Genetic Commonalities in Gynecologic Cancers Using Publicly Available Genome-Wide Association Study Summary Results: An Exploratory Meta-Analysis

by

Mark Francis Vater

BS Biology, Northern Kentucky University, 2014

Submitted to the Graduate Faculty of the

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Mark Francis Vater

It was defended on

August 4, 2021

and approved by

Jeanine Buchanich, Associate Professor, Department of Biostatistics

Ada Youk, Associate Professor, Department of Biostatistics

John Shaffer, Assistant Professor, Department of Human Genetics

Thesis Advisor: Jenna Carlson, Assistant Professor, Department of Biostatistics

Genetic Commonalities in Gynecologic Cancers Using Publicly Available Genome-Wide Association Study Summary Results: An Exploratory Meta-Analysis

Mark Francis Vater, MS

University of Pittsburgh, 2021

Abstract

Public Health Significance: Gynecologic cancers are responsible for millions of deaths worldwide every year. The goal of this research is to further scientific understanding of such cancers, potentially leading to improved diagnosis and treatment. The public health significance of this project is to uncover potential genetic underpinnings of gynecologic cancers, aiding in efforts to reduce the mortality rate of gynecologic cancers, which is imperative in protecting the health of susceptible persons across the globe.

Gynecologic cancers are those which arise in the female reproductive system, chiefly, those of the ovaries, cervix, vulva, endometrium, and vagina. These conditions present a serious threat to the health of susceptible persons, leading to nearly three million deaths worldwide each year. Questions remain about the genetic origins and risk factors for each of them. Specifically, there is uncertainty regarding potential overlap in genetic architecture for these conditions. This analysis sought to answer the questions: are there shared genetic variants between ovarian cancers and endometrial and cervical cancers? If so, what are the genetic implications of these overlapping variants? This was accomplished by gathering summary-level data from published genome-wide association studies (GWAS) and analyzing them using fixed effects inverse variance meta-analysis. Four datasets were included, three ovarian cancer datasets, one cervical and one endometrial cancer dataset. Three meta-analyses were performed: ovarian + endometrial cancer,

ovarian + cervical cancer, and ovarian cancer only. Several significant variants were found in the ovarian + endometrial cancer meta-analysis, including those located on genes *TERT* (p=3.1e-17), *ABO* (p=3.4e-11), and *ATAD5* (p=3.7e-12). The ovarian + cervical cancer meta-analysis was inconclusive, but the ovarian-only meta-analysis provided evidence for significant variants across datasets, including those located on genes *TIPARP* (p=2.3e-14) and *SKAP1* (p=3.0e-10). Some variants found may yield new insight if studied in conjunction with both types of cancer in the meta-analysis, such as *ABO* for endometrial cancer, a known genetic factor in ovarian cancer risk. Further study is needed to determine the relevance and level of involvement for shared genetic variants in multiple types of gynecologic cancers, potentially leading to new or improved treatments.

Table of Contents

List of Tables

List of Figures

Preface

The creator of this work wishes to thank the Department of Biostatistics and the Department of Human Genetics at the University of Pittsburgh Graduate School of Public Health, as well as family and friends that supported him throughout this journey.

1.0 Introduction

Gynecologic cancers are a class of cancers that affect people susceptible to cancers of the ovaries, cervix, endometrium, vulva, and vagina (i.e., those with female reproductive organs), and result in approximately 2.9 million cancer deaths every year globally (Sankaranarayanan and Ferlay, 2006). Specific to the United States, the incidence of gynecologic cancers is approximately 40 diagnoses and 12.5 deaths per 100,000 susceptible people per year (Phelan et al., 2017). As such, these cancers pose a serious threat to the health of susceptible people. Each of these cancers has been well researched singularly. However, the similarities and differences of the underlying sub-types of gynecologic cancers are not fully understood. That is, there may be commonalities in the mechanisms of gynecologic cancers in terms of tumorigenesis, genomic makeup, etc., that are currently not characterized in the body of literature. Investigating such commonalities may lead to improved or new methods of treatment or prevention for gynecologic cancers.

