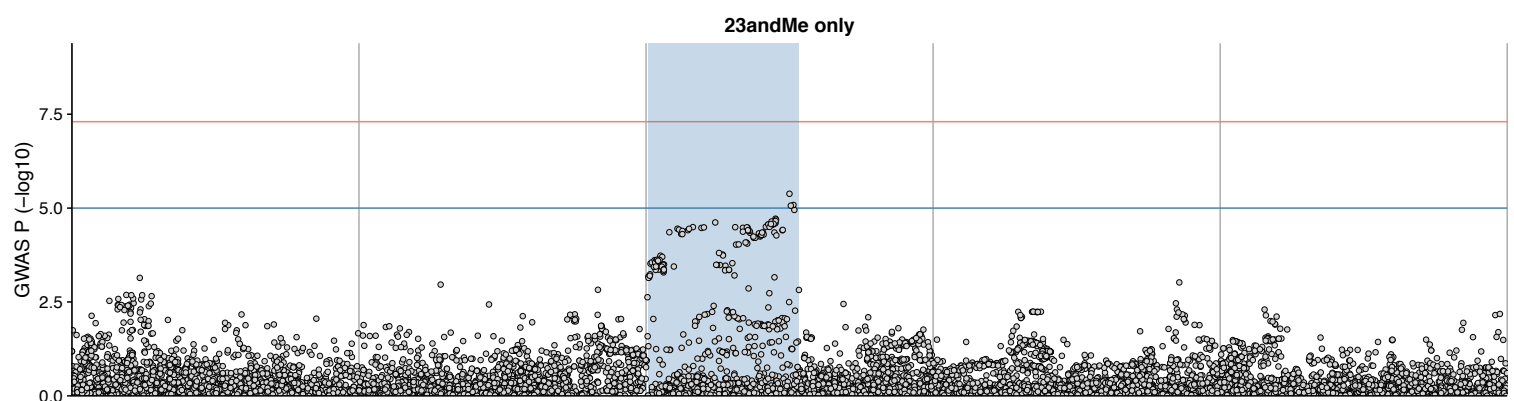
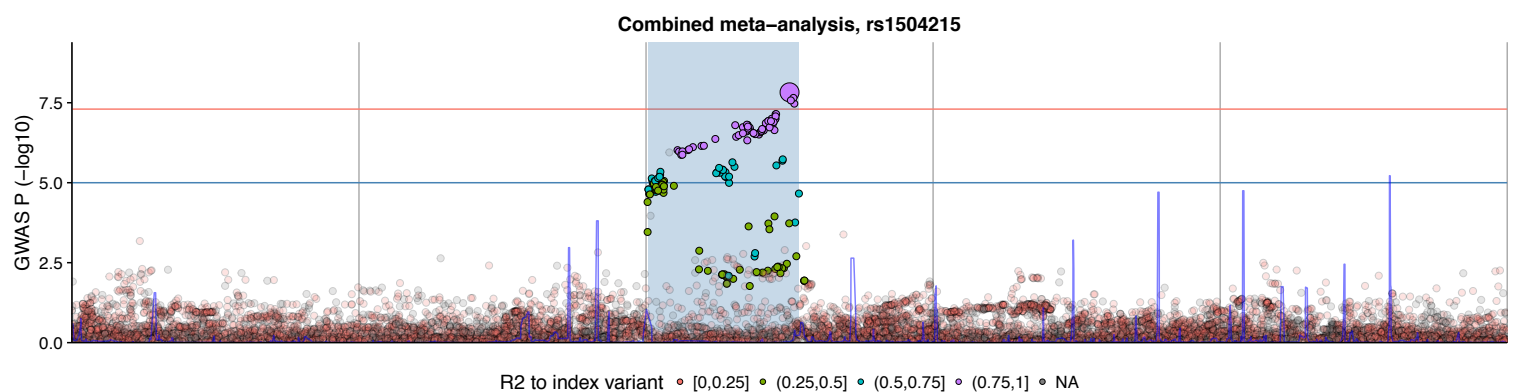
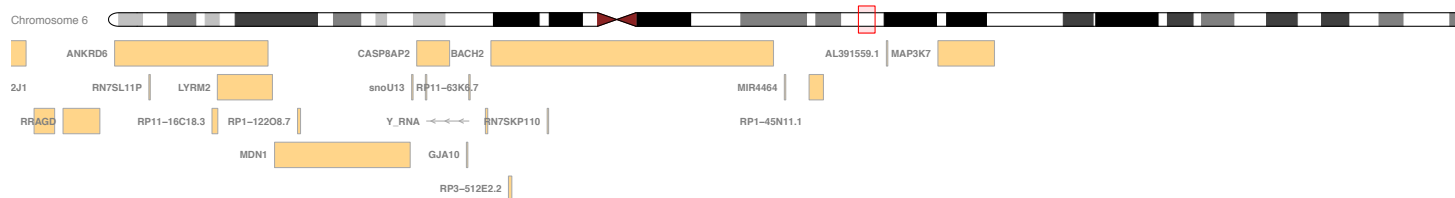
Conditional analysis on index marker

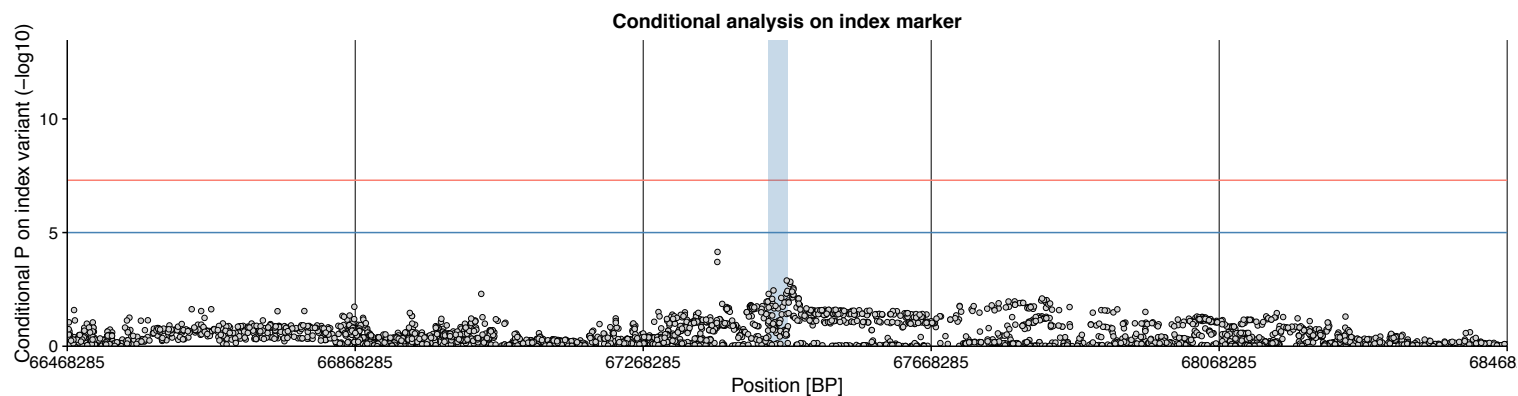
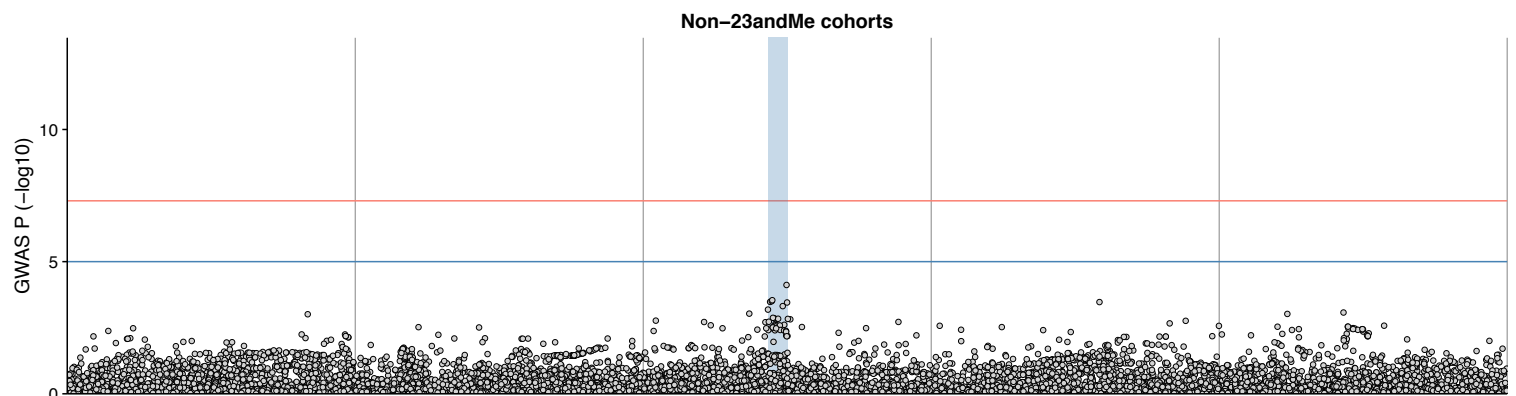
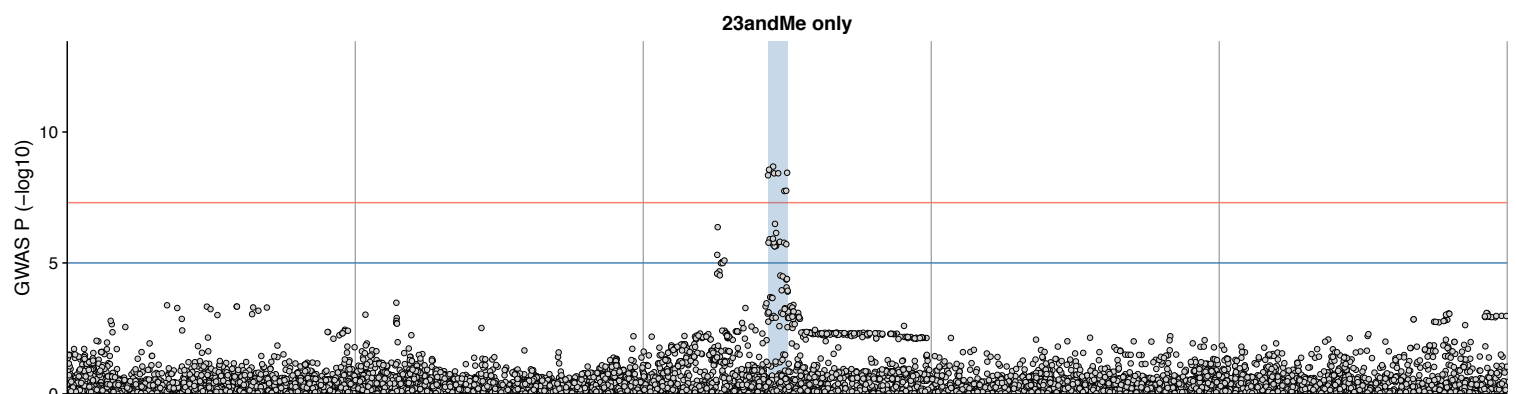
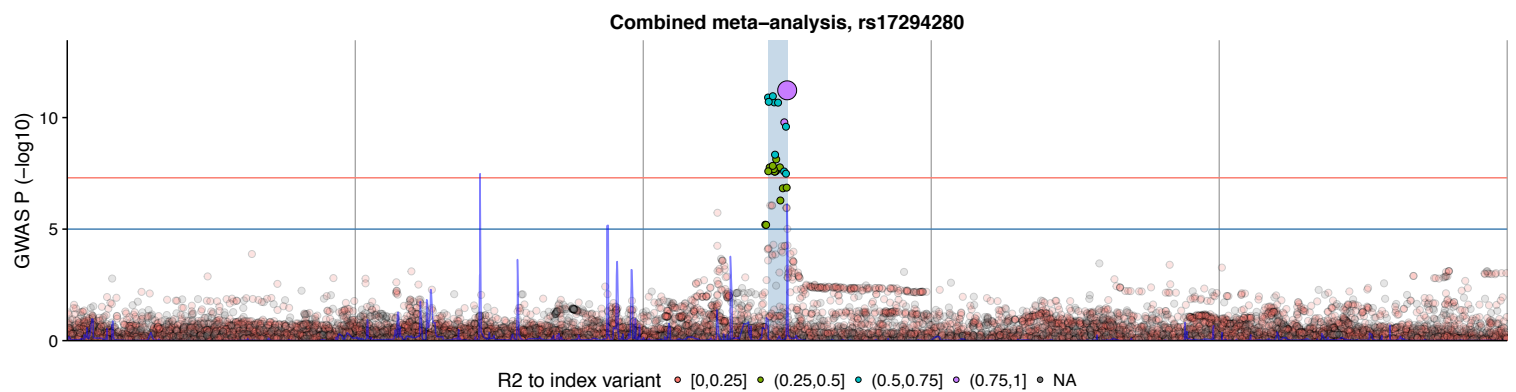
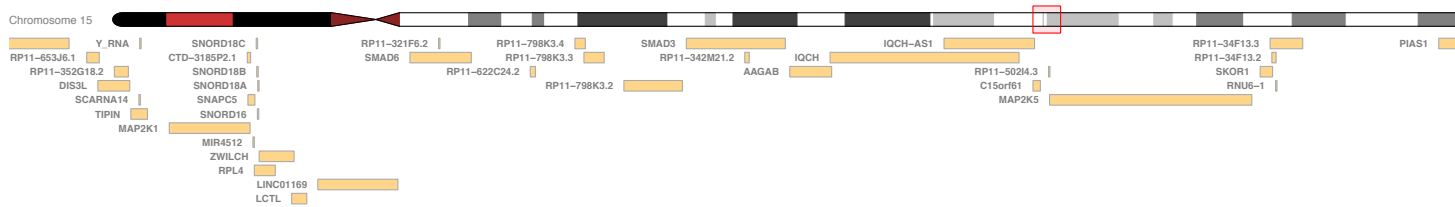


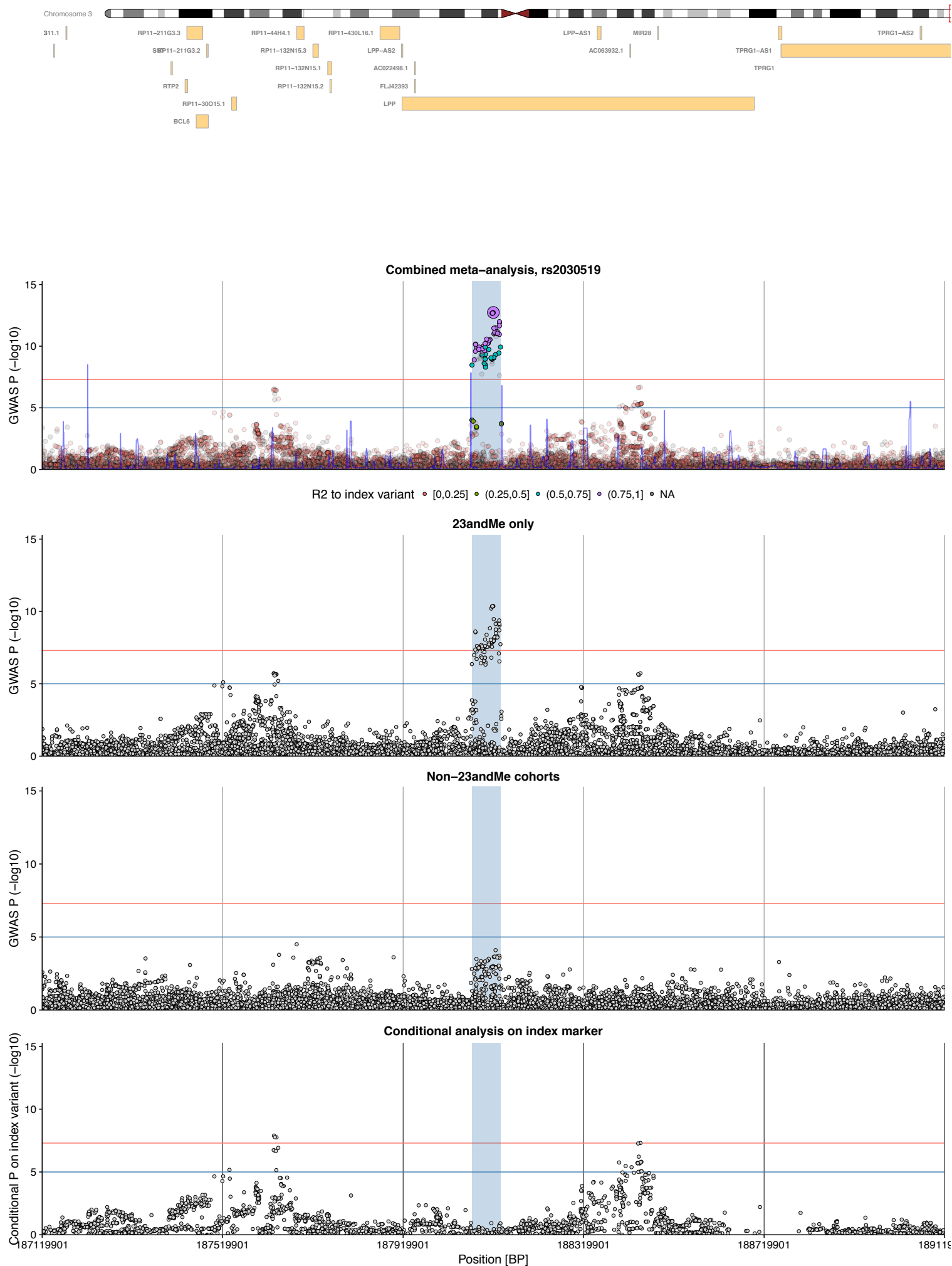


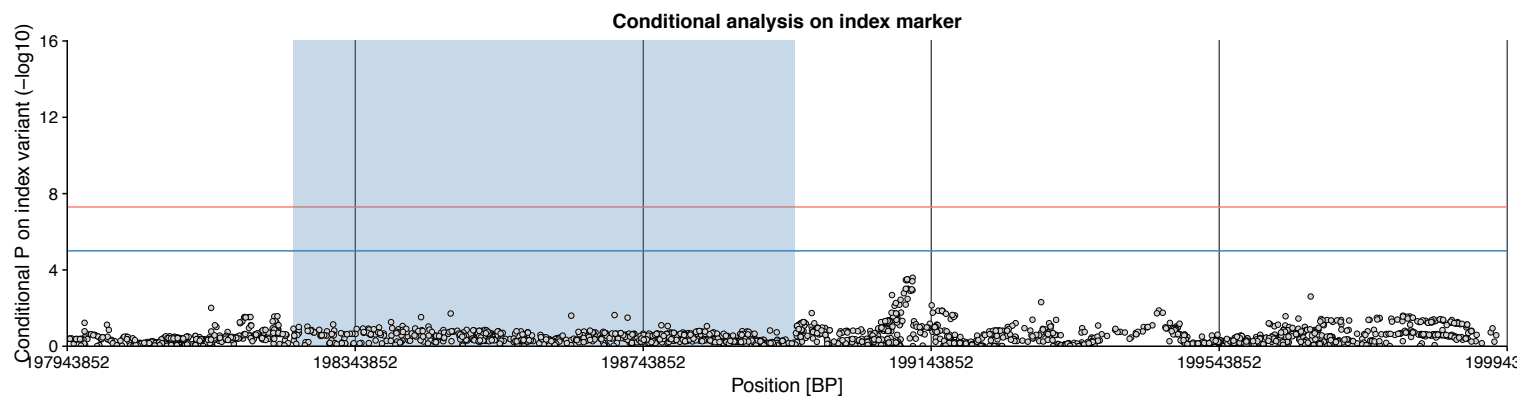
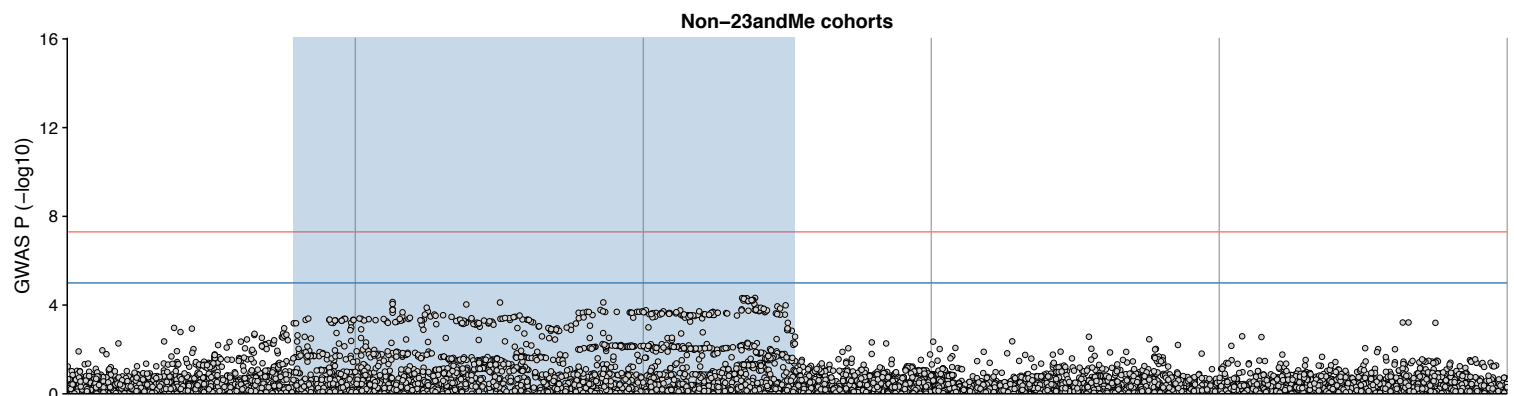
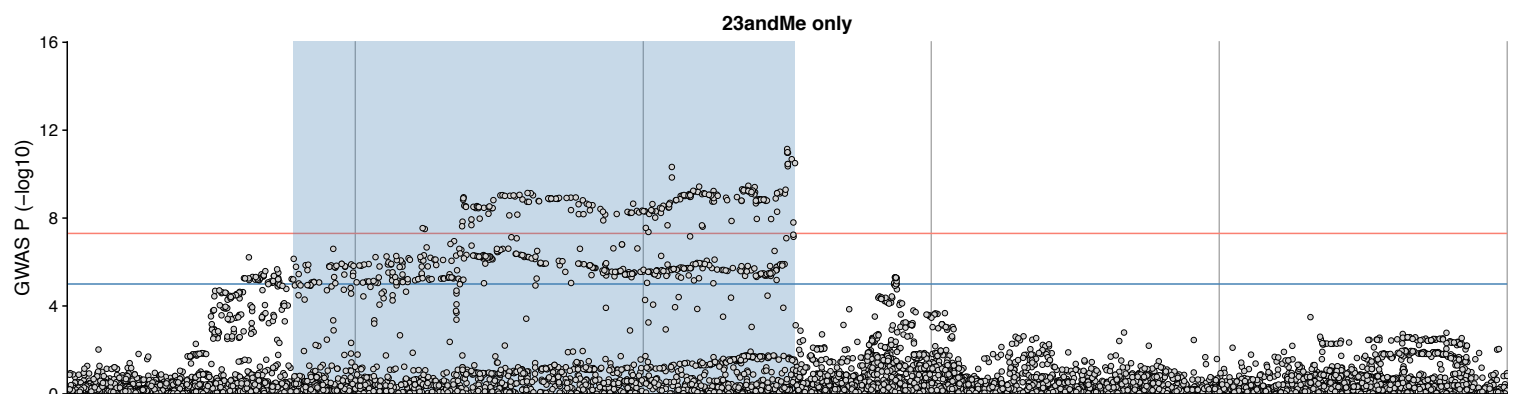
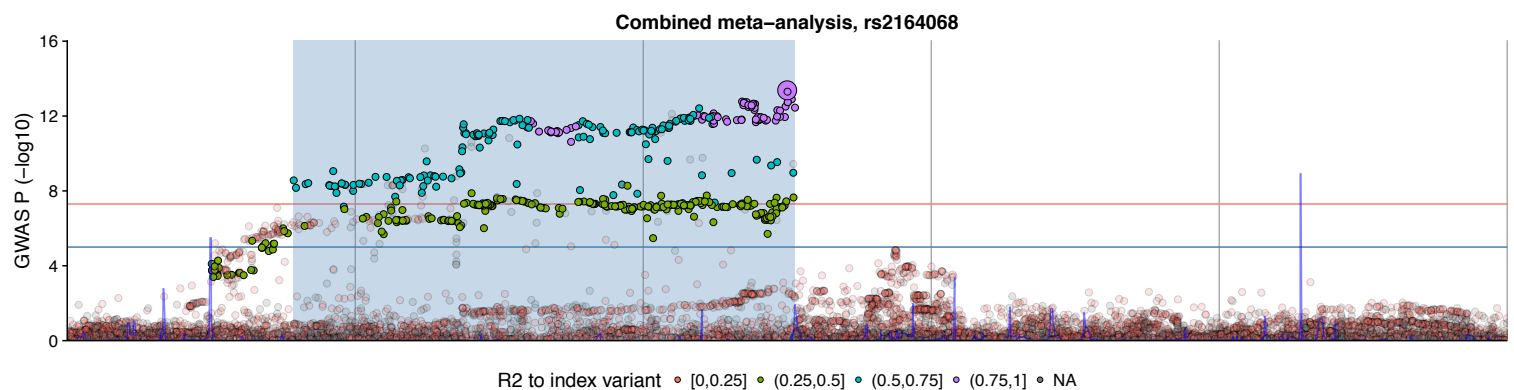
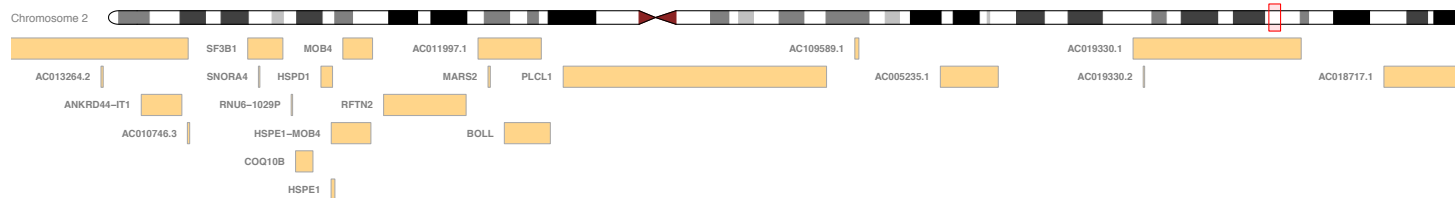