1.1 Distinct Characteristics of Gynecologic Cancer Sub-Types

Examining the overlapping characteristics of gynecologic cancers necessitates understanding the etiologic and prognostic differences across sub-types. Cervical cancer, for example, is thought to arise virtually always in cases of individuals with human papillomavirus (HPV) (Ledford and Lockwood, 2019). Contrasted to cervical cancer, endometrial cancer is thought to be attributed to imbalances in hormone regulation, as high body mass index is a risk factor for development of endometrial cancer; this is potentially due to the role of adipose tissue

in circulating estrogen in post-menopausal women, promoting endometrial growth (Guo et al., 2020; Ledford and Lockwood, 2019; Schmandt et al., 2011). The causes of ovarian cancers are less clearly understood; some are hypothesized to be related to endometriosis and others could arise in the fallopian tube epithelium (Zorn et al., 2005; Guo et al., 2020). Vaginal and vulvar cancers are rarer, both of which are related to HPV (Ledford and Lockwood, 2019). Prognosis also varies across gynecologic cancer sub-types. Ovarian cancer is more deadly than the other four gynecologic cancers, which is speculated to be due to late diagnosis times (i.e., higher severity of cancer stage at time of discovery) (Ledford and Lockwood, 2019). Ledford and Lockwood (2019) estimated that ovarian cancer accounted for approximately 14,000 deaths in 2018, compared to endometrial cancers at approximately 11,000, and the other three types accounting for less than 5,000 deaths in the United States, based on 2018 estimates (Ledford and Lockwood, 2019). According to a literature review of gynecologic cancer clinical outcomes and association with genetic polymorphisms, as reported by Diaz-Padilla et al. (2012), no consistent associations were repeatedly observed by searching databases including EMBASE, MEDLINE, and Cochrane for specific gene polymorphisms and cross-phenotype outcome (Diaz-Padilla et al., 2012).

1.2 General Commonalities Across Gynecologic Cancer Sub-Types

While each of the five cancers vary in cellular makeup and physical location of tumor, there have been some overall similarities observed between sub-types of gynecologic cancer. For example, all types of gynecologic cancers originate in the Mullerian ducts and are affected by female sex hormones (Berger et al., 2018; Guo et al., 2020). Endometrial and some ovarian cancers have been associated with Lynch syndrome, and cervical cancers and some ovarian cancers have

been associated with Peutz-Jeghers syndrome (Constantinou and Tischkowitz, 2017). Lynch syndrome is passed down through an autosomal dominant trait, with an approximate prevalence of 1:370 for susceptible populations globally, though not each case of Lynch syndrome results in the development of a carcinoma (Constantinou and Tischkowitz, 2017). Lynch syndrome is thought to develop by mutations in DNA repair genes *MLH1*, *MSH2*, *MSH6* and *PMS2* (Constantinou and Tischkowitz, 2017). Lynch syndrome has been associated with a number of cancers, generally, associated cases of which demonstrate microsatellite instability, redundant sequences which cause replication errors (Constantinou and Tischkowitz, 2017).

Wang et al. (2018) examined the role of driver genes across ovarian, cervical, and endometrial cancers by functional enrichment and pathway analysis. The authors found four differentially expressed genes to be noteworthy across the prognoses of the three cancer types, including lowered expression of *MCM2* (all sub-types), elevated expression of *MMP2* (primarily cervical), elevated expression of *COL1A1* (primarily ovarian), and elevated expression of *JUN* (primarily ovarian and endometrial) (Wang et al., 2018). Another study examined expression levels of *FAS* and *FAS*-ligand across ovarian, endometrial, and cervical carcinoma tissue, demonstrating decreased levels of *FAS* gene expression in all three sub-types compared with non-cancerous cell tissue (Das et al., 2000). Further, Zorn et al. (2005) assessed differential expression profiles of serous, endometrioid, and clear cell histotypes of ovarian and endometrial cancer. The authors found several common differentially expressed genes across ovarian and endometrial cancers and across histotypes, including *ANXA4*, *UGT1A1*, *FXYD2*, and *GLRX* (Zorn et al., 2005).