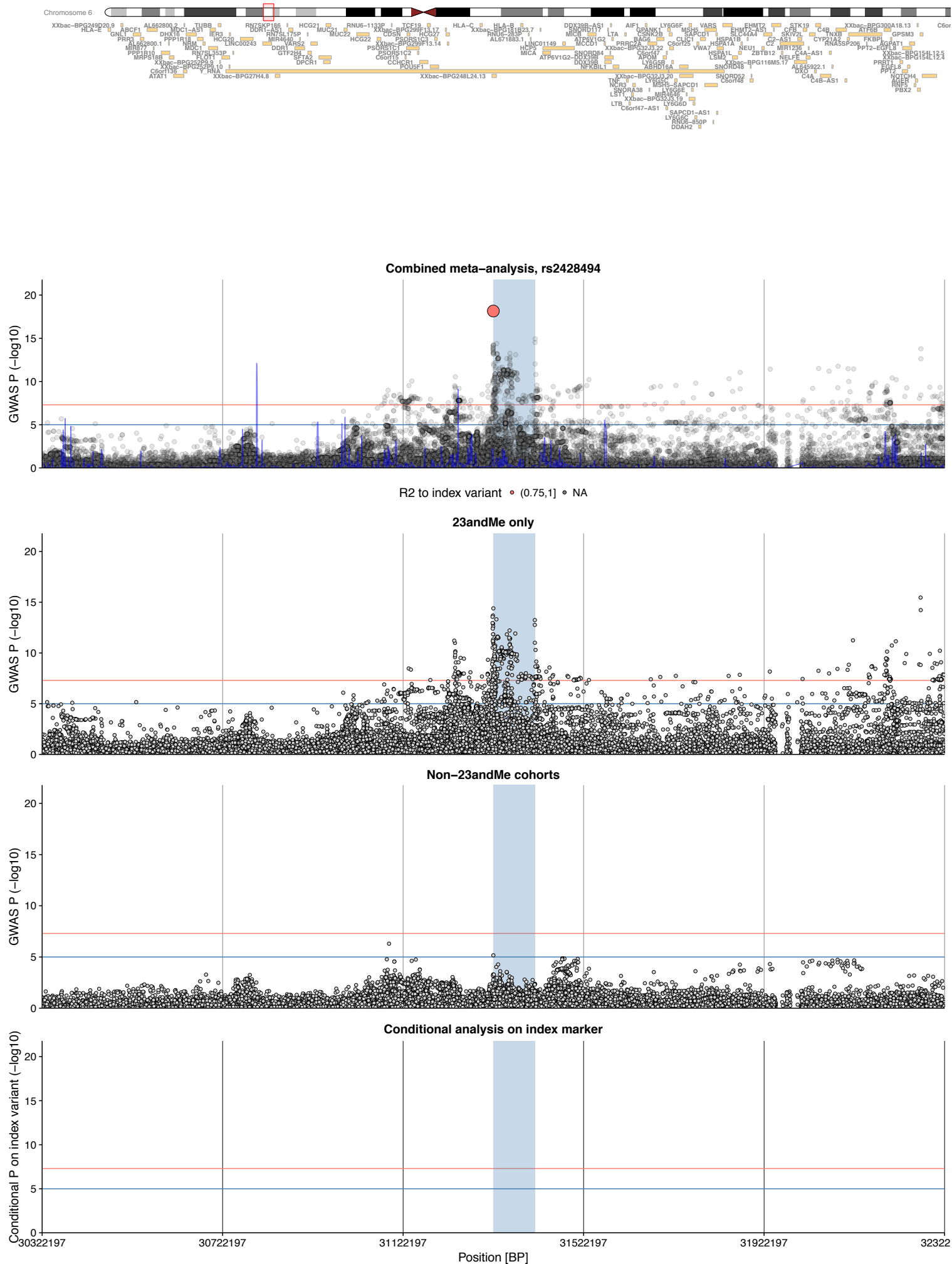


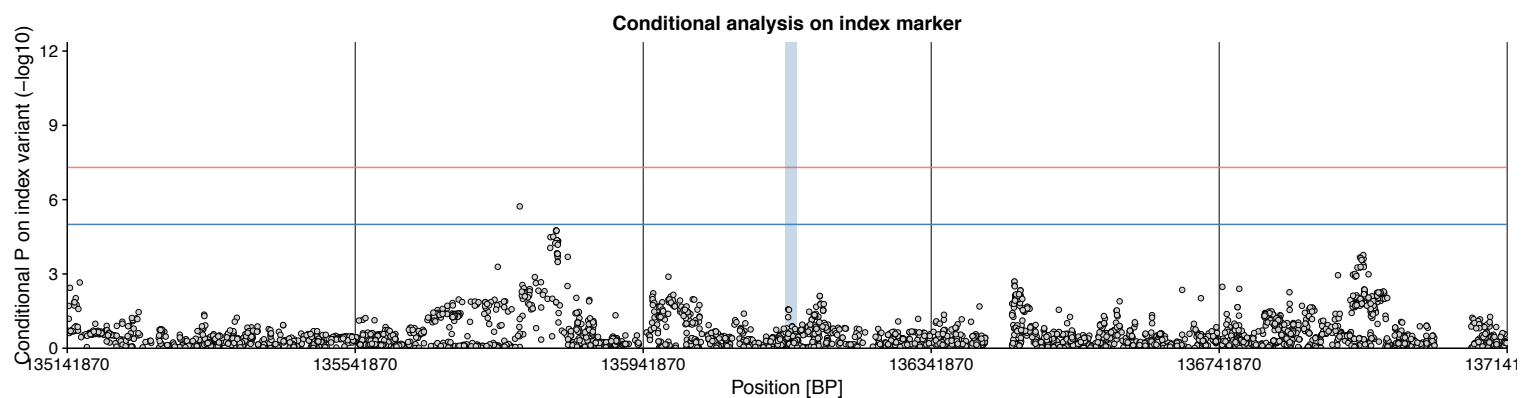
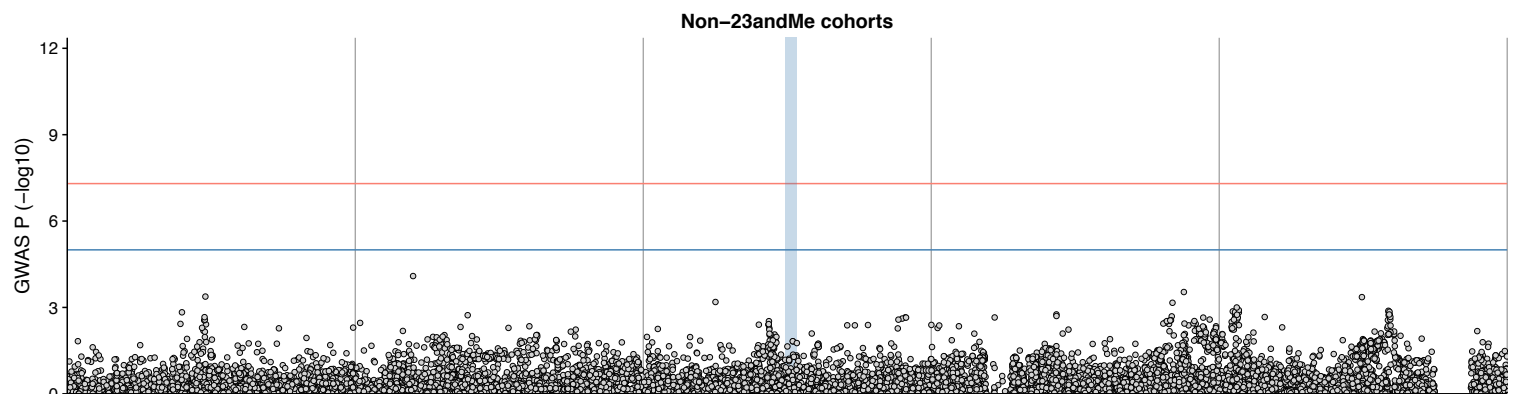
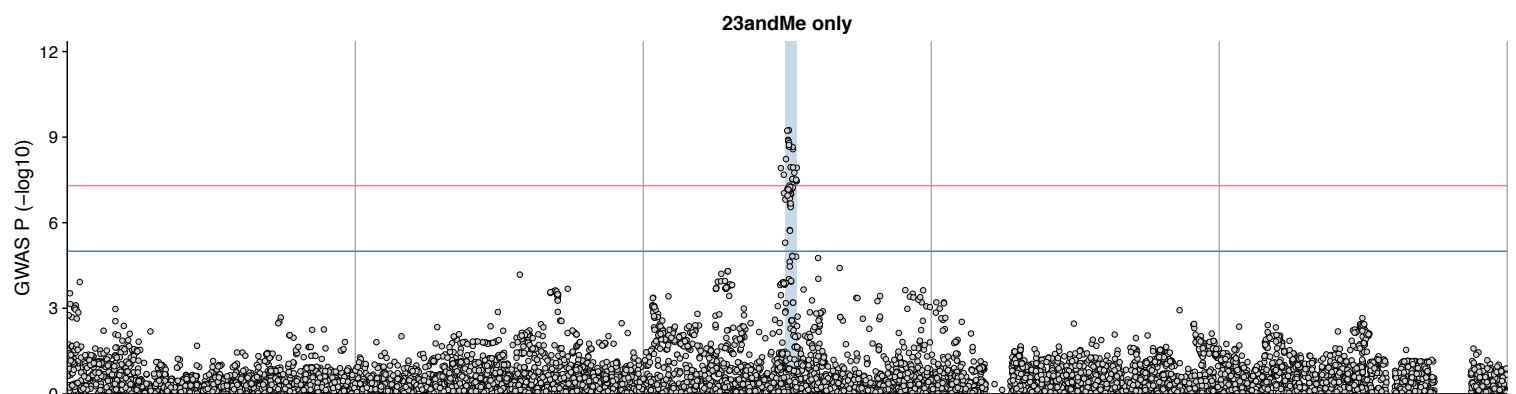
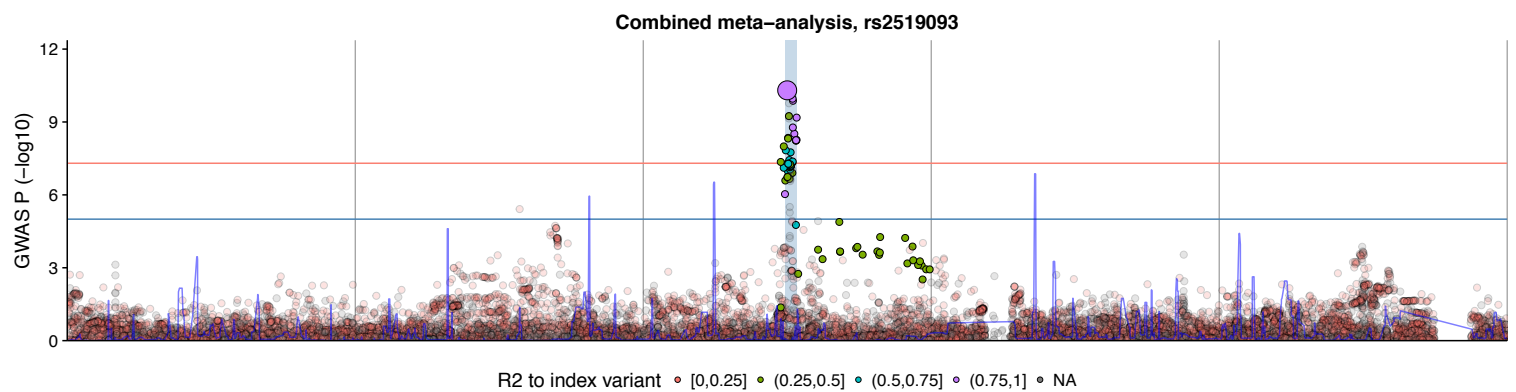
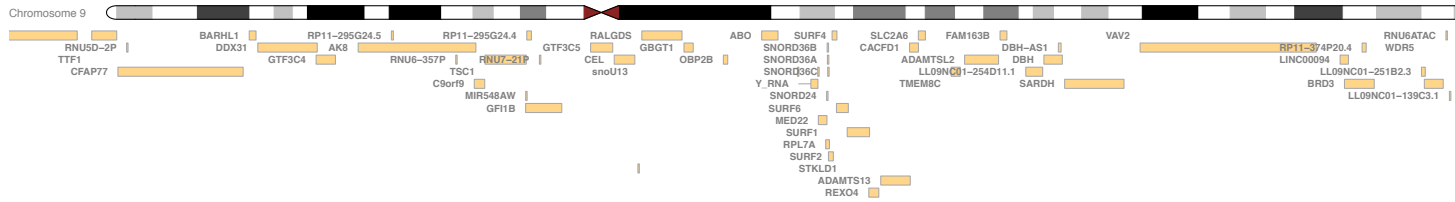


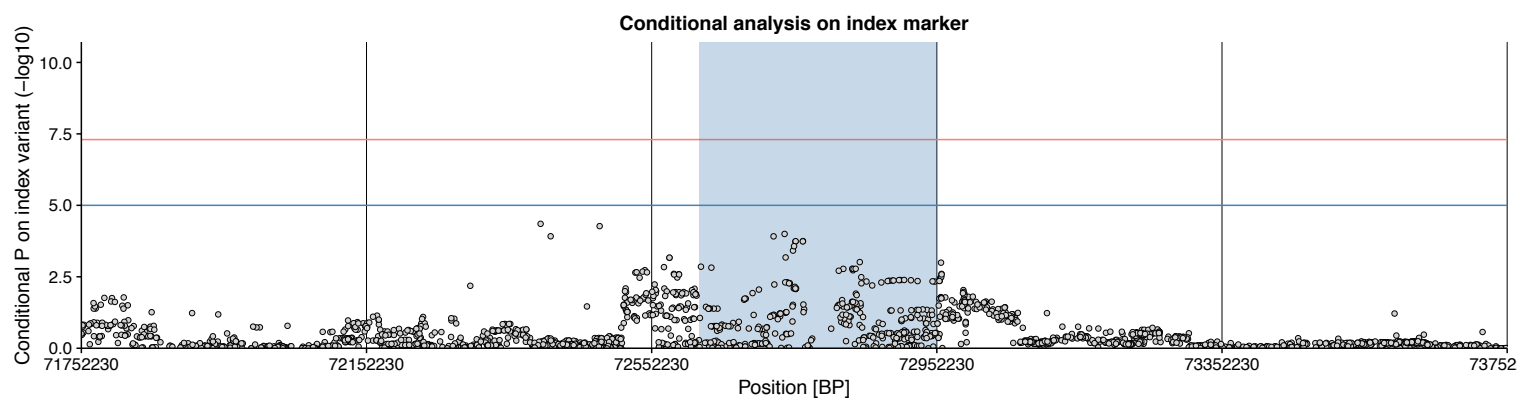
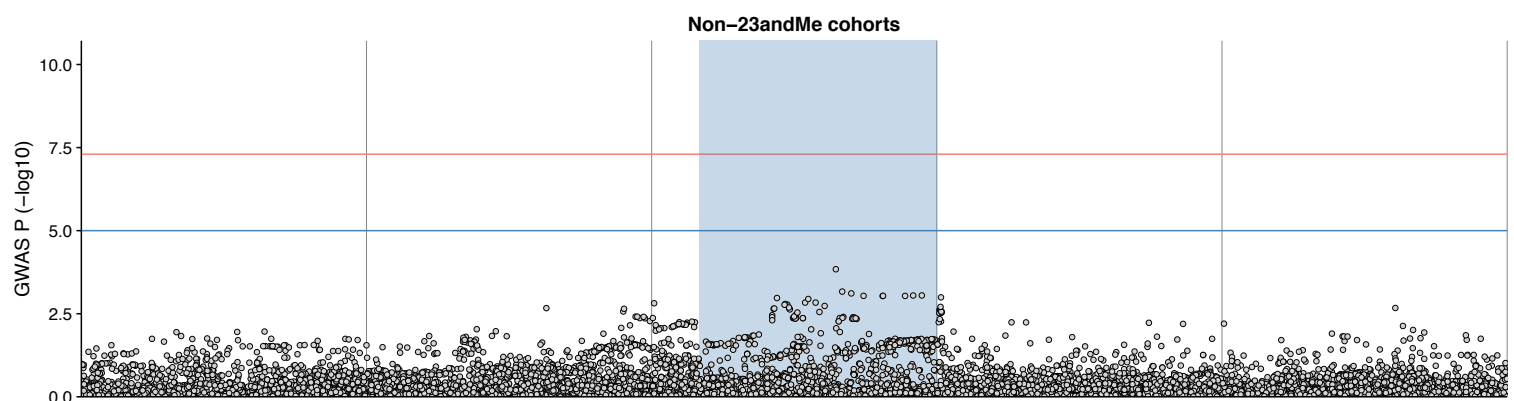
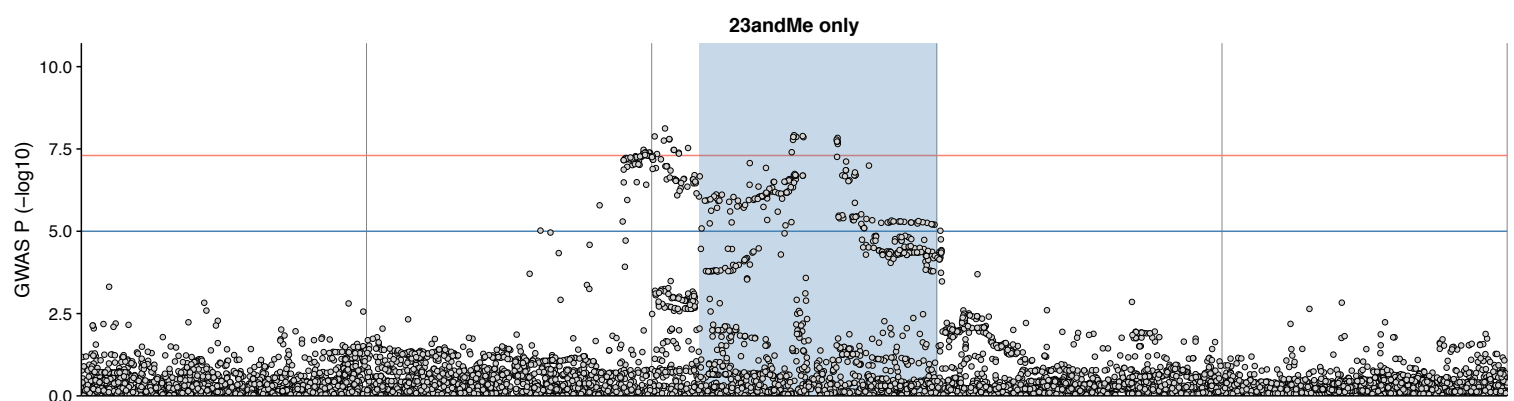
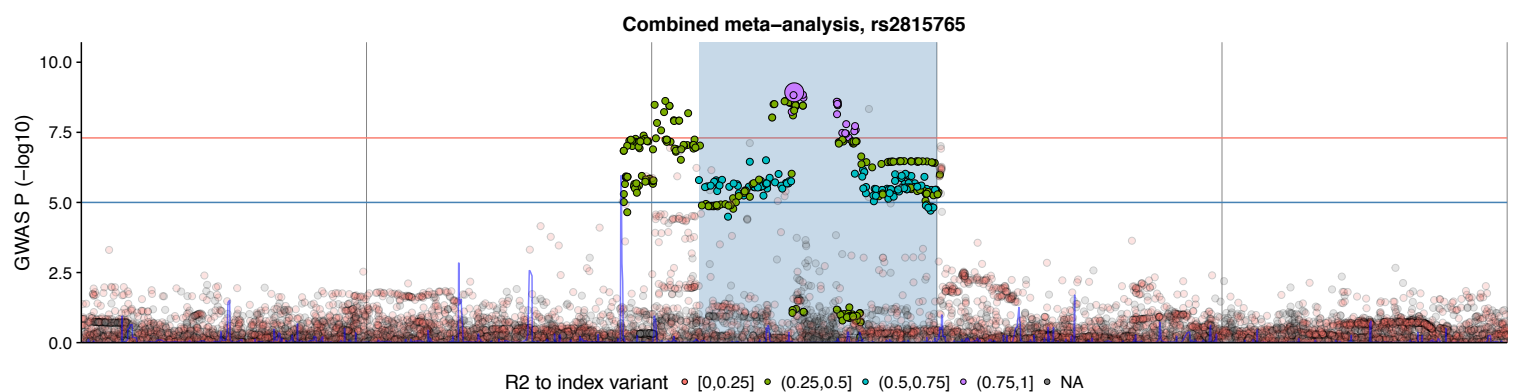
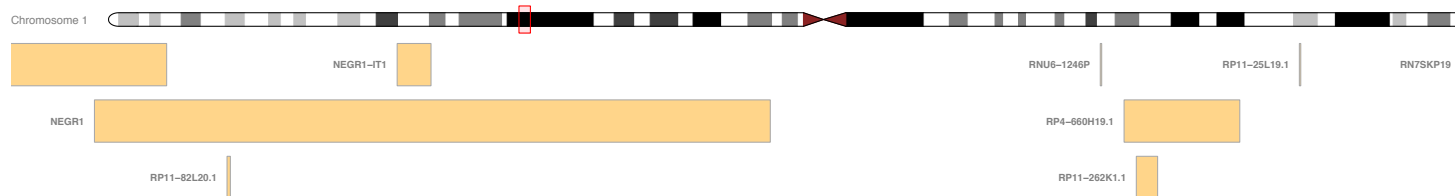