A whole-exome study in 209 Chinese gynecologic cancer patients (specifically, ovarian, endometrial, and cervical cancer) demonstrated some commonalities across the three cancer types (Guo et al., 2020). Primarily, the authors identified mutations in the cancer-associated gene *PIK3CA* across all three types studied (122/209, 58% of cases studied) (Guo et al., 2020). Other tumor associated gene mutations were also observed, including *PTEN*, *TP53*, *CDC27*, *ZFHX3*, *MUC16*, *ARIDIA*, *KMT2C*, *KRAS*, and *BRCA2*, which were observed in 9% to 25% of overall cases across the total sample (Guo et al., 2020). Further, Guo et al. reported that *MUC16/CA125*, associated with an ovarian cancer biomarker protein, was mutated in 20% of endometrial cancer cases, 29% of cervical cancer cases, and 12% of ovarian cancer cases, as well as the cilia motility associated gene *HYDIN* mutated in 43% (90/209) of the total sample (Guo et al., 2020).

Guo et al. (2020) also assessed copy number variants (CNVs) across the 209 Chinese endometrial, cervical, and ovarian cancer patients cohort. Specifically, the authors identified two amplified regions (12p13.33 (75/209), 15q26.3 (33/209)) and two deleted regions (9p24.3 (69/209), 11p15.5 (47/209)) across the three cancer types (Guo et al., 2020). The amplified regions covered retinal coding proteins and the deleted regions covered tumor suppressor genes (Guo et al., 2020).

Additionally, Guo et al. described molecular similarities between the three cancer types. Specifically, they described evidence in each cancer sub-type for DNA mismatch repair deficiencies, noted by mutation signature (Guo et al., 2020). They further demonstrated a number of shared enriched biological pathways across cancer types, two of which were associated with cilia movement (GO:0044782, GO:0001539), one associated with chromatin modification

(GO:0016569), and a number associated with cell or tissue morphogenesis (GO:0022604, GO:0002009, GO:0003007) and cellular organization (GO:0098742, GO:0034330) (Guo et al., 2020: Supplementary Table 3).

Berger et al. (2018) also examined the molecular properties of gynecologic cancers, specifically, ovarian, cervical, and uterine/endometrial cancers using Cancer Genome Atlas data (TCGA). They found 23 genes that were more frequently mutated in gynecologic cancers compared with other TCGA cancer types via enrichment, including *FBXW7, PIK3CA, PIK3R1*, *TP53*, and *PTEN* (Berger et al., 2018). The authors additionally performed hierarchical unsupervised clustering on tumor mutation signatures, resulting in ten distinct clusters. Two of the clusters were associated with mutations related to DNA mismatch and strand repair (Berger et al., 2018). Another cluster showed a high number of microsatellite instability mutations (primarily uterine/endometrial samples), and another showed a high number of *APOBEC* mutations (primarily cervical samples) (Berger et al., 2018). The authors additionally performed hierarchical clustering over CNVs, resulting in six clusters, two of which were associated with *TP53* mutations, and the majority of samples in these two clusters at least one whole genome doubling event (Berger et al., 2018).

Lastly, Zhang et al. (2018) performed a meta-analysis across case-control studies of gynecological cancers that included one or both of two gene promoter polymorphisms, *MDM2 T309G* and/or *MDM2 G285C*. The authors found the risk of gynecologic outcome was increased or decreased, in both polymorphic cases, depending on the underlying genotype at those locations (Zhang et al., 2018).

There also exists some evidence of multiple types of gynecologic cancers synchronously manifesting in the same individuals. Specific to ovarian and endometrial cancers, co-occurrence arises in approximately 5% of endometrial cancer cases and 10% of ovarian cancer cases (Takeda et al., 2018). Herrinton et al. (2001) reviewed three population-based case-control studies across 56 women presenting synchronous ovarian and endometrial tumors (Herrinton et al., 2001). Specifically, the participants in the underlying studies were those that were diagnosed with both cancers within a 12-month period and were between 20 and 54 years of age (Herrinton et al., 2001). The authors found that, compared with only ovarian cases, participants with both ovarian and endometrial cancers were more likely to have endometrioid (43% observed, 10% expected) and less mucinous (31% observed, 54% expected) ovarian cancers (Herrinton et al., 2001). Takeda et al. (2018) examined one synchronous ovarian and endometrial case with Lynch syndrome, and reported that this case exhibited microsatellite instability as well as reduced *MSH2* and *MSH6* expression in both tumors (Takeda et al., 2018). There is also evidence of ovarian and cervical tumor synchronicity, however, cases are rare and isolated (Bacalbasa et al., 2020).

As such, there is foundational evidence suggesting a number of both shared and unique traits among sub-types of gynecologic cancers. The most obvious contrast between all the sub-types is the ultimate location of tumor growth, but there are a number of shared genes that have been associated with more than one type of gynecologic cancer. For instance, Berger et al. (2018) and Guo et al. (2020) both identified the *PTEN* and *TP53* genes in some association with ovarian, endometrial, and cervical cancers. And, while co-occurrences are rare, the presence of synchronous cases further indicates that there may be specific similarities affecting multiple

phenotypes of gynecologic cancer arising due to a similar genetic or environmental risk factor. There are some genes that potentially may be associated with such cancers but have not yet been identified in association, presenting an opportunity for study.

1.2.1 Genome-Wide Association Commonalities Across Sub-Types

Masuda et al. (2020) conducted a genome-wide association study (GWAS) meta-analysis across 647 ovarian cancer cases, 909 endometrial cancer cases, 538 cervical cancer cases, 5236 cases of uterine fibroid, and 645 cases of endometriosis from the Biobank Japan Project (39,556 controls). They specifically conducted random-effects meta-analyses for individual case groups and respective controls, then compared these with both logistic and linear-mixed regression models across all cases and shared controls. Additionally, the authors performed reverse regression analysis on imputed genotypes and covariate-adjusted phenotypes, such that phenotypes were assembled in the best combination to maximize log-likelihood estimates, with covariates including age, squared-age, BMI, and the top 20 principal components. They found one genome-wide significant locus by conducting random effect meta-analysis on the linear mixed model summary statistic results (*LOC730100*, OR: 1.16 (95% CI: 1.01-1.33), $p=2x10^{-8}$) of three conditions (ovarian cancer, endometrial cancer, and endometriosis) (Masuda et al., 2020). Three genome-wide significant loci were detected by reverse regression analysis (*GABBR2*, $p=4.8x10^{-8}$; *SH3GL3/BNC1*, $p=3.3x10^{-8}$; *LOC107985484*, $p=3.8x10^{-8}$), however, only the last locus included more than one cancer endpoint in its best model, ovarian and endometrial cancers (Masuda et al., 2020).

1.3 Research Question

There is some genetic and molecular evidence that gynecologic cancers are related to some degree. These shared properties may present enormous opportunities in future research of detection and treatment for gynecologic cancers. However, what is missing is a robust amount of data on shared genetic properties in a genome-wide context. This is an important gap to fill, and genome-wide association studies can be used to address this gap, as they assess the frequency of variants across the genome in a large number of cases and controls for a particular disease outcome. The variant-outcome results can then inform previously unrelated or undiscovered biological mechanisms and risk factors for the specific outcome. These results can be translated into more practical or effective medical treatments. The Masuda et al. (2020) study is one of few cross-phenotype meta-analyses over genome-wide gynecologic cancer data, and perhaps the only study assessing more than two gynecologic phenotypes in a genome-wide context. Specifically, GWAS analyses have built a foundation for personalized medical care, that is, treatments and therapies tailored to a patient's specific genome-wide makeup. Particularly, as it relates to pleiotropy, GWAS analyses have demonstrated that the same variant or variants can be associated with multiple phenotypic outcomes across cohorts (Li and Zhu, 2017; Visscher et al., 2017). However, the relationship between shared variants identified by GWAS results and biological implications is not always clear, necessitating further investigations from a genetic and molecular standpoint. Many of the variants identified in GWAS analyses are located in non-coding regions, so combining GWAS results with functional analyses can grant insight on the mechanisms by which specific variants influence a phenotypic outcome (Lichou and Trynka, 2020).