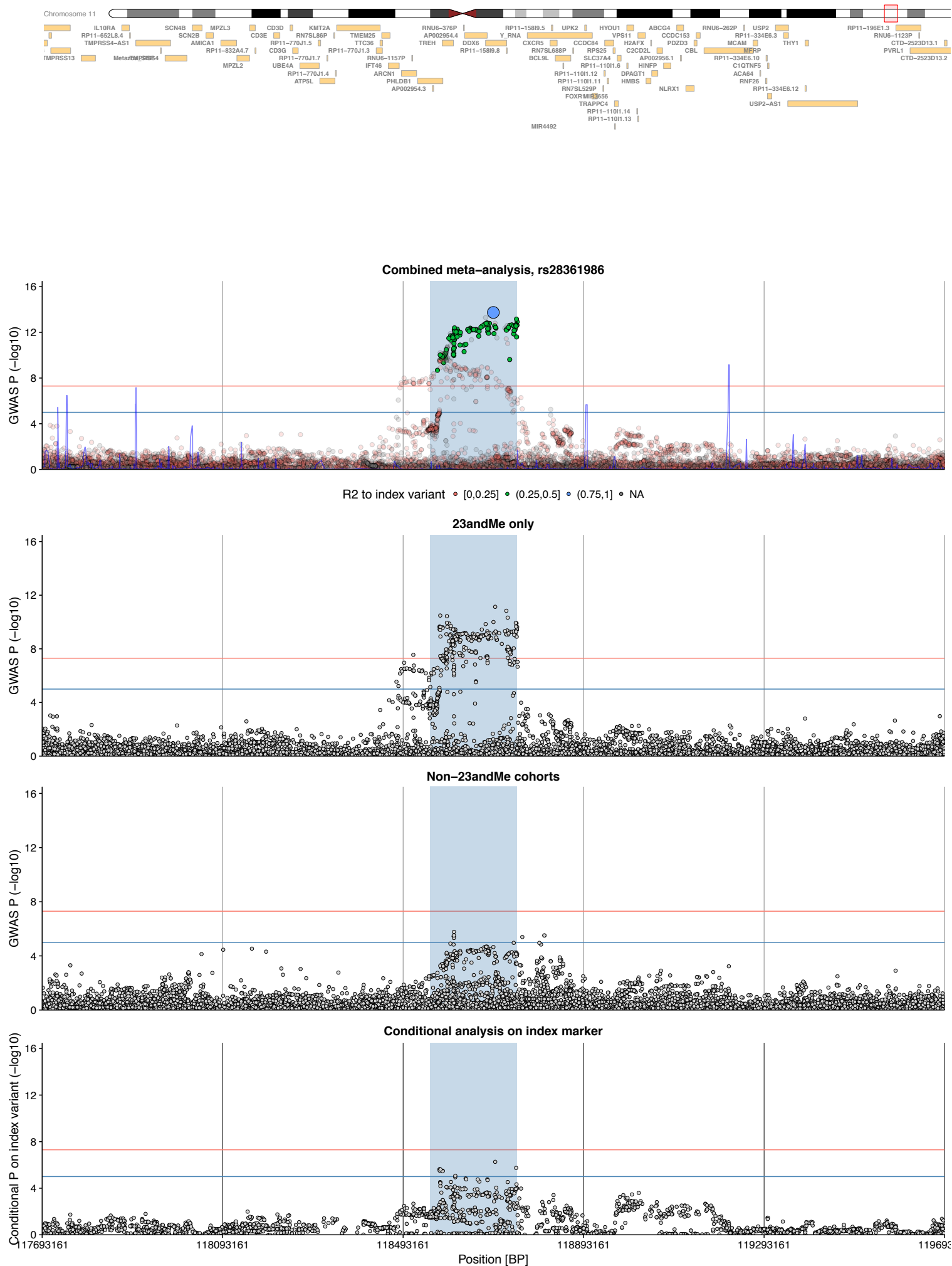


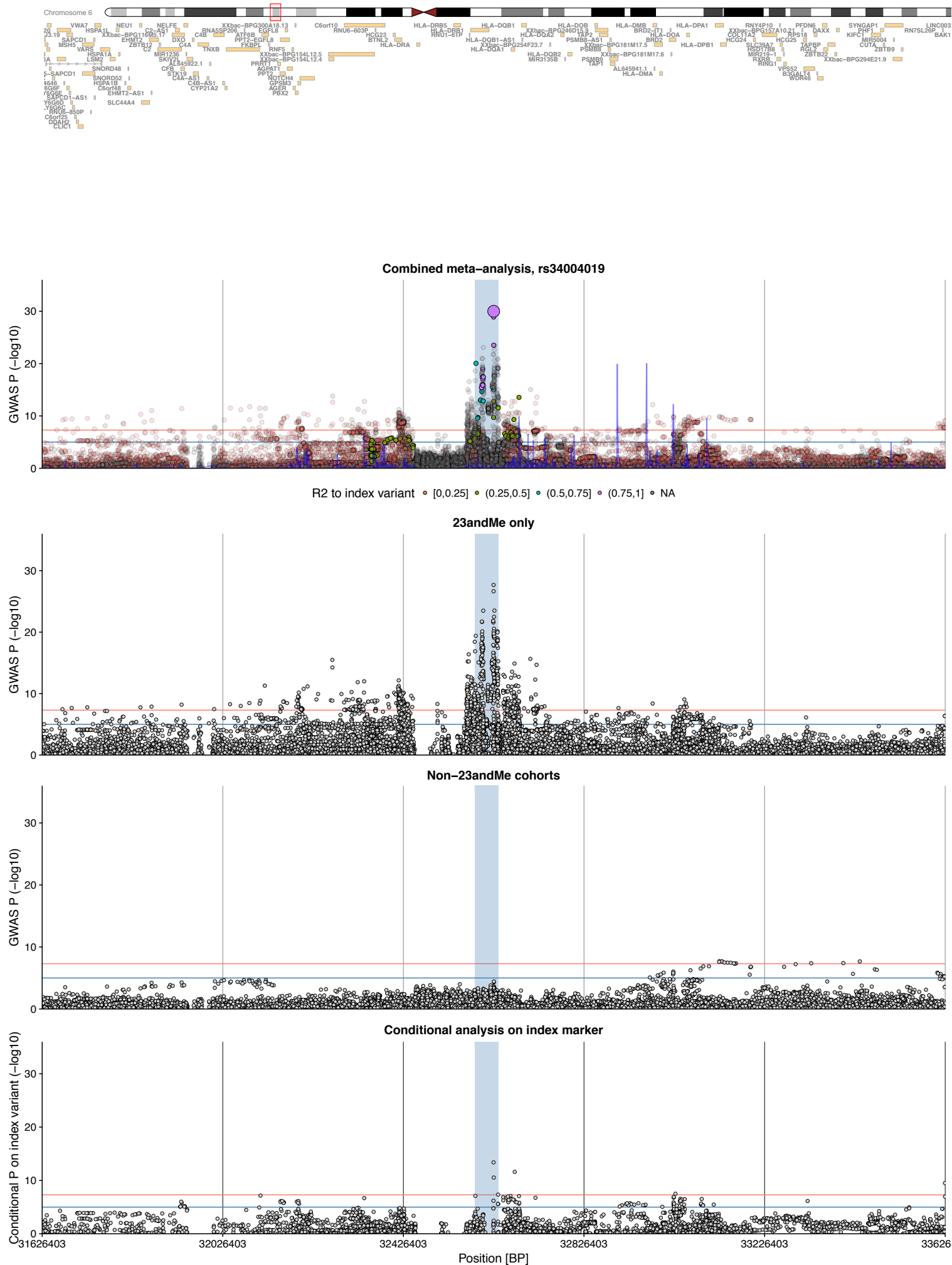


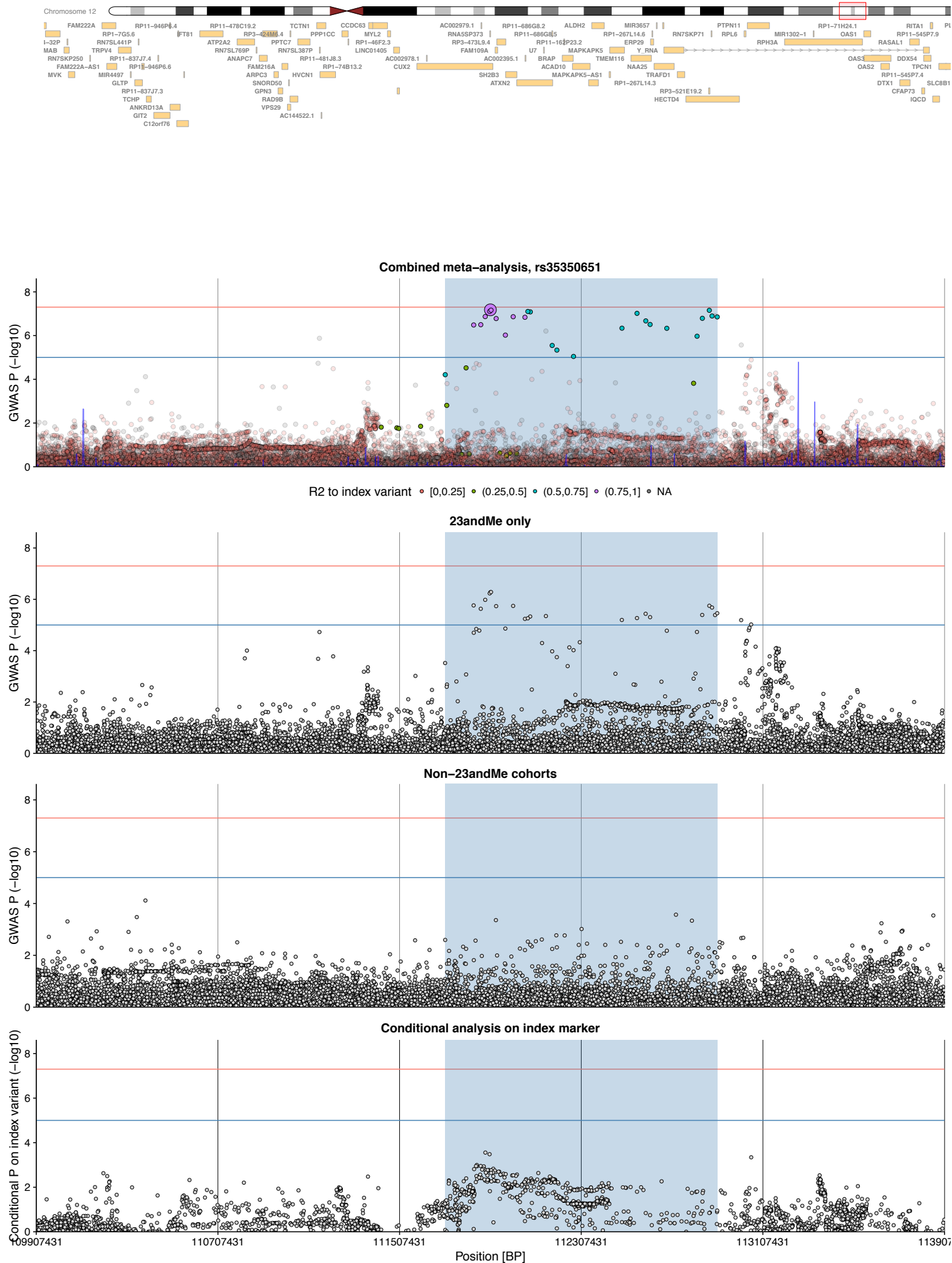


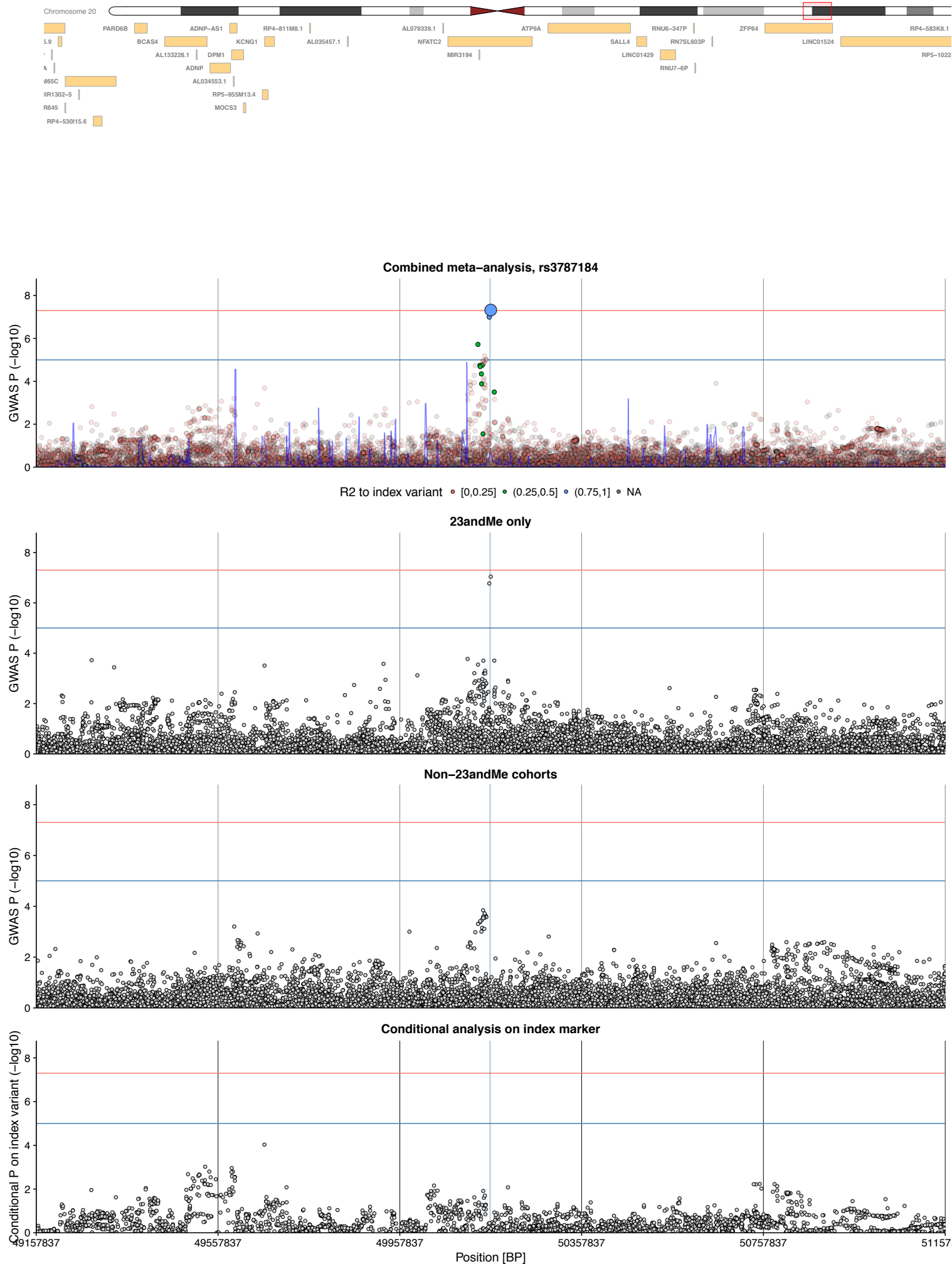


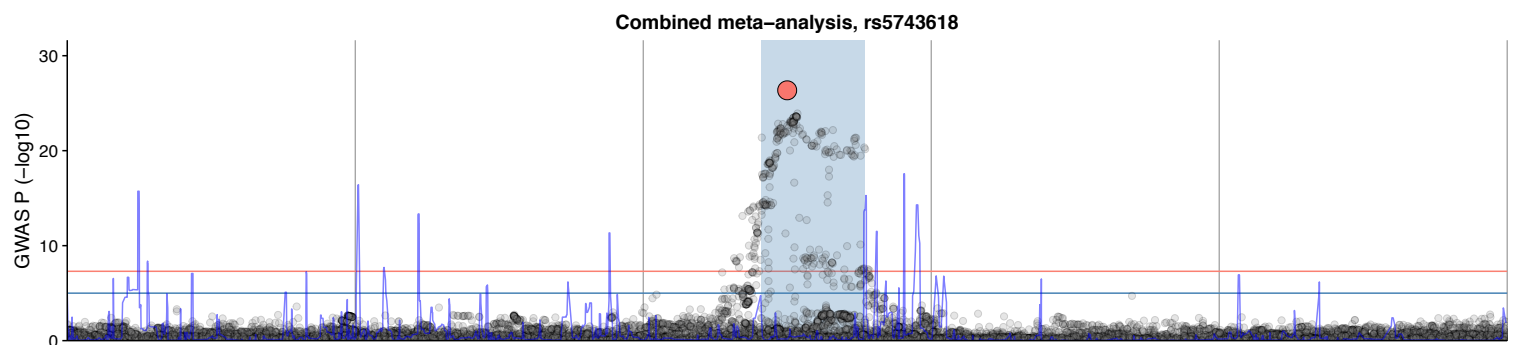
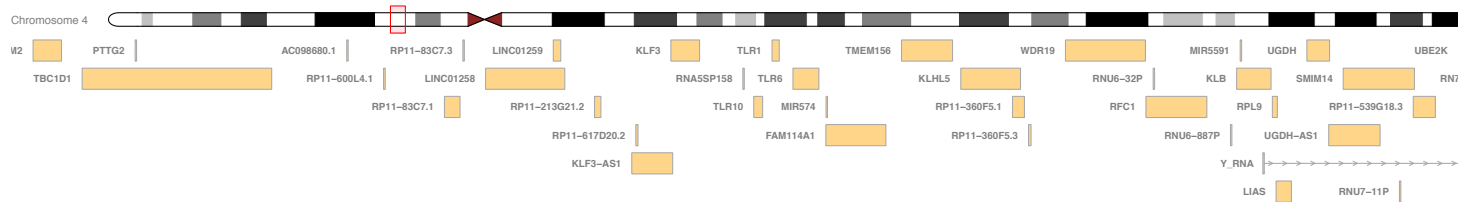




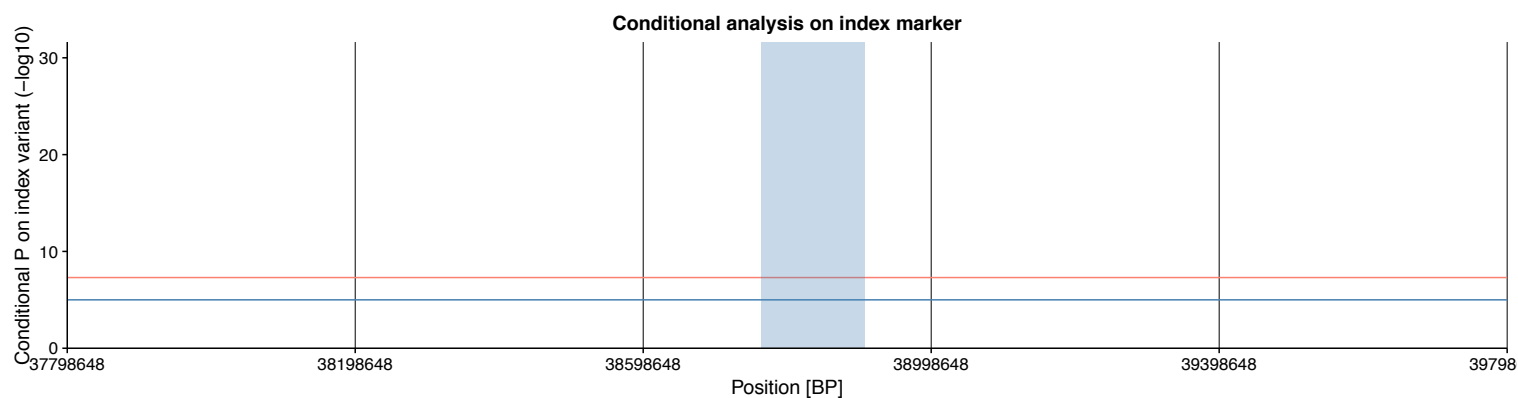
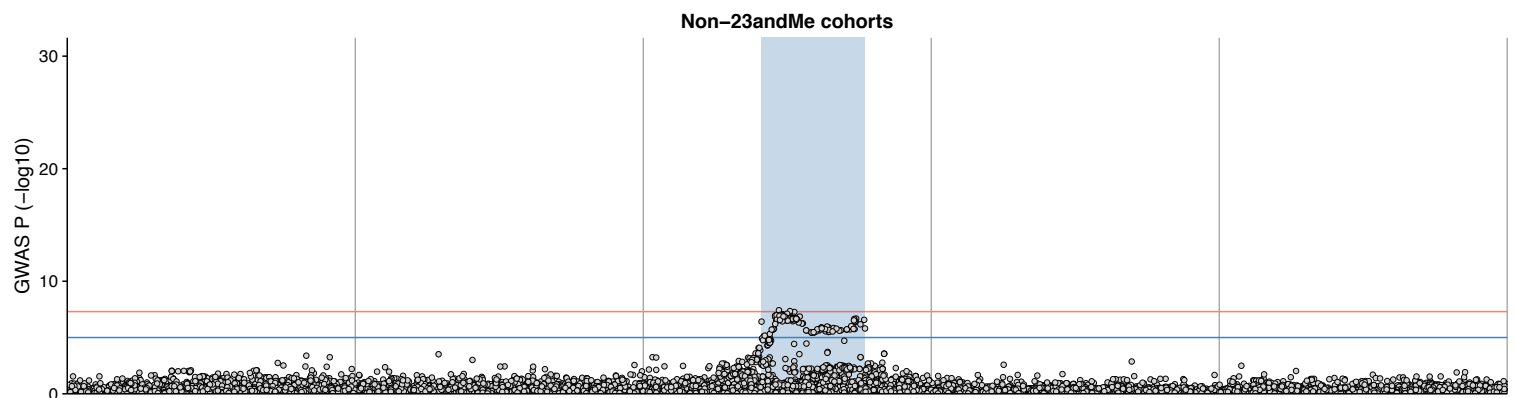
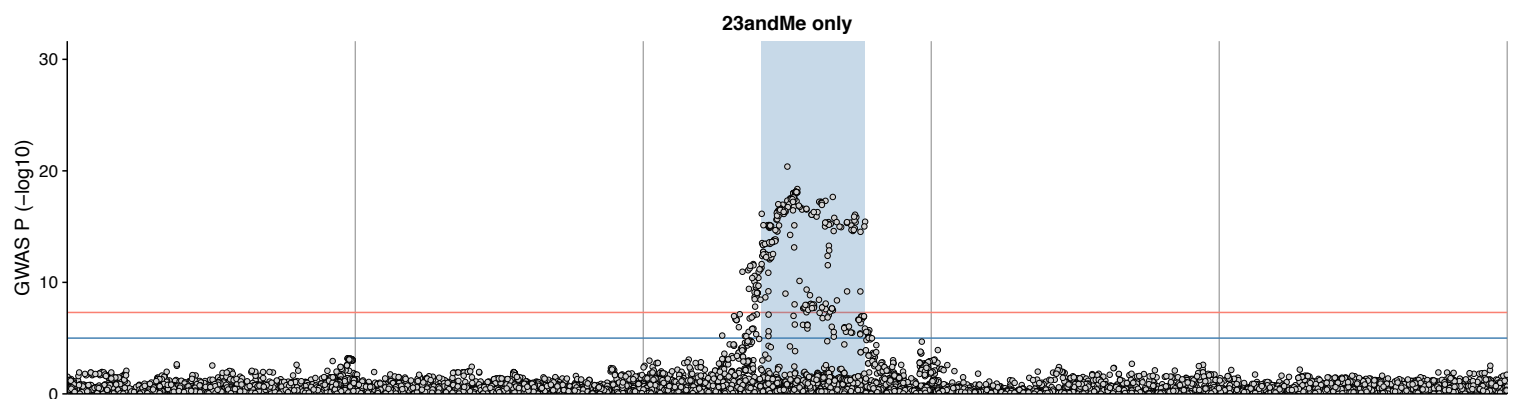