As such, the focus of this analysis is to identify shared genetic association signals within ovarian cancer as well as between cancer groups (ovarian + cervical and ovarian + endometrial). This will be accomplished using publicly available data from genome-wide association studies in the form of de-identified summary-level data, for each variant across studies to determine shared genetic architecture, which could shed light on new variants associated with one or more cancer phenotype.

2.0 Methods

2.1 Datasets Included

Datasets were identified by searching the GWAS Catalog (www.ebi.ac.uk/gwas/downloads/summary-statistics) for available sets of data containing the keywords “ovar”, “cerv”, and “endo”. Additionally, to identify available summary statistic datasets from existing publications not included in the GWAS Catalog, PubMed was queried with the keyword phrases “(((gynecol*) AND (cancer)) AND (gwas)” as well as “ovar*”, “endo*”, and “cerv*” variations for ‘gynecol*’, but no additional datasets were identified that met the inclusion criteria (described below) and did not overlap with already identified studies. To identify existing meta-analyses of gynecologic GWAS results, the keyword set (((meta) AND (gynecol*)) AND (cancer)) AND (gwas) was also queried in PubMed.

To be included in these meta-analyses, studies must have assessed a gynecologic cancer via GWAS, with available accompanying summary statistic packages in complete form, including translatable regression statistics (i.e., beta coefficients measuring the magnitude of effect, as well as associated standard errors and p-values), both alleles per variant with allele frequencies, and not overlapping samples with other included studies. This was the case for two studies (Rashkin et al., 2020 and UK Biobank (2018)), so Rashkin et al. was excluded; Rashkin et al. (2020) was further missing allele frequencies, presenting an issue with preparing the data. An additional source did not provide comprehensive study information (Leo et al., 2017), so it was excluded on this basis. As such, exclusion criteria were studies with incomplete summary statistic packages (either

missing alleles and/or frequencies, or untranslatable analysis statistics in the form of something other than a beta value or log odds ratio), or studies without freely available or any complete summary statistic results, resulting in the two study exclusions above.

Phelan et al. (2017) assessed data from the Ovarian Cancer Association Consortium (OCAC), as well as data from the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA), including 25,509 epithelial ovarian cancer cases and 40,941 controls in European ancestry women for GWAS analysis (Phelan et al., 2017). Histotypes of epithelial ovarian cancer that the authors studied included all invasive, serous invasive, high grade serous, low grade serous, mucinous, endometrioid, and clear cell carcinomas (Phelan et al., 2017). Genotyping for this study was performed by OncoArray and iCOGS at five sites, University of Cambridge, Center for Inherited Disease Research, National Cancer Institute, Genome Quebec and Mayo Clinic (Phelan et al., 2017). Genotyping QC was conducted by excluding samples with call rate 95%, low or high heterozygosity, samples that were not female, or duplicates (Phelan et al., 2017). They further applied QC measures on SNPs including those with call rate < 95%, those not in Hardy-Weinberg equilibrium, and those with < 98% concordance (Phelan et al., 2017). Imputation was performed for any region with a SNP with p-value < 10^{-6} using the 1000 Genomes Project reference panel (v3, October 2014), and by employing the IMPUTE2 method, with boundaries at +/- 500kb from the most significant regional SNP, with exclusion of imputed variants or accuracy $r^2 < 0.3$, or minor allele frequency < 0.01 (Phelan et al., 2017). Association analyses were conducted separately for both data sources, categorizing the disease outcomes as: all invasive disease, high grade serous, low grade serous, all invasive serous, serous borderline, low grade serous + borderline, endometrioid, clear cell carcinoma, or mucinous invasive + mucinous borderline

(Phelan et al., 2017). For OCAC, initial genome-wide analyses performed on each of the underlying studies were meta-analyzed by inverse variance fixed effects, adjusted for study and population substructure. After imputation for genotype in regions of interest, logistic regression was performed, adjusting for genotype-study combination and principal components (Phelan et al., 2017). For CIMBA data, the underlying studies were assessed by fixed effect meta-analysis, and a time-to-cancer survival analysis framework was further implemented. To address non-random sampling between *BRCA1* and *BRCA2* carriers, test statistic scores of retrospective likelihood were employed to assess the association between genotype and risk of ovarian cancer (Phelan et al., 2017). Both CIMBA and OCAC data were pooled in an inverse variance fixed effects meta-analysis (Phelan et al., 2017).