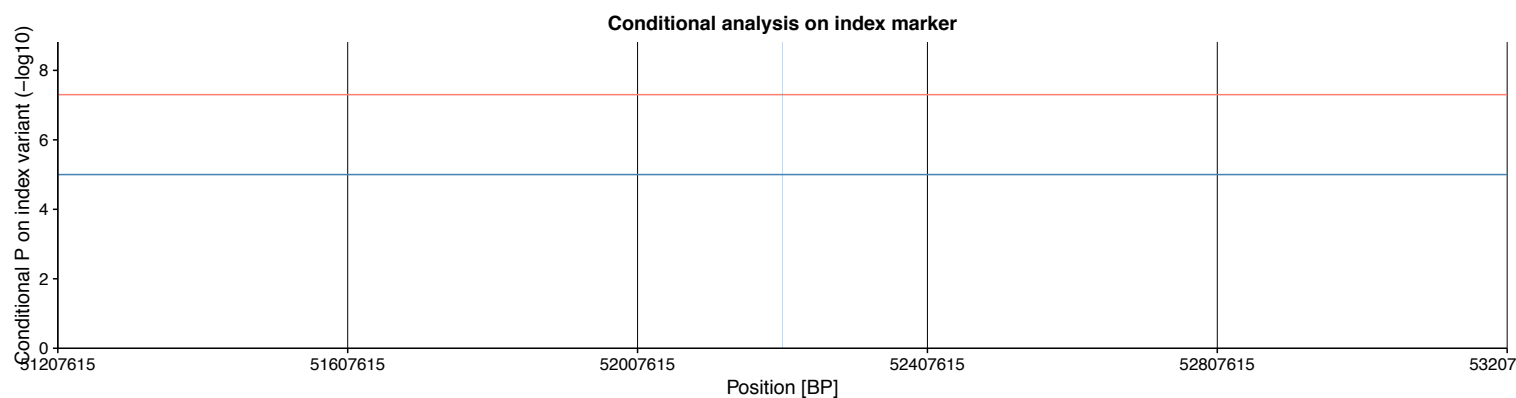
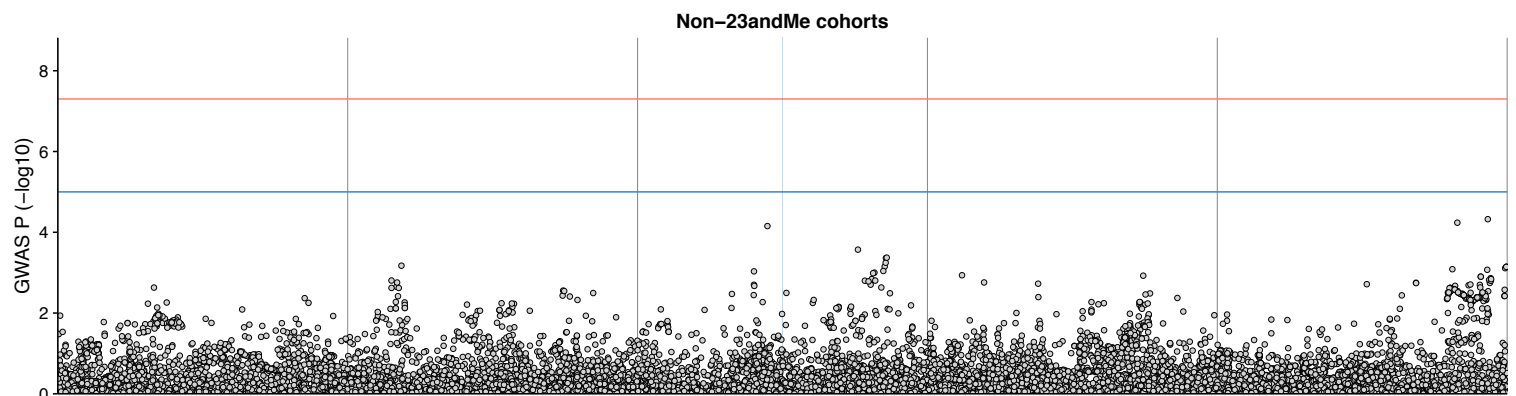
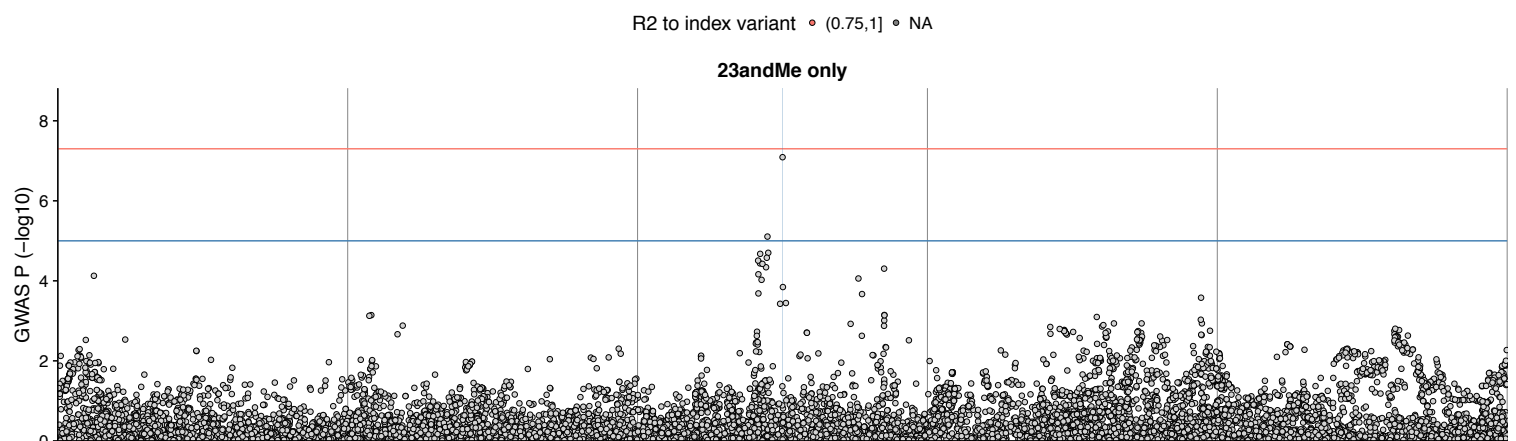
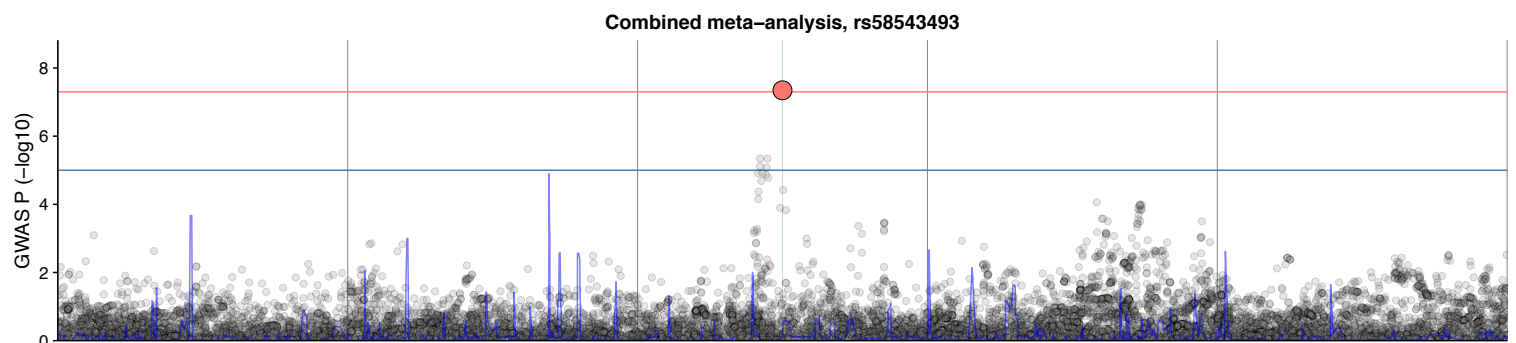
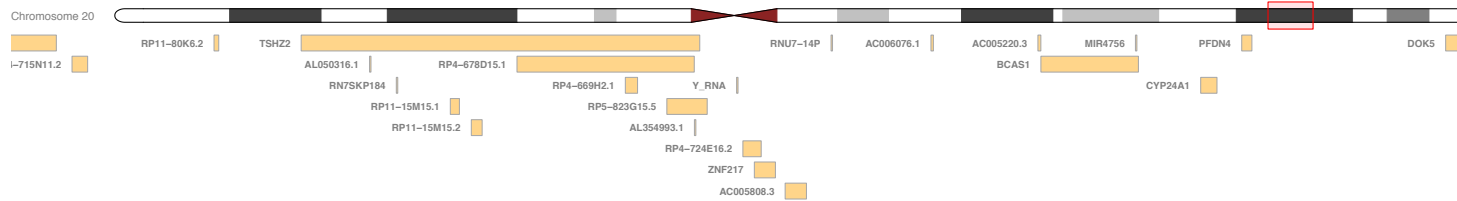


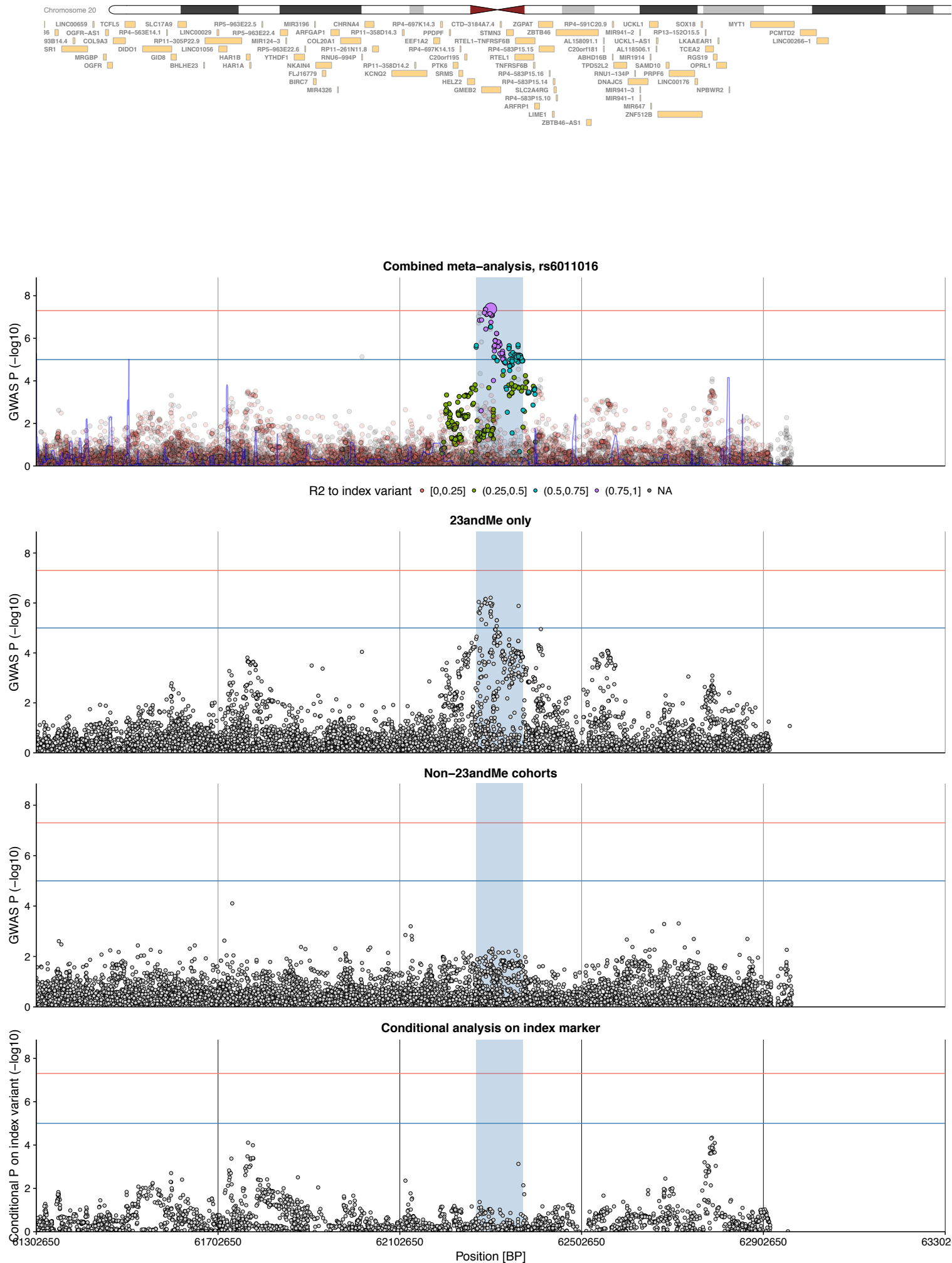


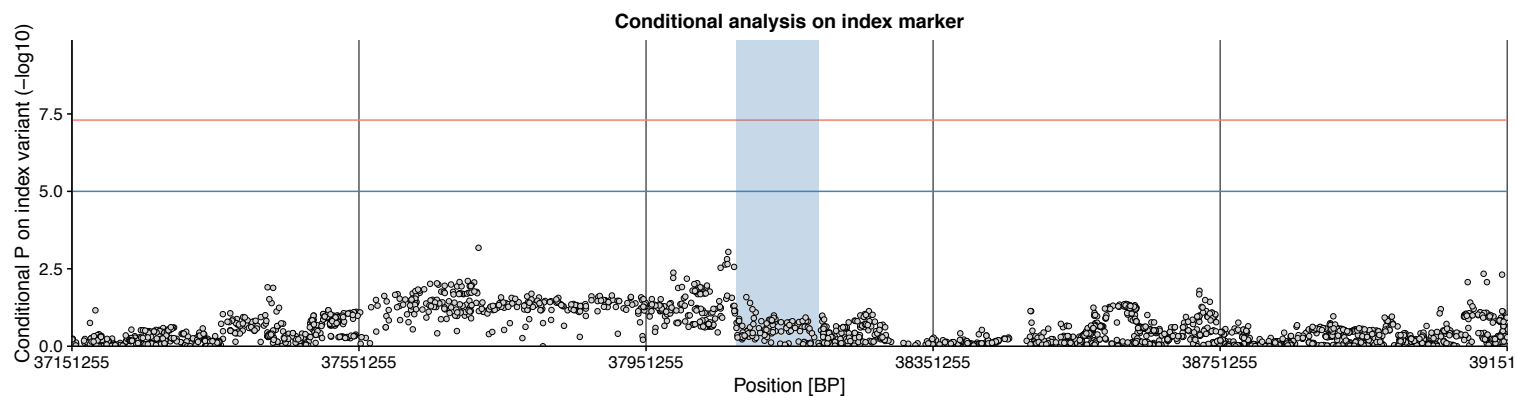
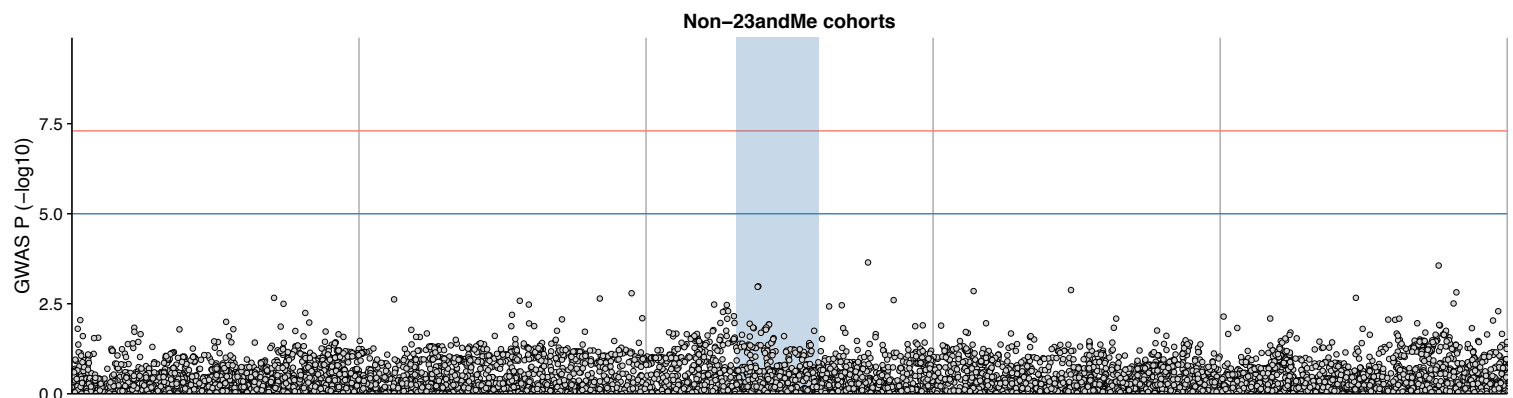
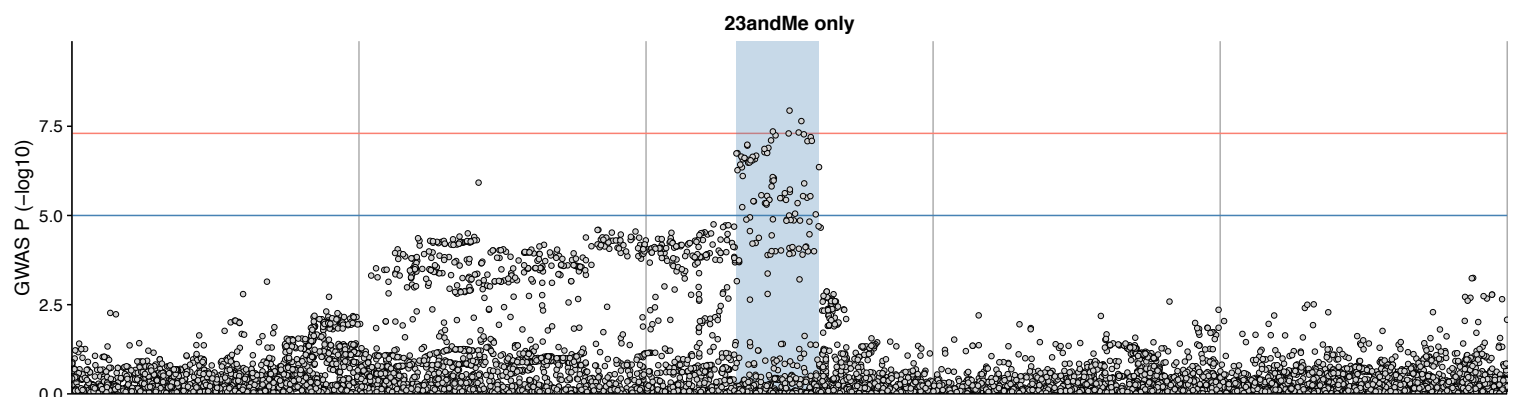
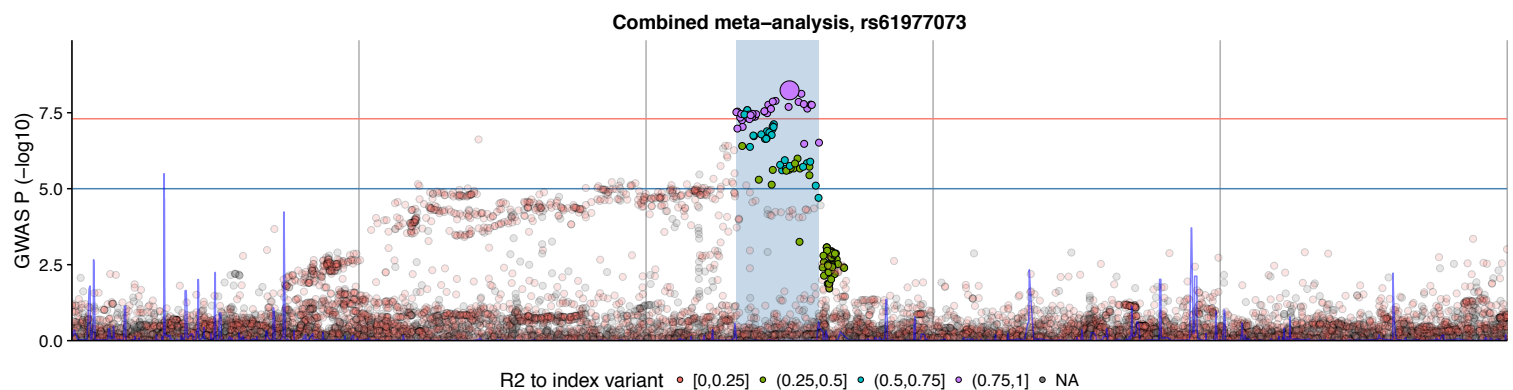
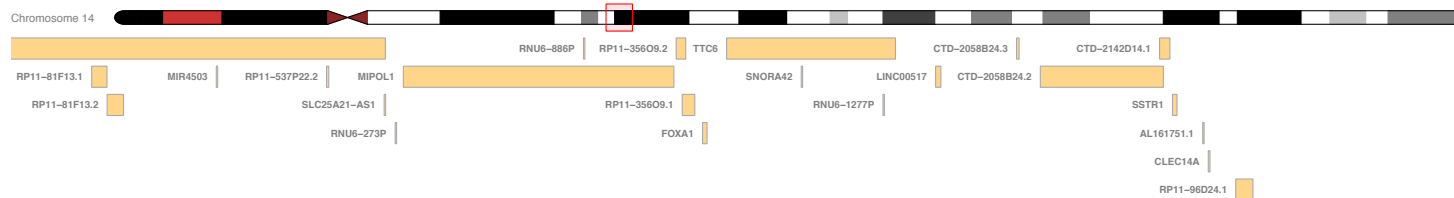


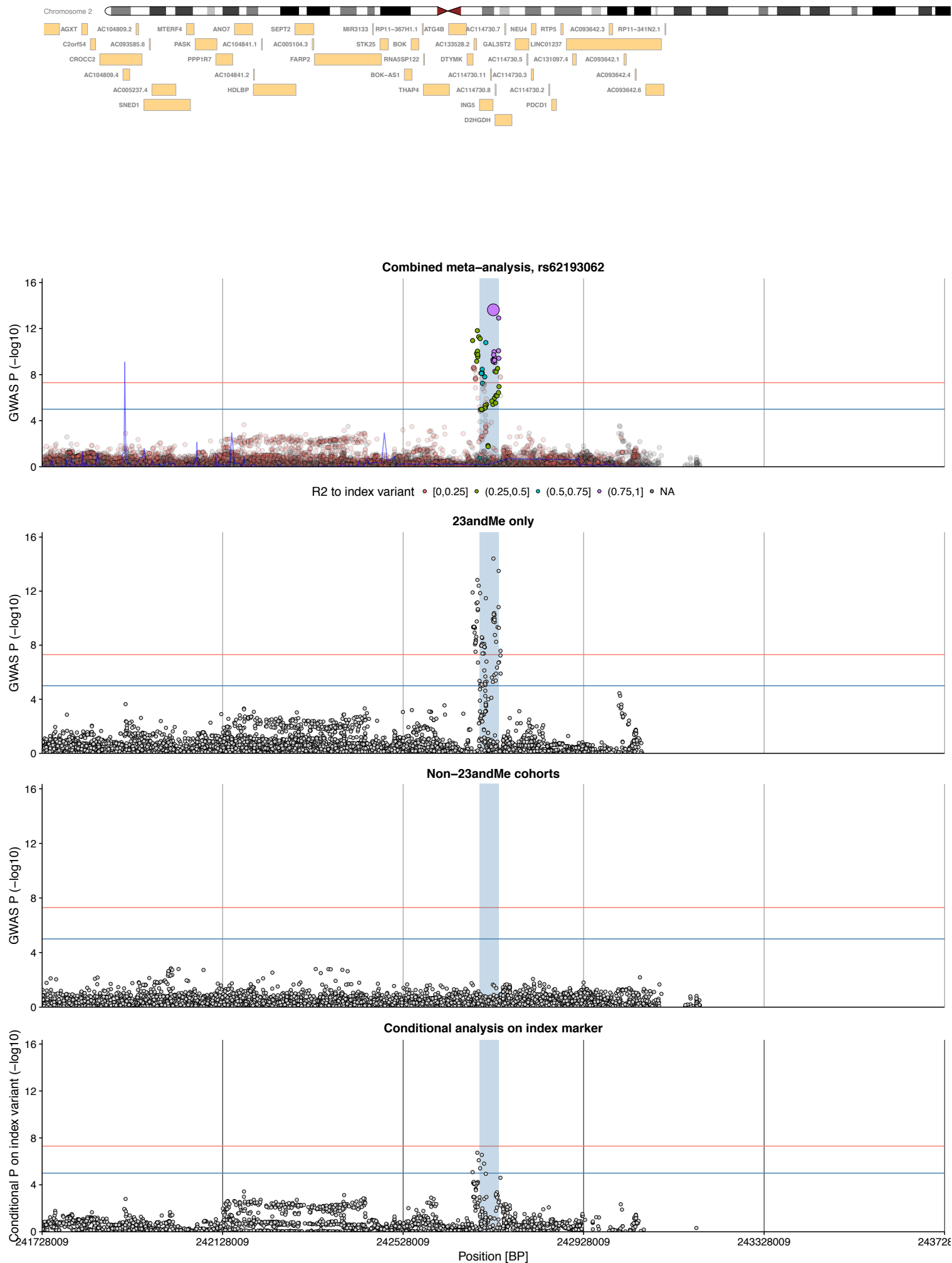
R2 to index variant • (0.75,1] • NA

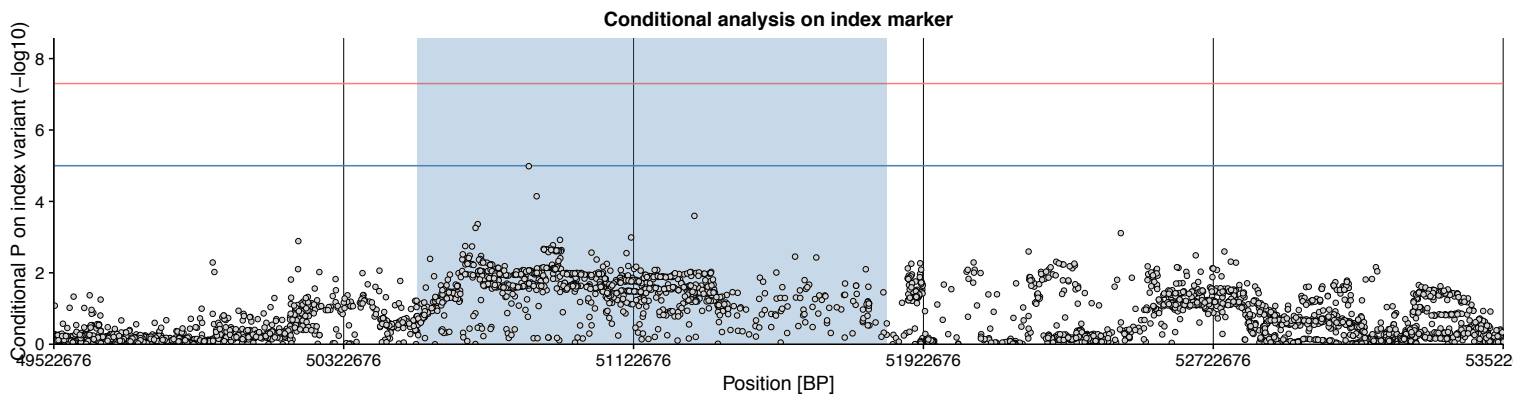
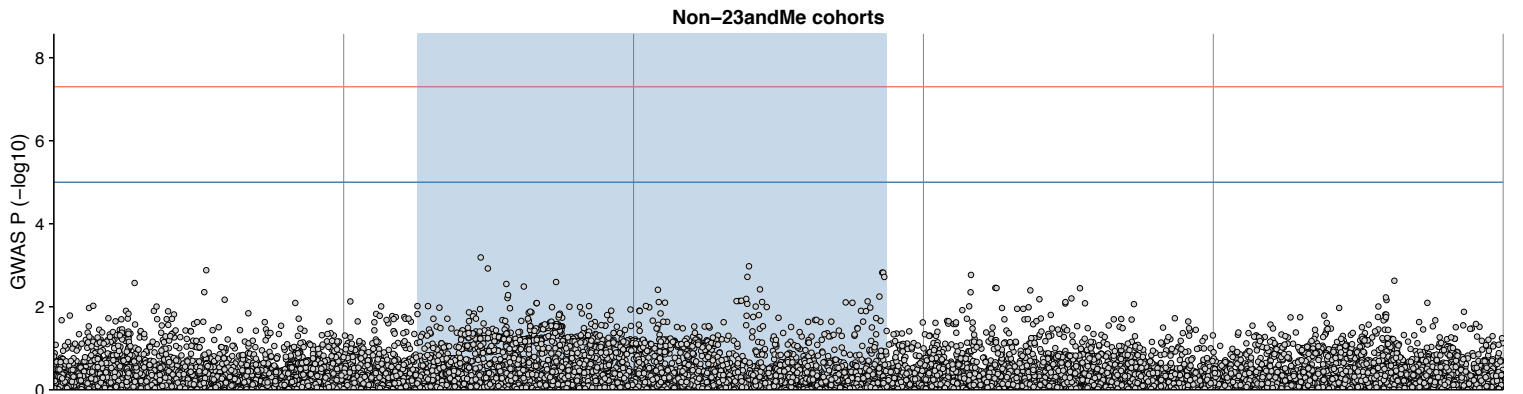
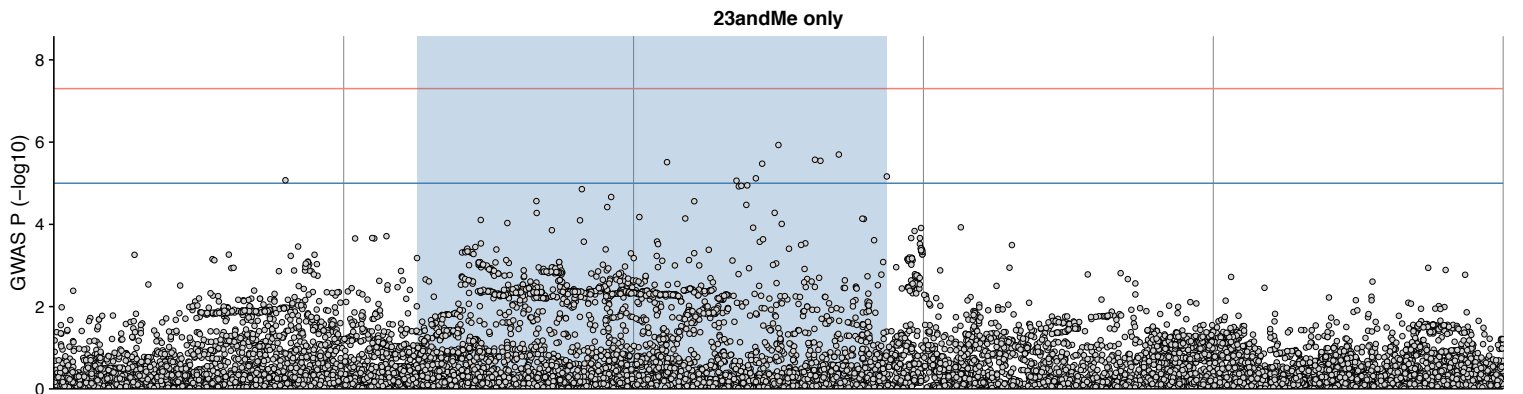
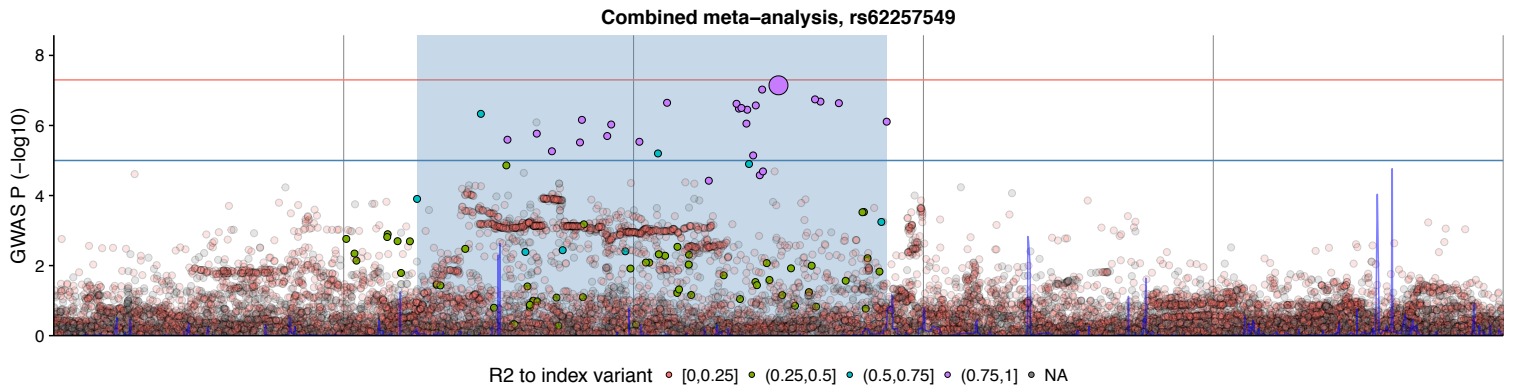
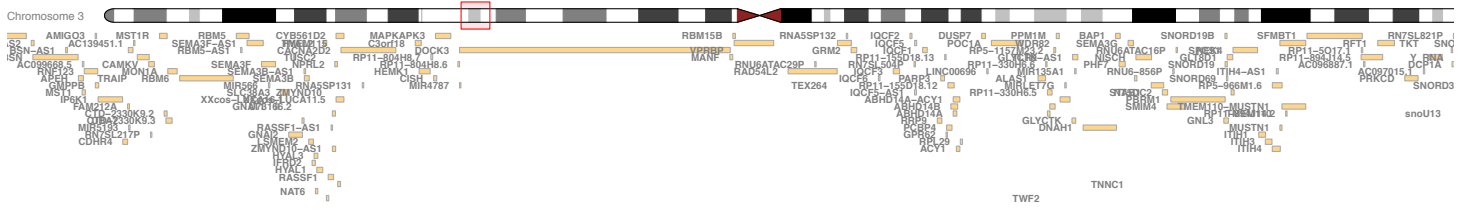


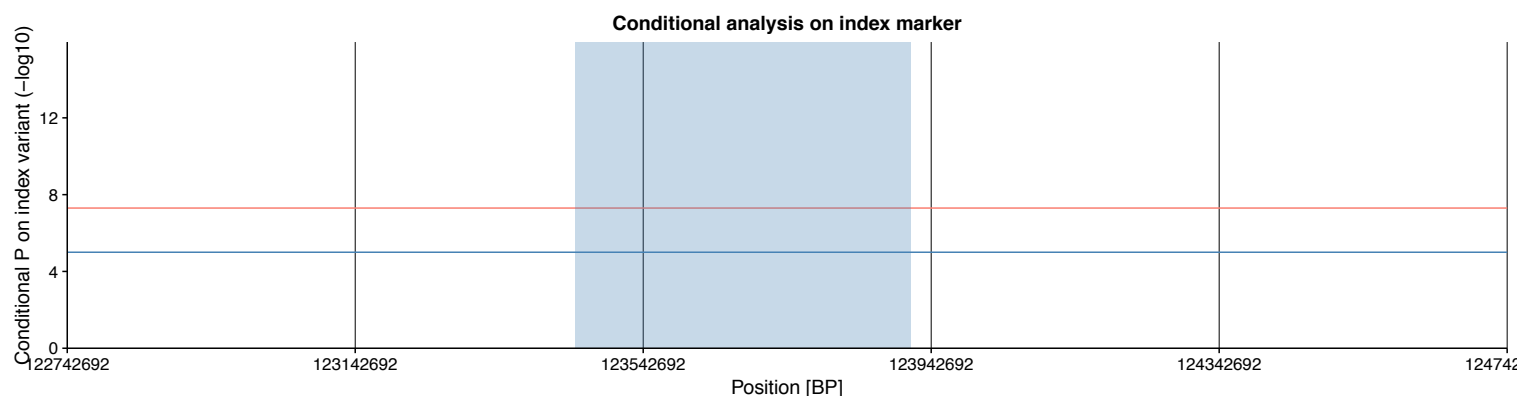
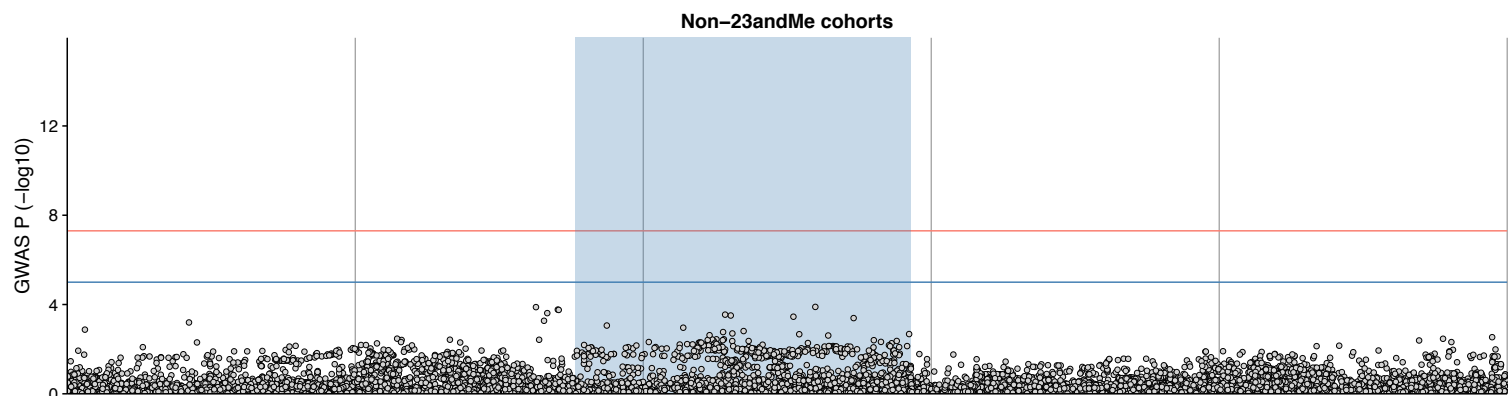
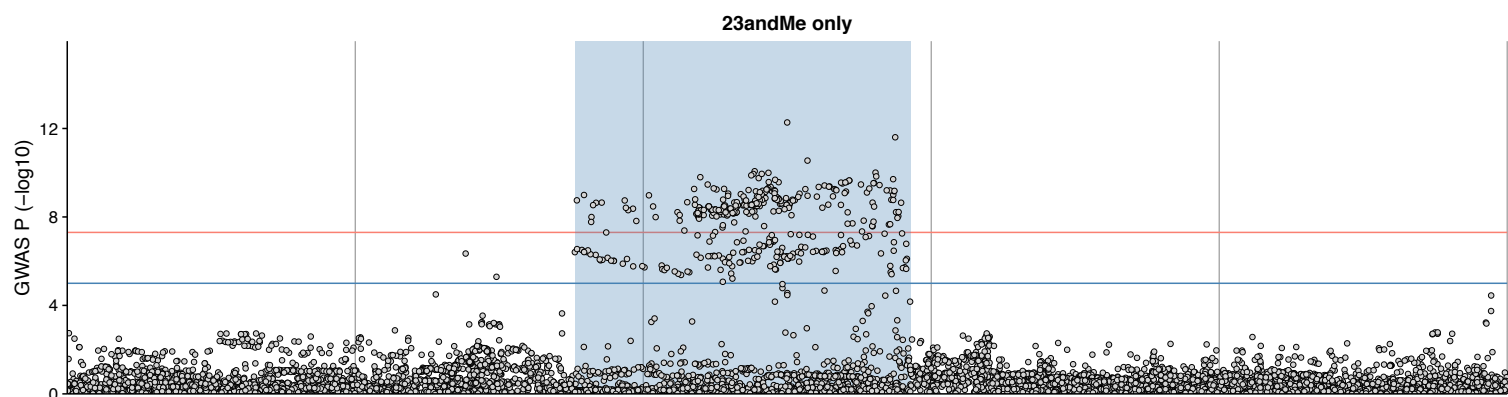
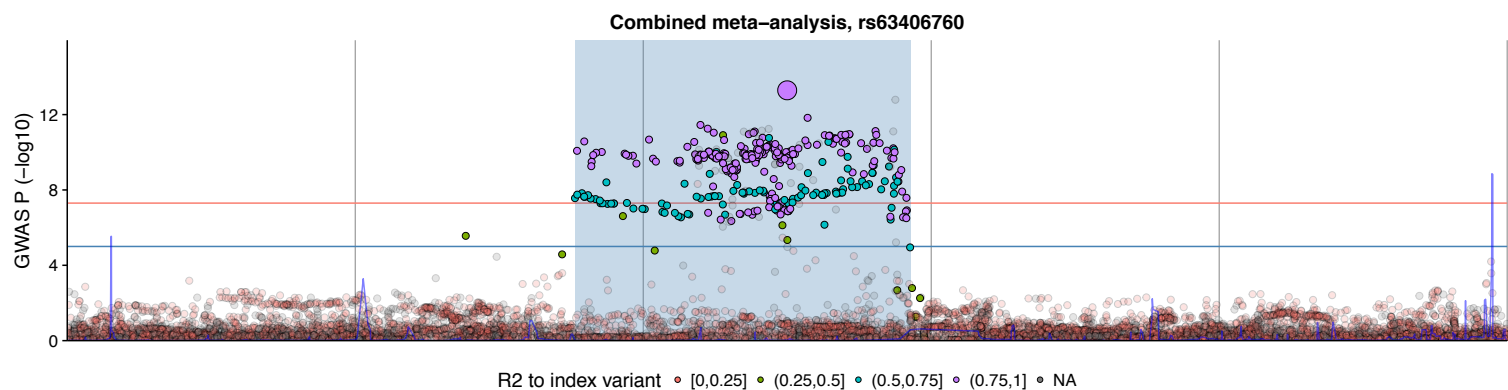
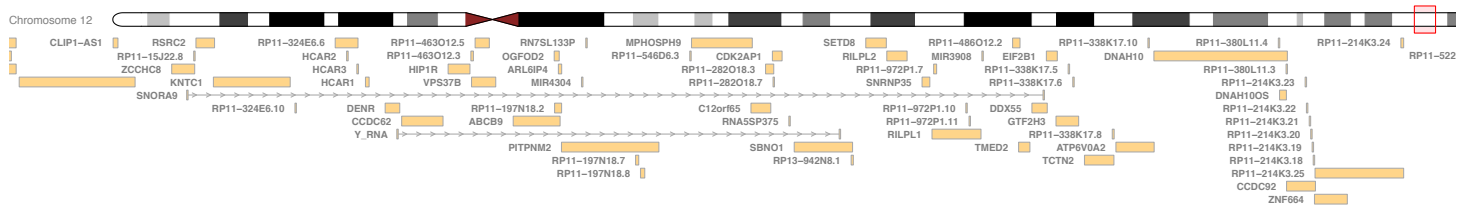