Lawrenson et al. (2019) meta-analyzed data from 63 studies under the OCAC, compiling 7,321 cases and 4,083 controls of Asian ancestry descent individuals (Lawrenson et al., 2019). Similarly, genotyping for this study occurred by OncoArray at one of five sites: University of Cambridge, Center for Inherited Disease Research, National Cancer Institute, Genome Quebec and Mayo Clinic (Lawrenson et al., 2019). Genotyping QC was characterized by excluding samples with call rate < 95%, low or high heterozygosity, samples that were not female, and duplicates (Lawrenson et al., 2019). Phenotypic outcomes were categorized as one of the following histotypes of epithelial ovarian cancer: high grade serous, low grade serous, endometrioid, clear cell, mucinous, 'other', borderline serous, and borderline mucinous (Lawrenson et al., 2019). Similarly, SNP QC was characterized by excluding SNPs with call rate < 95%, those not in Hardy-Weinberg equilibrium, those with concordance < 98%, those with minor allele frequency < 0.01, or those with imputation accuracy $r^2 < 0.3$ (Lawrenson et al., 2019). Imputation was executed by using the

1000 Genomes Project reference panel, and using the IMPUTE2 method (Lawrenson et al., 2019). Boundaries were set for +/- 500kb from the most significant regional SNP (Lawrenson et al., 2019). Initial genome-wide analyses were conducted for each of the datasets included, the results of which were meta-analyzed by fixed effects (Lawrenson et al., 2019). Following imputation, a simple logistic regression model was used, adjusting for genotype-study combination and principal components; association analyses were conducted for all histotypes, all invasive, all borderline, and each histotype individually (Lawrenson et al., 2019). Confounding by population substructure was addressed by calculating inflation in test statistics (Lawrenson et al., 2019).

O'Mara et al. (2018) meta-analyzed endometrial cancer data from 13 individual studies, as well as 45 additional consortia studies, to perform genome-wide association. The authors did not report specific criteria for phenotypic classification. Underlying consortium sources included the Endometrial Cancer Association Consortium (ECAC), Epidemiology of Endometrial Cancer Consortium (E2C2), Women's Health Initiative, and UK Biobank Round 1 results; as such, the O'Mara dataset was not included in analyses with the UK Biobank data sources (O'Mara et al., 2018). In total, they meta-analyzed 12,906 endometrial cancer cases and 108,979 matched controls of European ancestry descent. Genotyping was performed by OncoArray at one of two sites, The Center for Inherited Disease Research or the University of Melbourne; data from the E2C2 were genotyped by Illumina Human OmniExpress array and Illumina Human 660W array, while the data from the Women's Health Initiative were genotyped on five different arrays (O'Mara et al., 2018). The UK Biobank data were genotyped using the Affymetrix UK BiLEVE Axiom array and Affymetrix UK Biobank Axiom array (O'Mara et al., 2018). Genotyping QC was performed by removing samples with call rate < 95%, low or high heterozygosity, estimated European ancestry

< 80%, or those that were suspected to be from non-female individuals (O'Mara et al., 2018). SNP QC was performed by removing SNPs with call rate < 95%, those violating Hardy-Weinberg equilibrium, those with concordance < 98%; before imputing, SNPs with minor allele frequency < 0.01 and those with concordance < 98% were excluded (O'Mara et al., 2018). For the OncoArray data, imputation was performed using the 1000 Genomes Project reference panel, via the SHAPEIT2 method, with boundaries set at 5mb non-overlapping intervals (O'Mara et al., 2018). Data from the E2C2 were imputed by using the 1000 Genomes reference panel, and Women's Health Initiative were imputed by using both 1000 Genomes and UK10K reference panels, both by using the minimac2 software (O'Mara et al., 2018). The UK Biobank data were imputed by using both 1000 Genomes and UK10K reference panels, and the SHAPEIT3 and IMPUTE3 services (O'Mara et al., 2018). Logistic regression was used to model the associations between each of the datasets (to a total of 17 strata) and the outcome, adjusting for principal component relative to dataset (O'Mara et al., 2018). Population substructure was addressed by computing test statistic inflation, and the resulting effect estimates were combined in a fixed effects inverse variance meta-analysis (O'Mara et al., 2018).