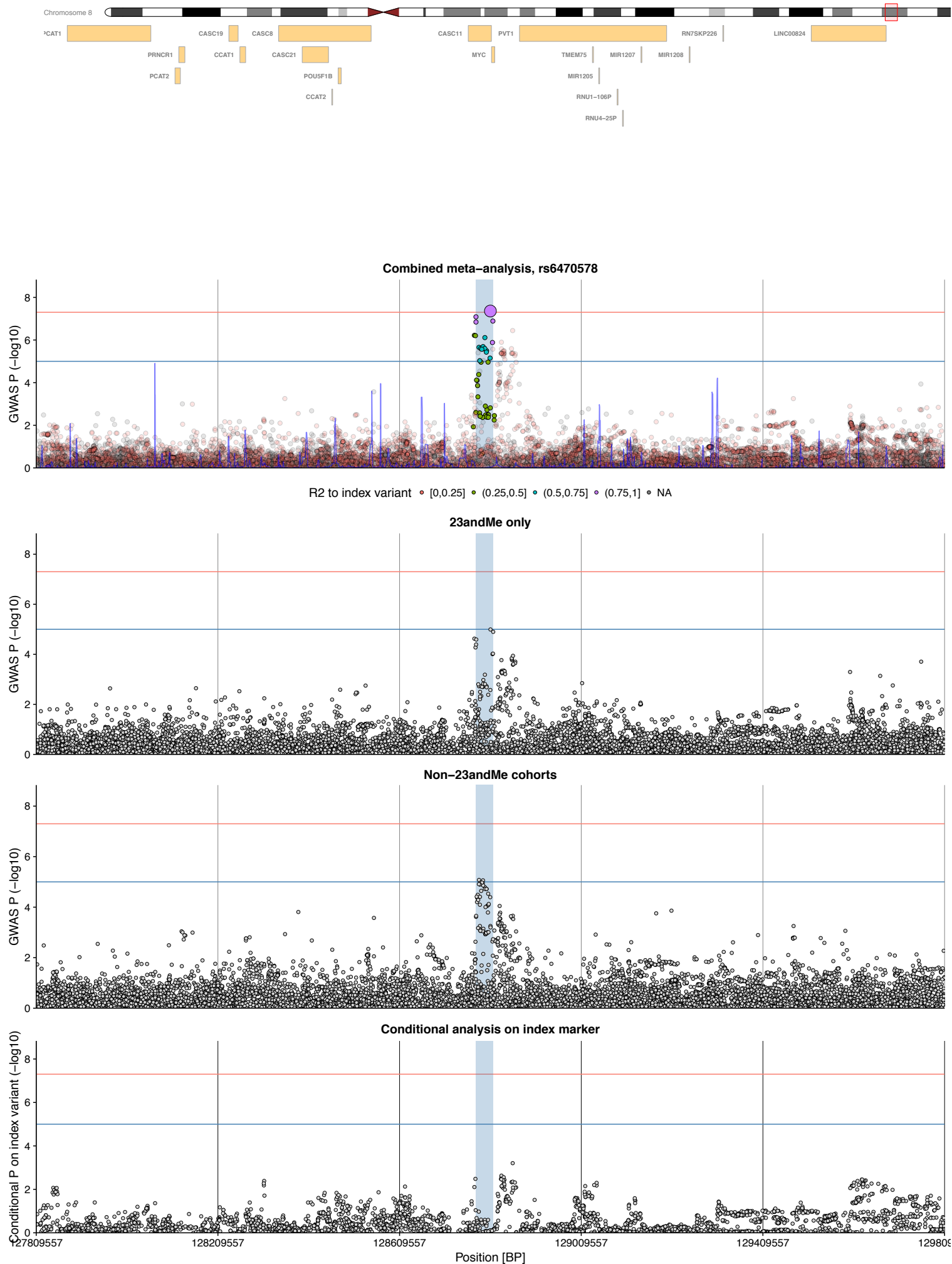


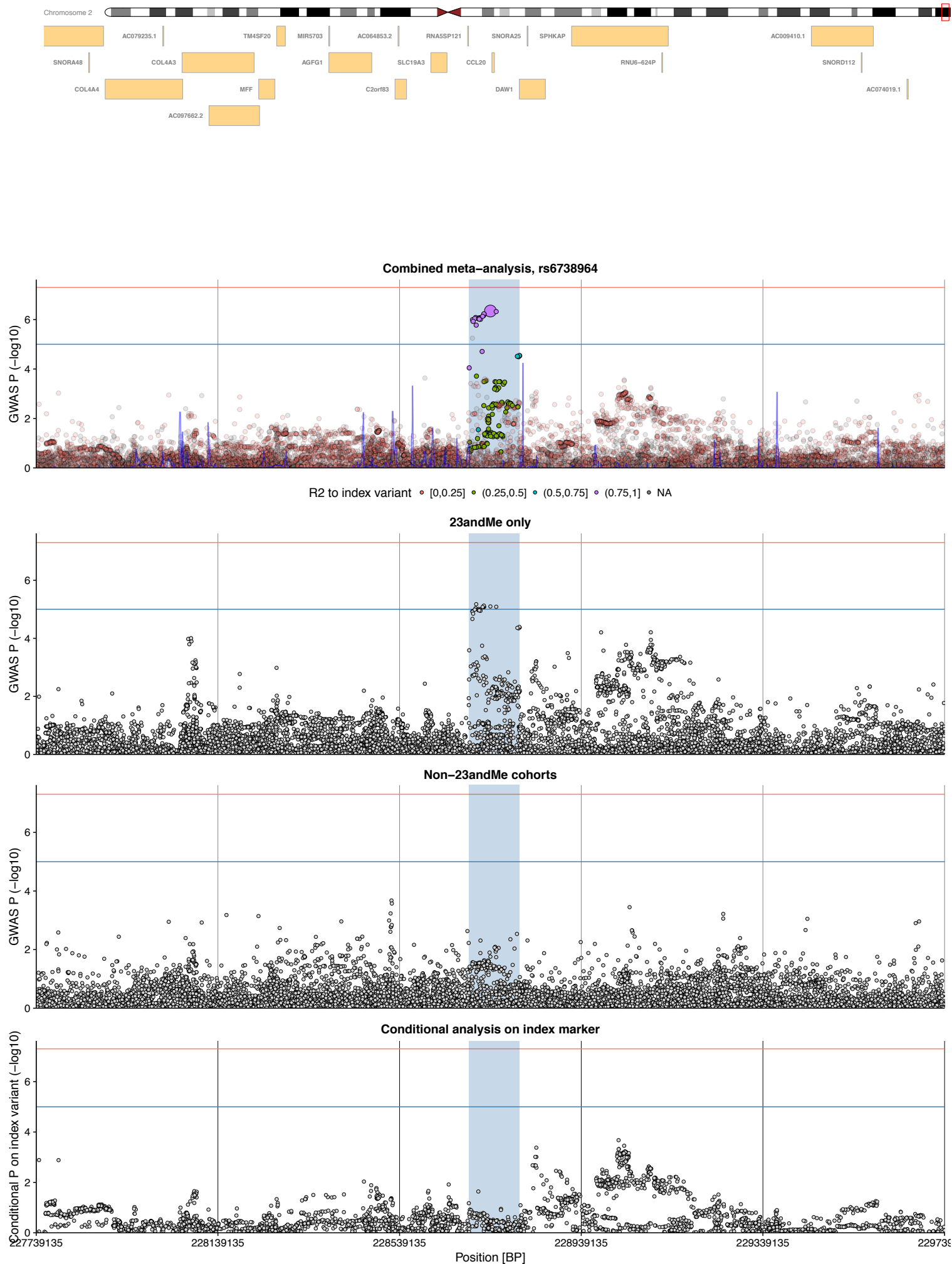


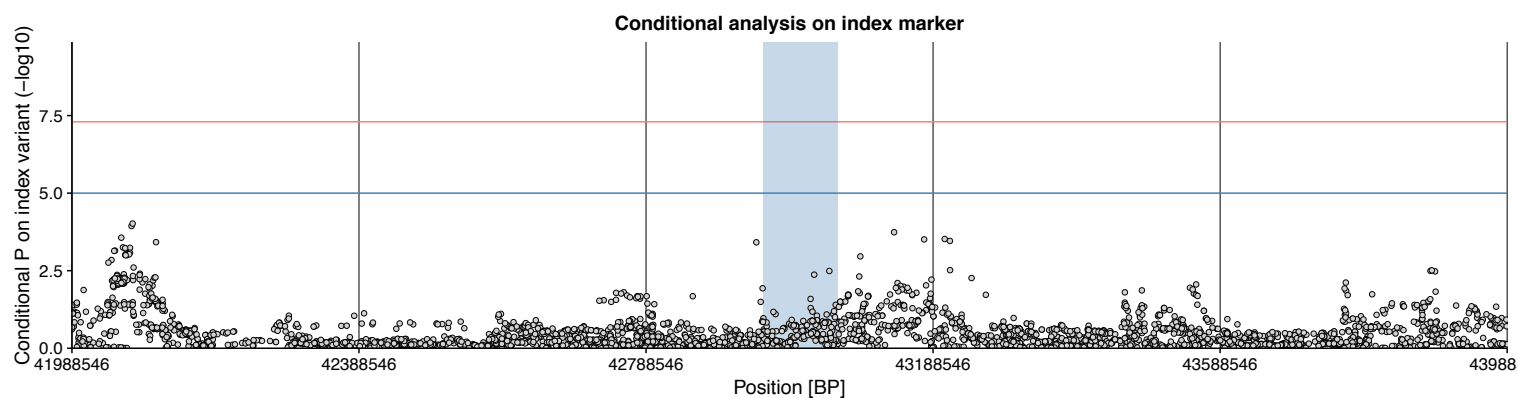
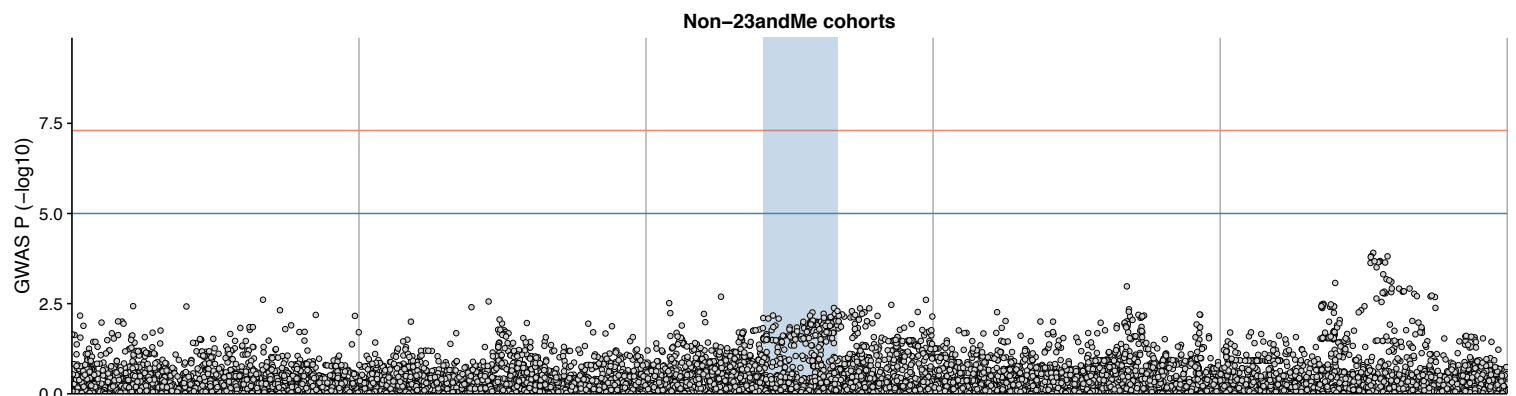
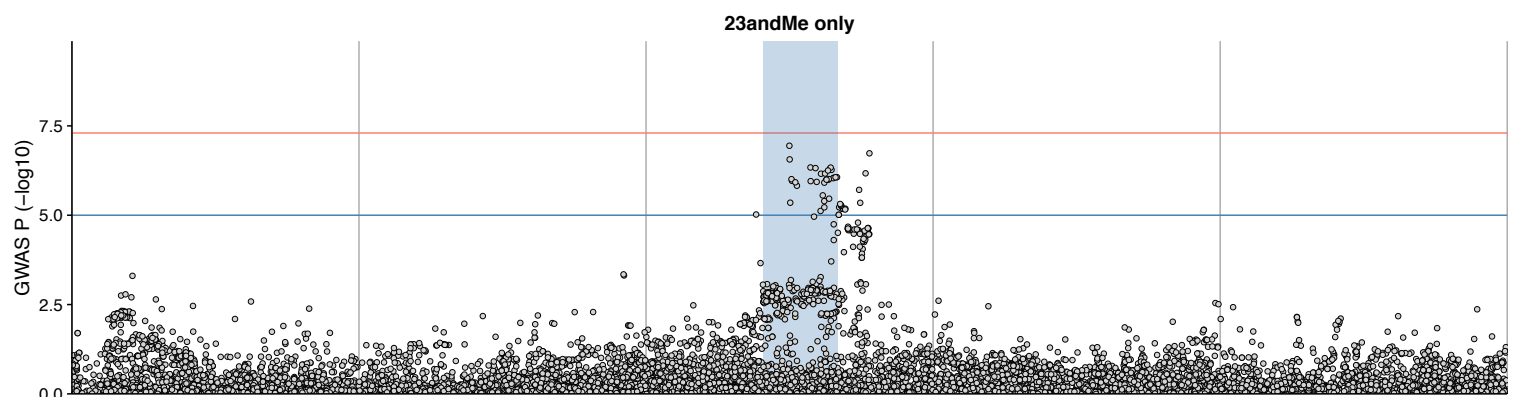
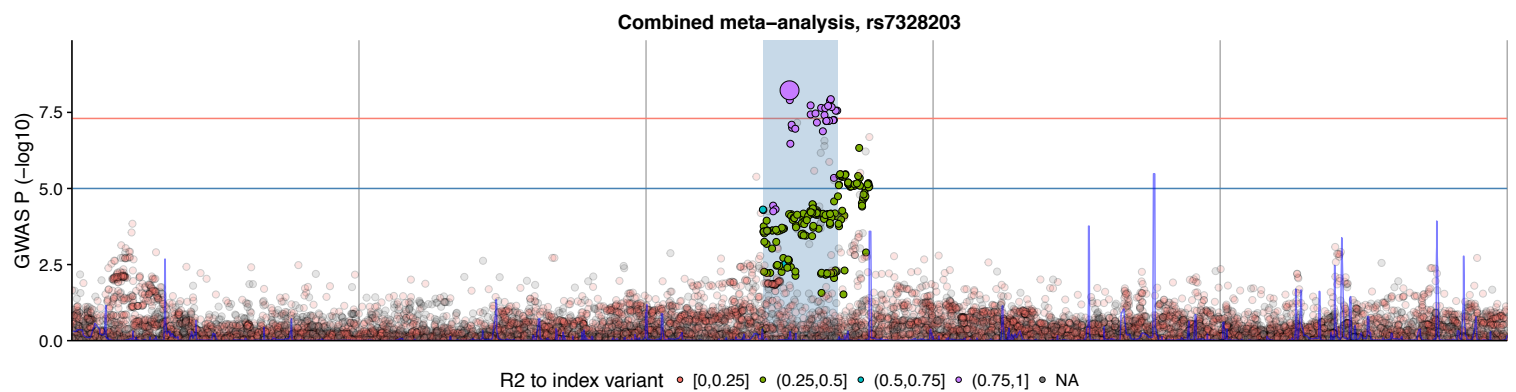
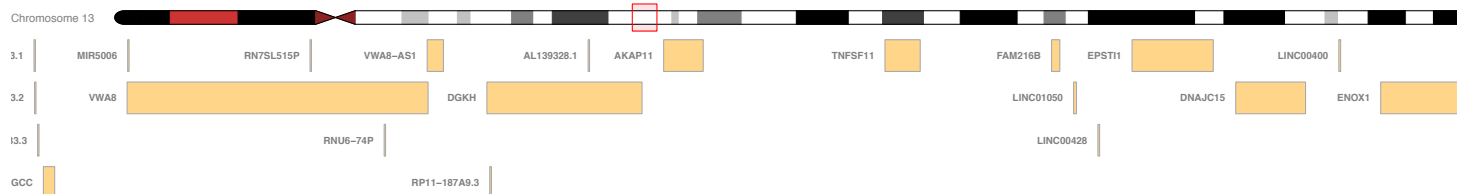


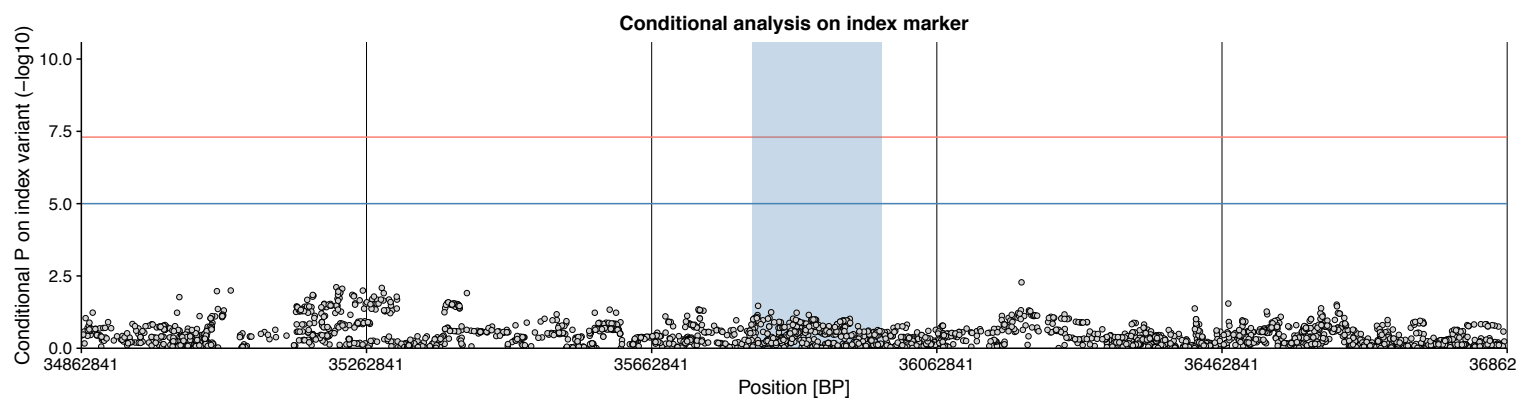
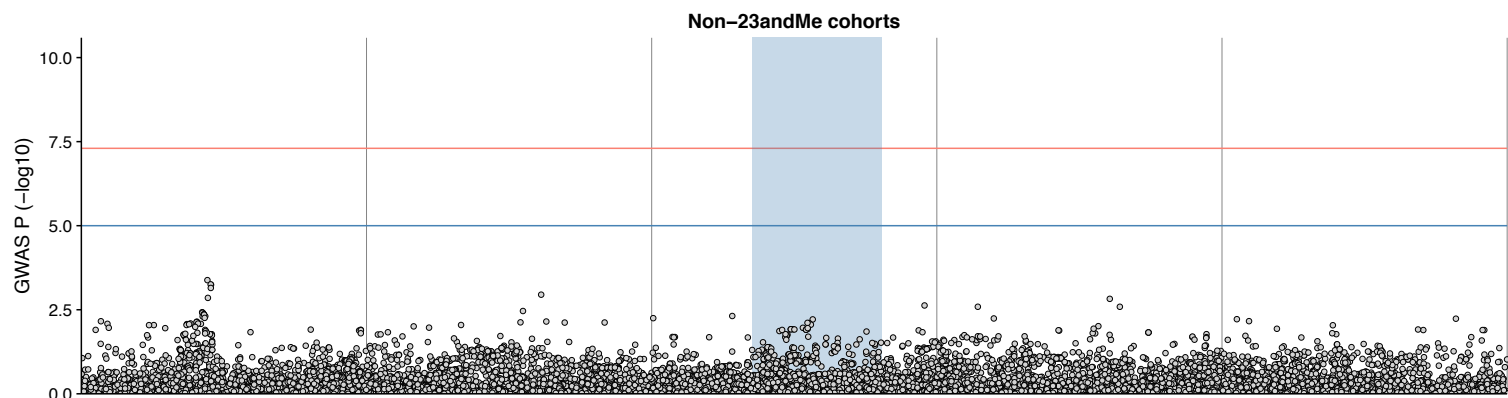
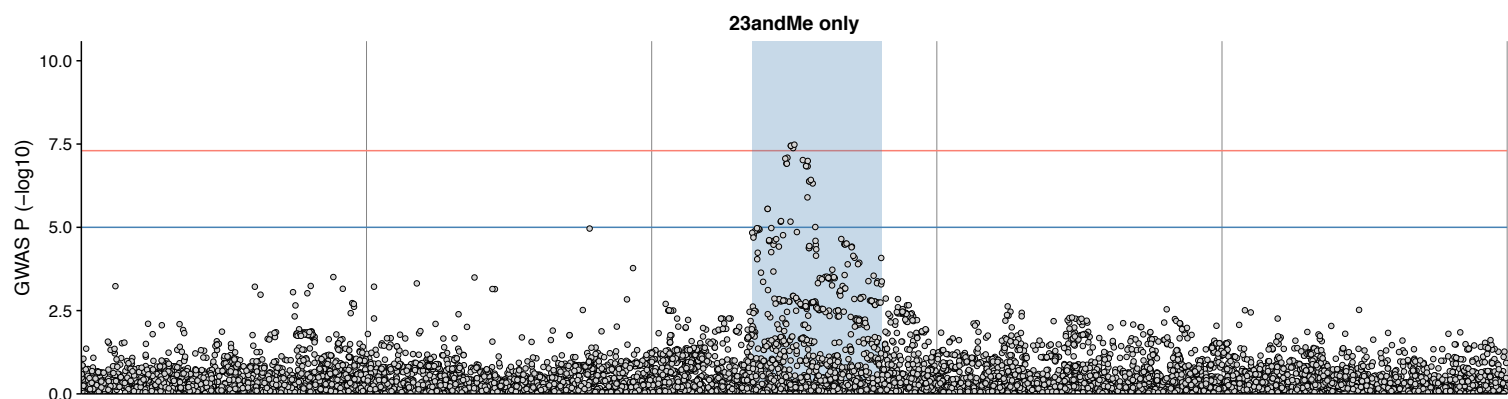
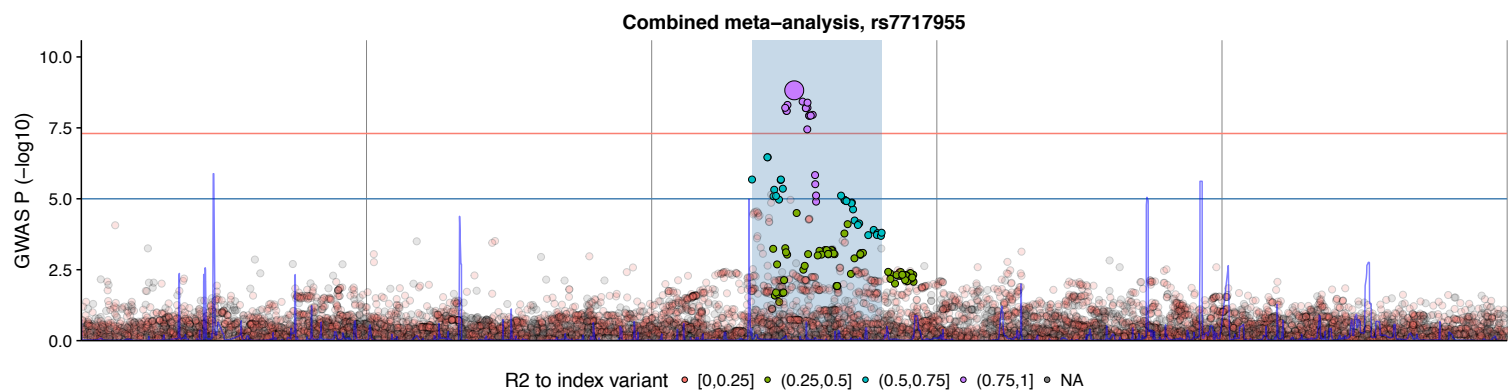
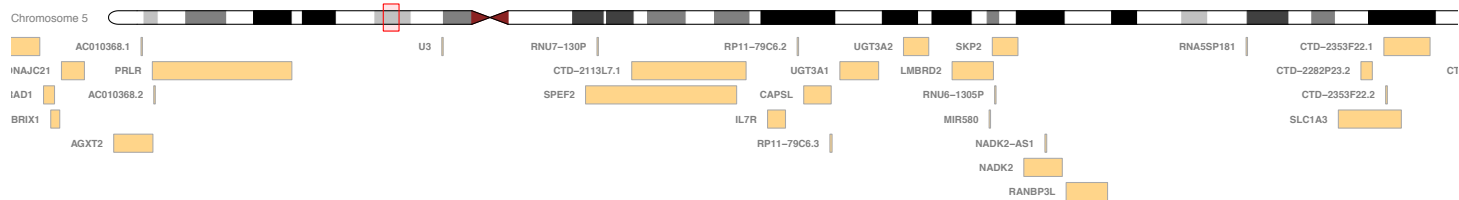


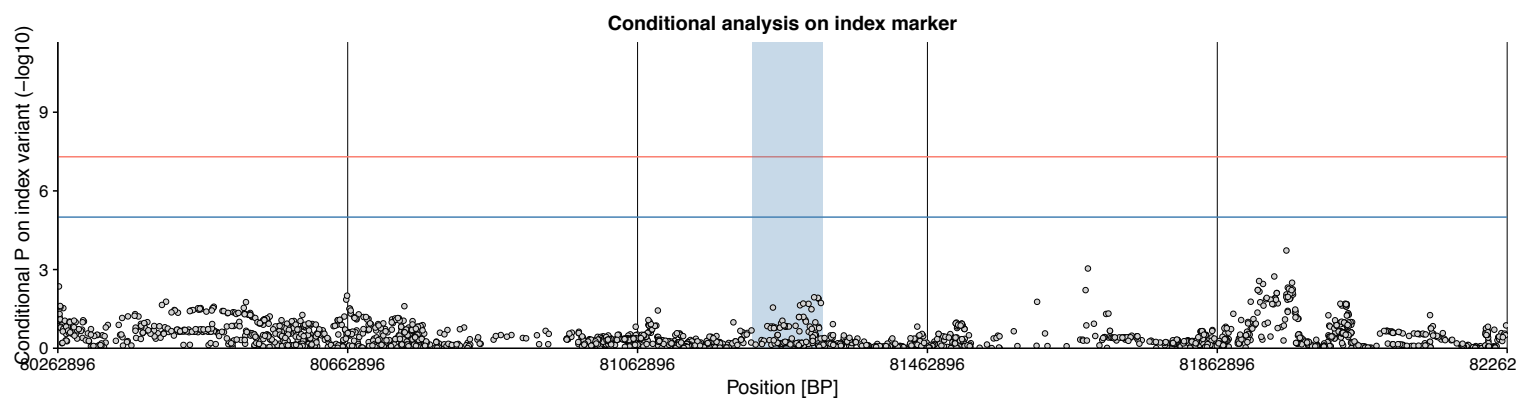
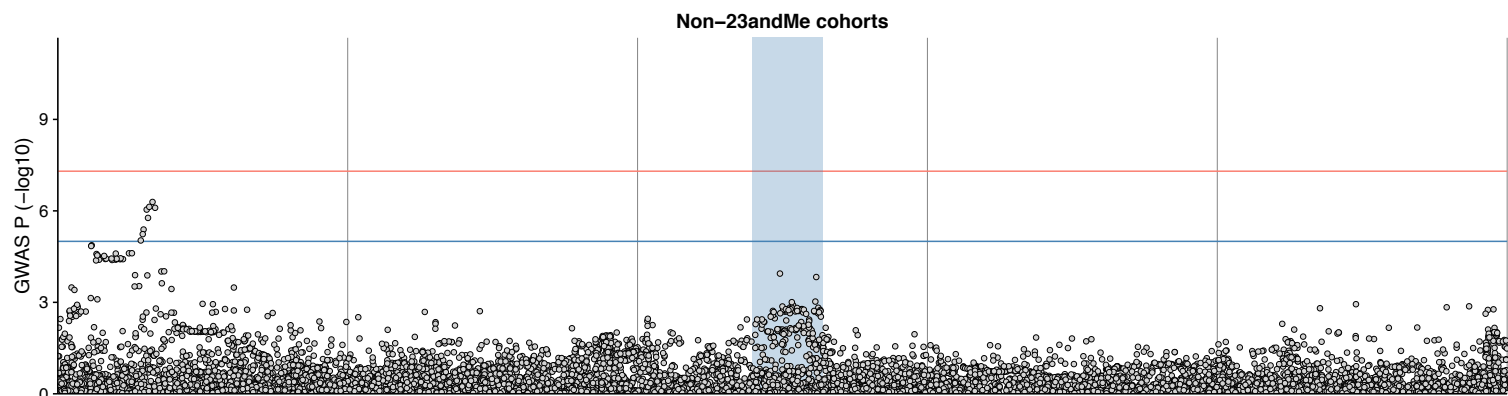
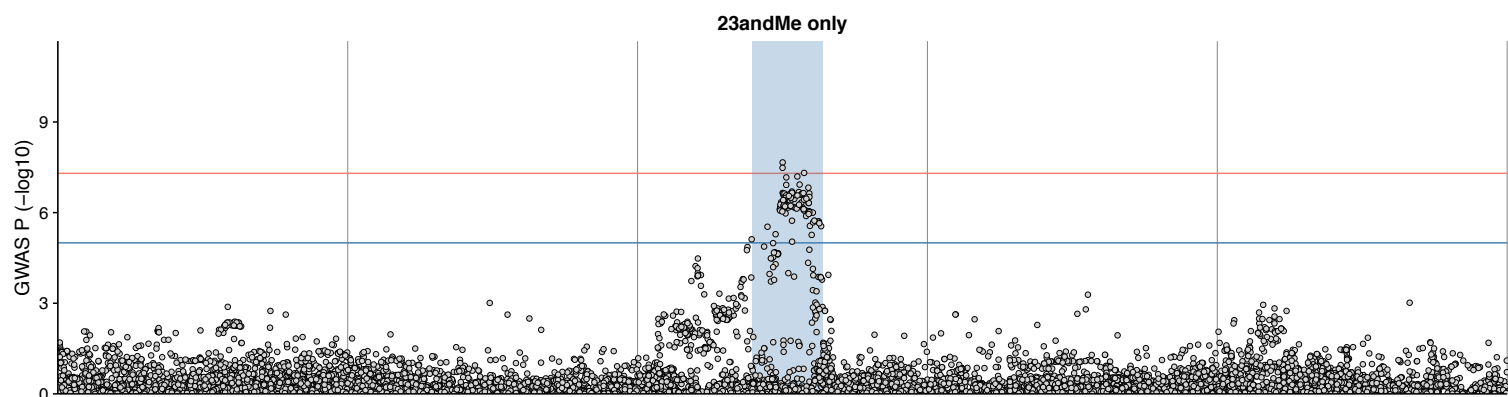
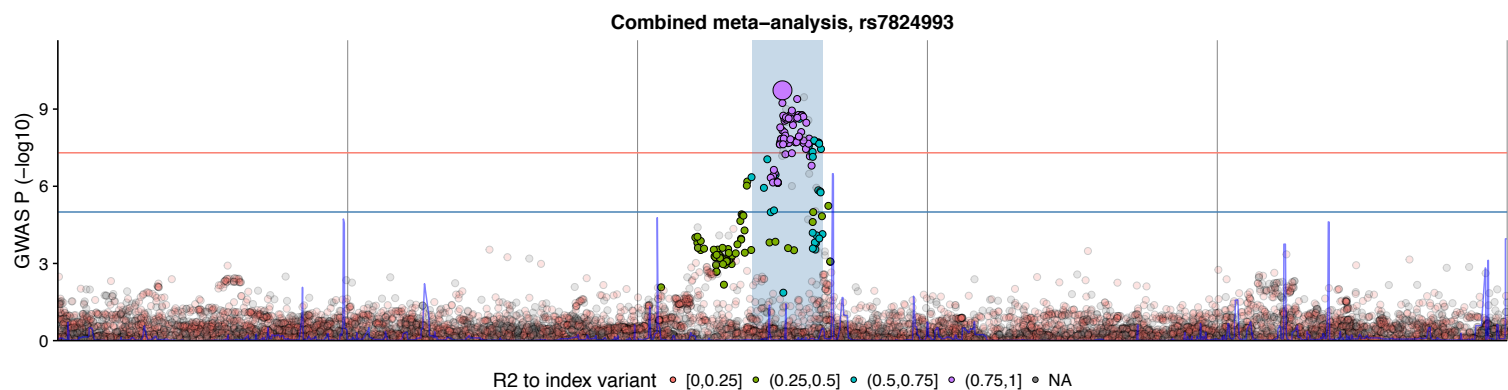
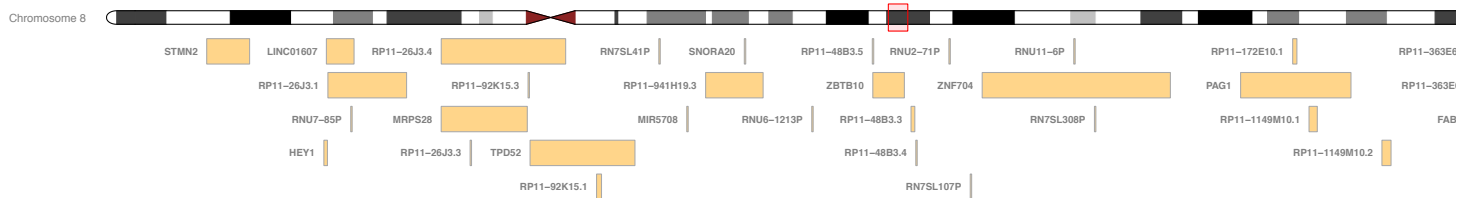


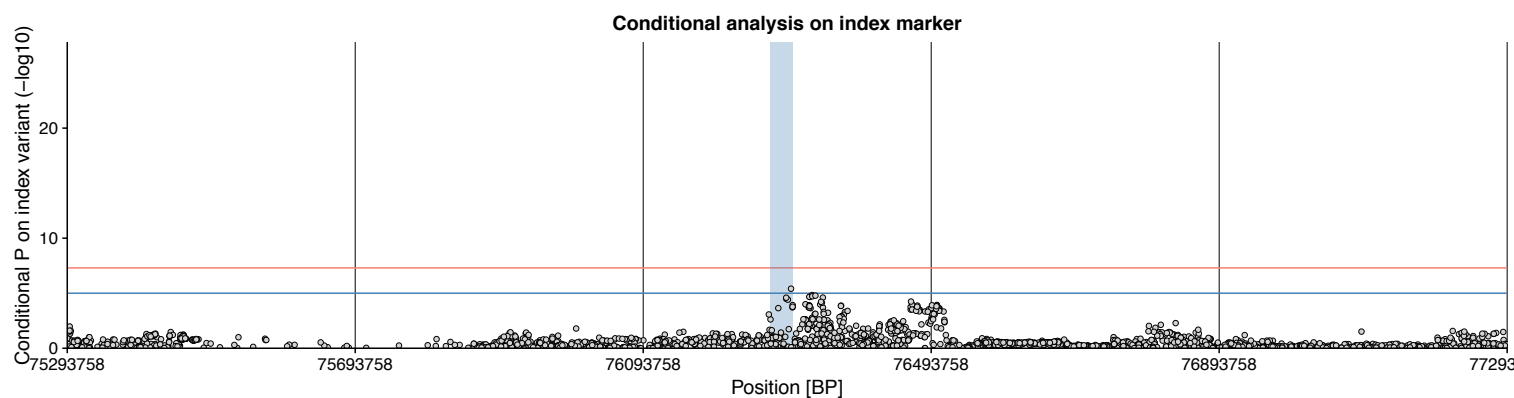
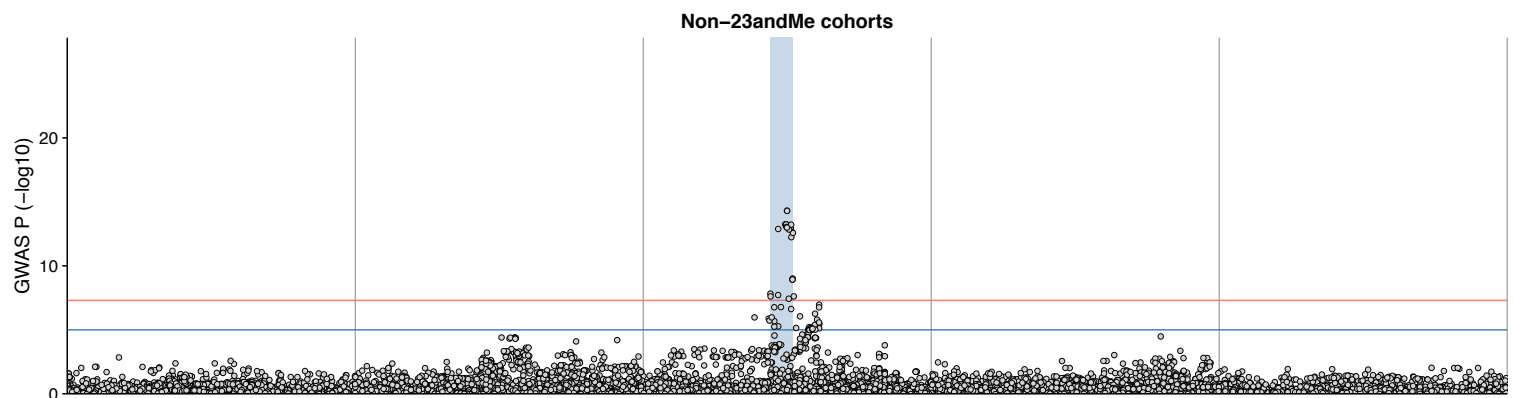
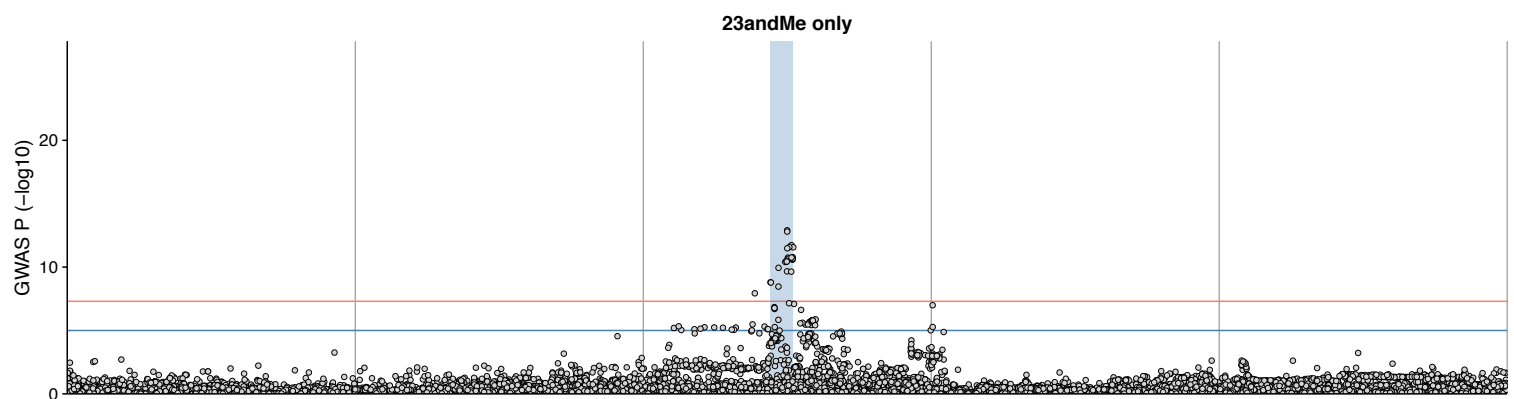
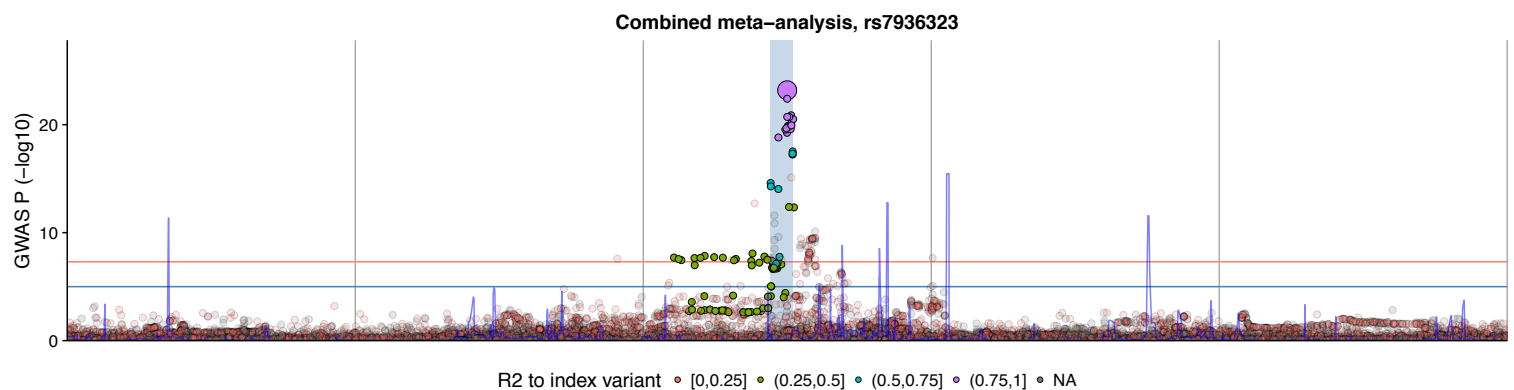
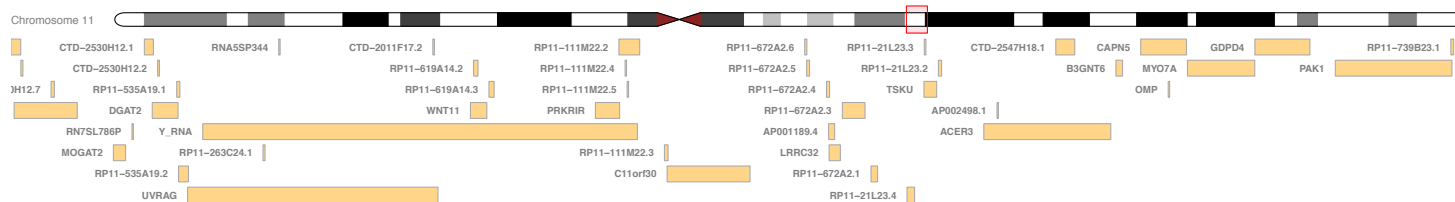


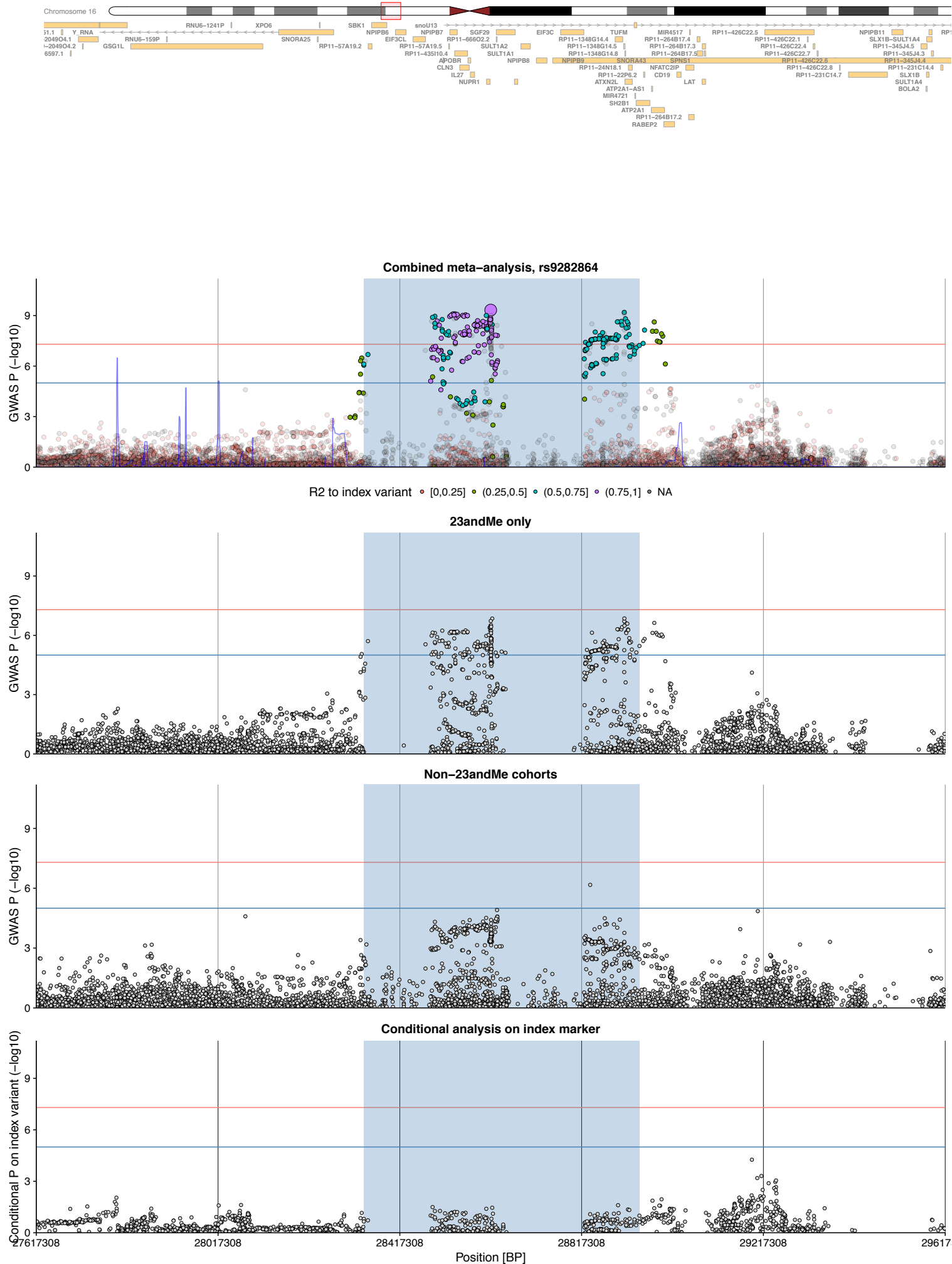


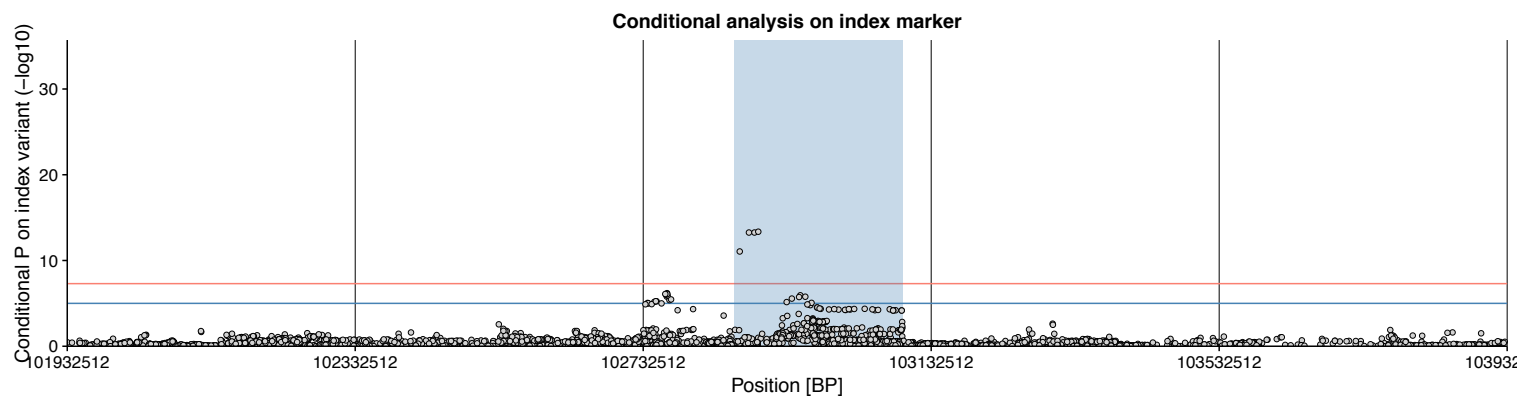
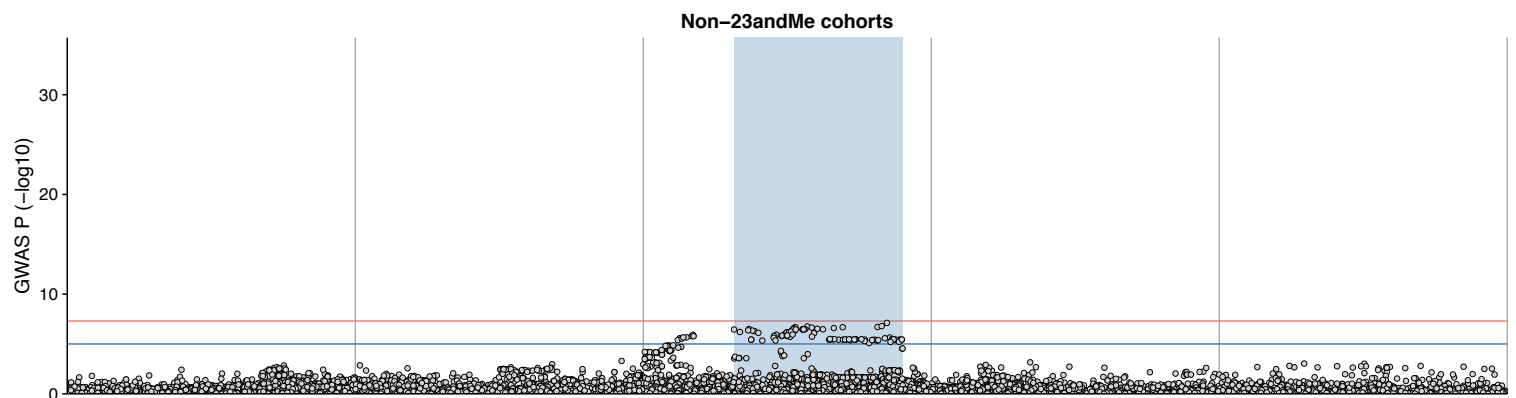
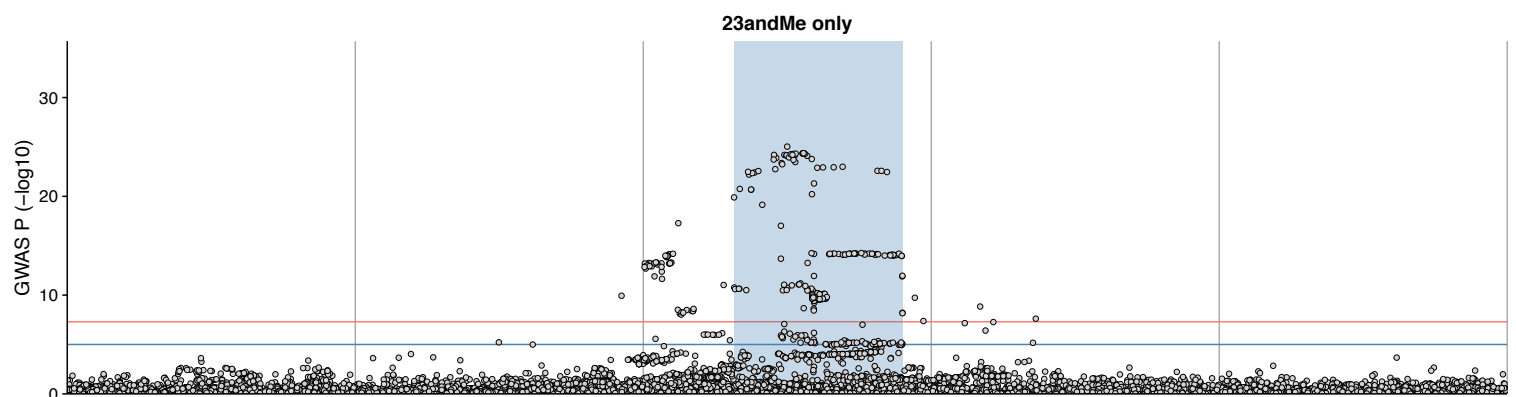
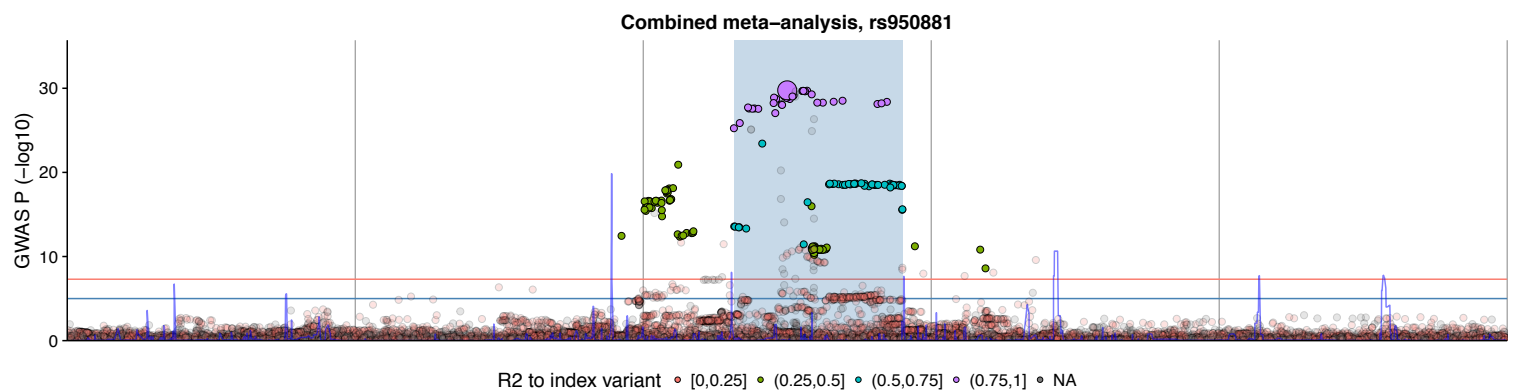
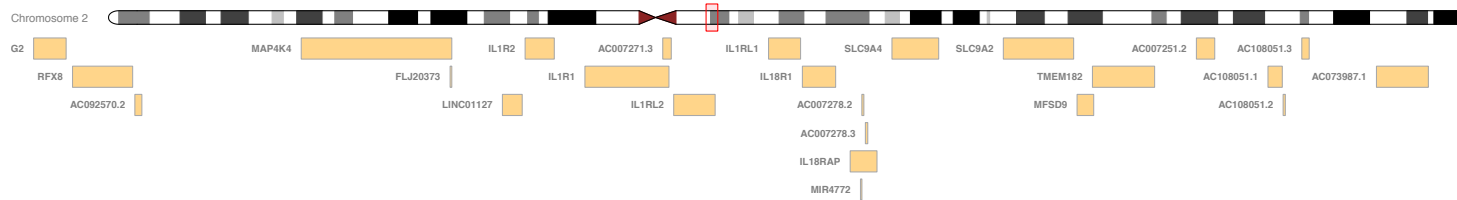


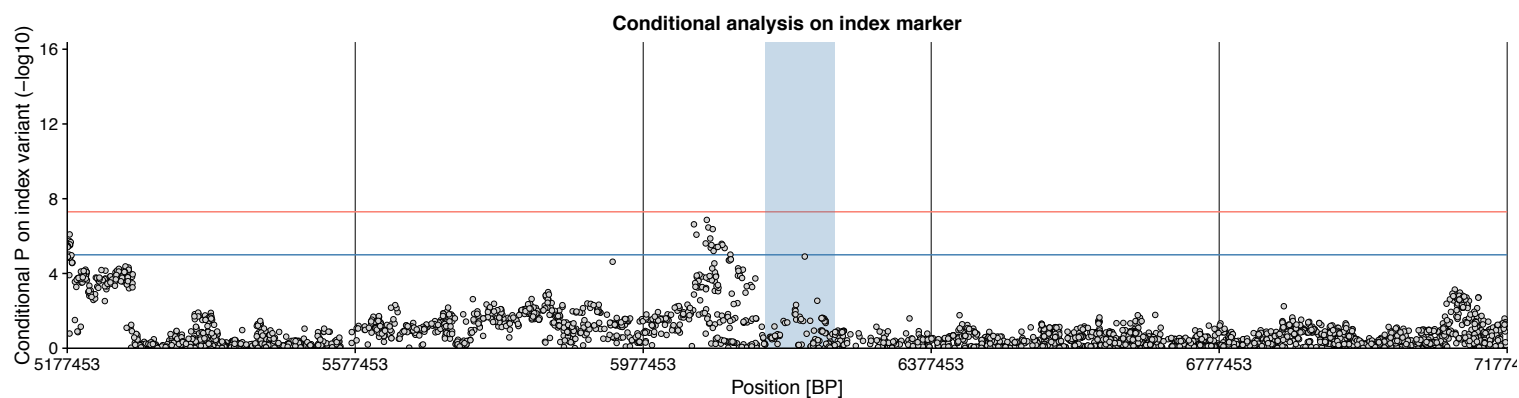
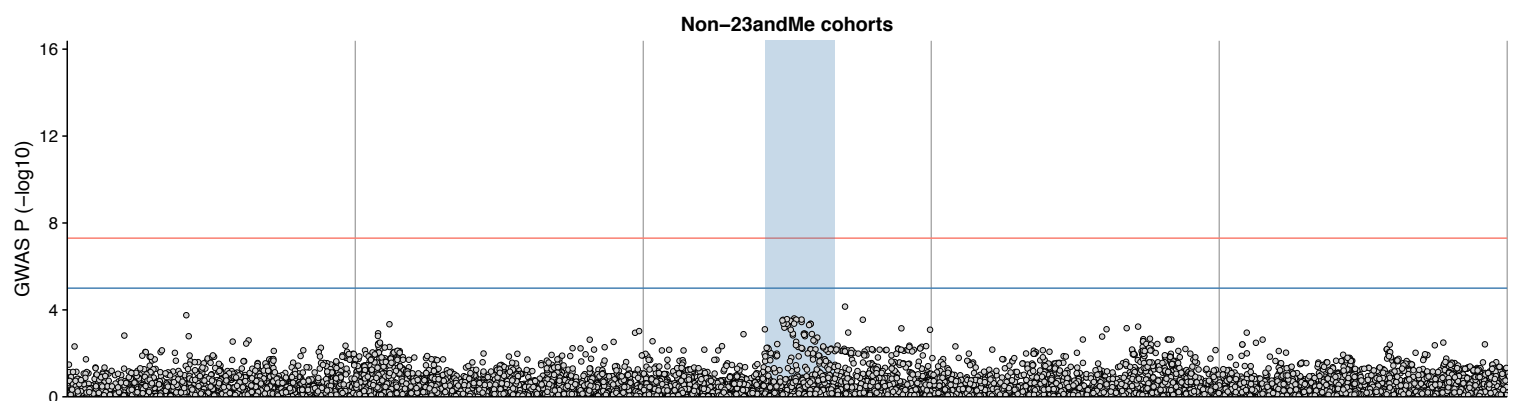
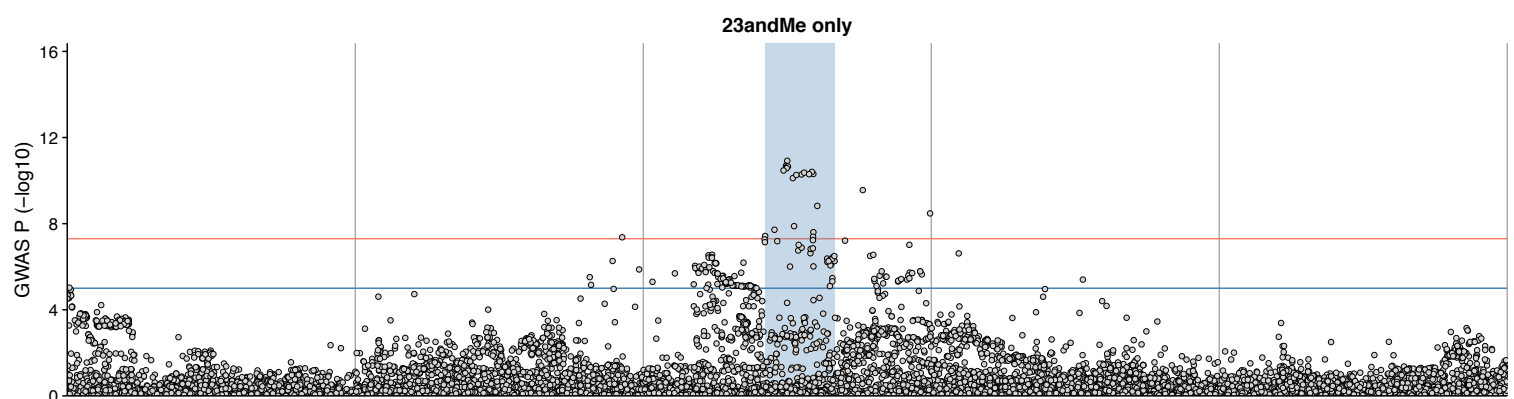
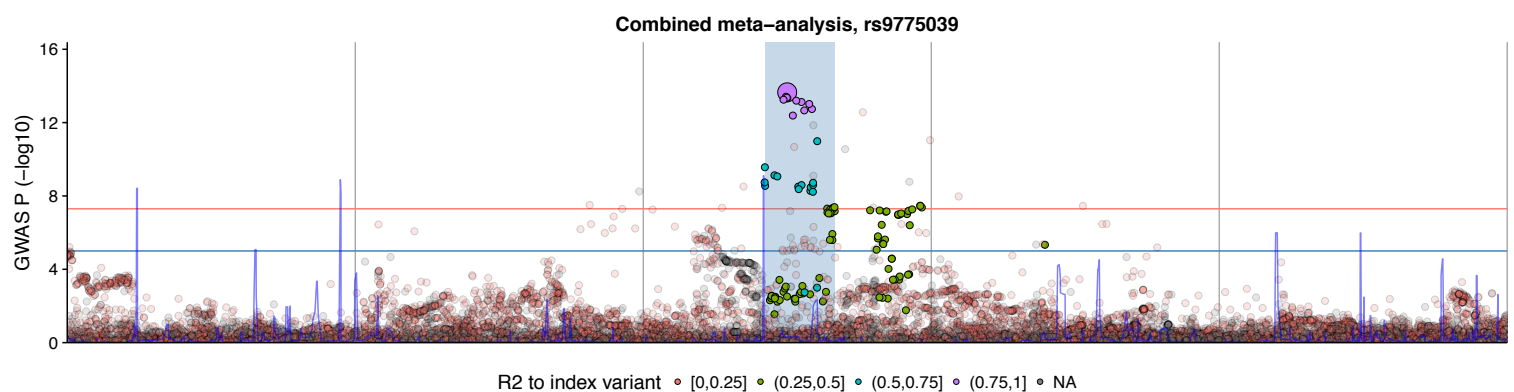
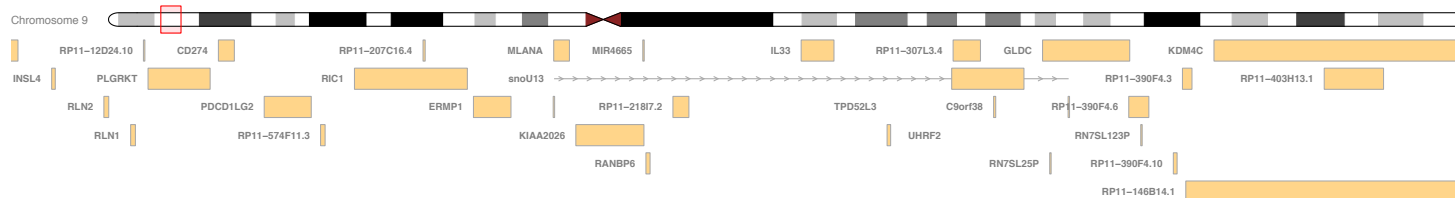










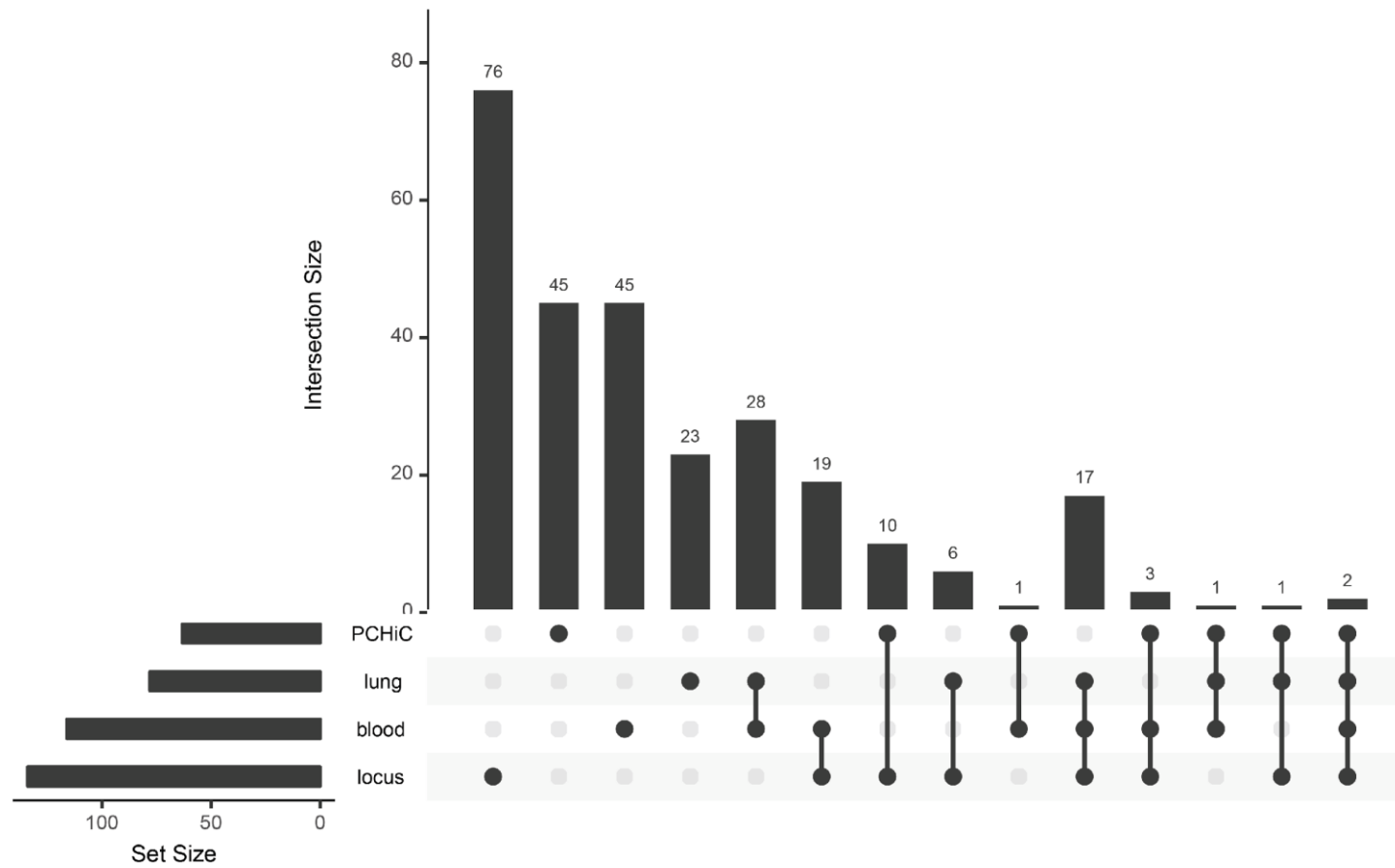


Enrichment plot from the GARFIELD regulatory element enrichment pipeline, showing tissue specific enrichment of DNASE hypersensitivity hotspots. Each central layered densitogram color represents enrichment values for increasing marker p-value thresholds (1 through 1e-8). The GARFIELD analysis was run on p-values from genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals.



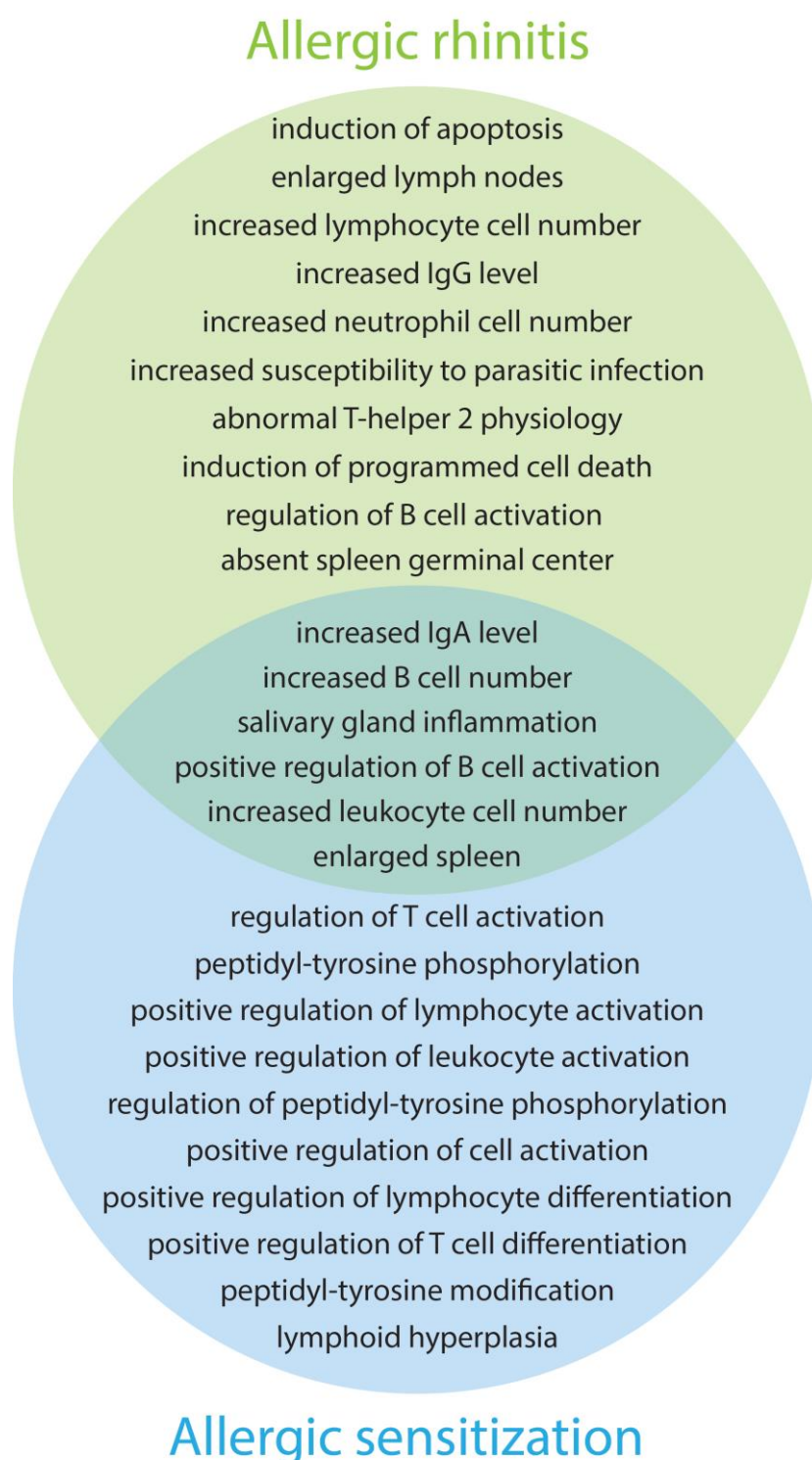
Supplementary Figure 7

“UpSet” set visualisation plot, indicating set intersections between e/meQTL blood, e/meQTL lung, PCHiC, and locus genes.



Supplementary Figure 8

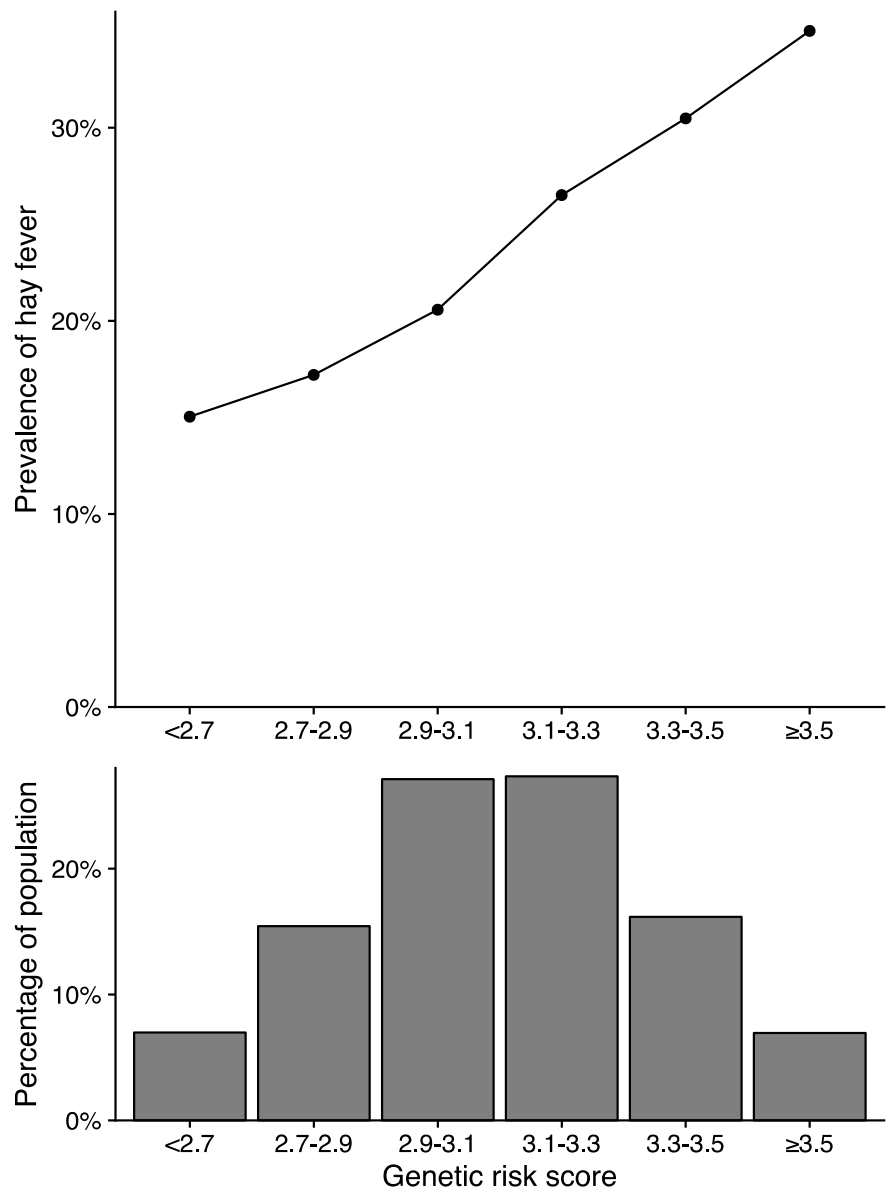
Overlap of enriched gene sets from the allergic rhinitis and allergic sensitization GWAS. All significant overlapping gene sets are shown together with the 10 gene sets with most statistically significant enrichment from each phenotype. Results for all gene sets are shown in Supplementary Tables 21 and 22. Enriched gene sets were calculated from an inverse variance weighted fixed-effect meta-analysis of genetic marker association of 16,531,985 genetic markers to allergic rhinitis from the discovery phase, including 212,120 individuals, and from inverse variance weighted fixed-effect meta-analysis of genetic marker association of 16,531,985 genetic markers to allergic sensitisation, including 24,481 individuals.



Supplementary Figure 9

Relationship between allergic rhinitis genetic risk score and disease prevalence.

Combined impact of risk alleles from the 41 genome-wide significant and replicating loci on prevalence of allergic rhinitis (hay fever) in the population-based B58C study. For each individual, we calculated a genetic risk score by applying per-allele risk estimates from the replication sets to the number of higher-risk alleles at the 41 genome-wide significant and replicating loci (one SNP per locus). The risk score thus represents an index of the number of weighted risk alleles. Along the x axis, individuals in each risk score interval are shown (with the lower bound included and the upper bound excluded in each interval). The prevalence of hay fever in each interval is plotted in the upper panel, and the histogram in the bottom panel shows the percentage of the B58C population in each risk score interval.



Supplementary Figure 10

Regional locus plot of association values from the allergic rhinitis GWAS for marker rs193243426 (N=203,356). The locus has an index marker with $p < 1e5-8$ (p-value from genetic marker association from the discovery phase), but lacks a credible locus LD structure, and was excluded from the study.