The UK Biobank study, a large prospective cohort study, is an ongoing study collecting genotypic and phenotypic data on over 500,000 individuals of European descent (Abbott et al., 2018; Sudlow et al., 2015). For this analysis, the Round 2 results were used for both cervical cancer and ovarian cancer separately, which included 691 cases of ovarian cancer and 193,483 controls, and 192 cases of cervical cancer and 193,982 controls, respectively (Abbott et al., 2018). Samples were genotyped by the UK BiLEVE Axiom Array by Affymetrix and Applied Biosystems UK Biobank Axiom Array (Bycroft et al., 2018). Genotype QC measures included extensive

missingness and heterozygosity, and samples that were not female (Bycroft et al., 2018). Variant-based QC measures including testing for batch effects, plate effects, violation of Hardy-Weinberg equilibrium, sex effects, array effects, and discordance in control replicates (Bycroft et al., 2018). Imputation was performed using a custom version of IMPUTE2 with 250 kb buffer regions (Bycroft et al., 2018). Association tests to generate summary statistics were performed by least-squares linear modeling using disease or phenotype as the outcome predicted by additive genotype, with sex and the first 10 principal components as covariates (Howrigan et al., 2017).

After excluding studies not meeting inclusion criteria, the above four data sources remained. Two ovarian cancer studies (Phelan et al. 2017, Lawrenson et al. 2019) were used in all three analyses, while the endometrial cancer dataset (O'Mara 2018) was used in the ovarian+endometrial ('o+e') cancer meta-analysis, the UK Biobank ovarian cancer dataset was used in the 'ovarian only' ('o-o') cancer meta-analysis and the UK Biobank cervical cancer dataset was used in the ovarian+cervical ('o+c') cancer meta-analysis. Lastly, because the O'Mara dataset was reported in the genome build 'hg38', in order to properly align markers with the other datasets under meta-analysis, this data was converted to the genome build 'hg19' using the tool LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver), due to the other datasets having been reported in the 'hg19' build. Build conversion via LiftOver (available through the UCSC genome browser web tool) was accomplished by uploading the chromosome, position start, and position end data along with the marker name to match the LiftOver product to the existing O'Mara data (Kent et al., 2002). The position start component was populated by subtracting the 'position' field in the O'Mara dataset by one unit, as the authors only reported the position end data.

2.2 Data Processing

The data for these analyses was obtained from five sources (including two distinct phenotypes from the same consortia): three with ovarian cancer data (Phelan et al., 2017; Lawrenson et al., 2019; UK Biobank, 2018), one with endometrial cancer data (O'Mara et al., 2018), and one with cervical cancer data (UK Biobank, 2018). First, the summary statistic result files were downloaded from either the GWAS Catalog (www.ebi.ac.uk/gwas/downloads/summary-statistics) or the Neale Lab data repository (for data from the UK Biobank) (Buniello et al., 2019; Abbott et al., 2018). Each set of summary statistic files contained resulting GWAS statistics from an individual study including the following information for each variant, at a minimum: chromosome, physical chromosomal position, beta regression coefficient or log odds ratio, standard error, reference or baseline allele, minor or effect allele, and minor or effect allele frequency. Files were organized and quality-controlled SNP-wise using RStudio (R Core Team, 2020). Summary statistic files were read into RStudio in their base form, and subsetted by chromosome (or, in the case of per-chromosome files, simply read in one chromosome at a time). Because not all files contained rsID information, SNP names were constructed by combining chromosome and position (in the GRCh37 human genome assembly), separated by ':', to harmonize between datasets. Quality-control measures included removal of SNPs in each dataset followed these criteria: SNPs with negative p-values as well as p-values greater than 1, SNPs with negative standard error, and SNPs with minor allele frequency less than 0.01 (or effect allele frequency less than 0.01 or greater than 0.99). Details on variant filtering for each step are present in Figure 1. Files were then exported as .txt files for analysis.

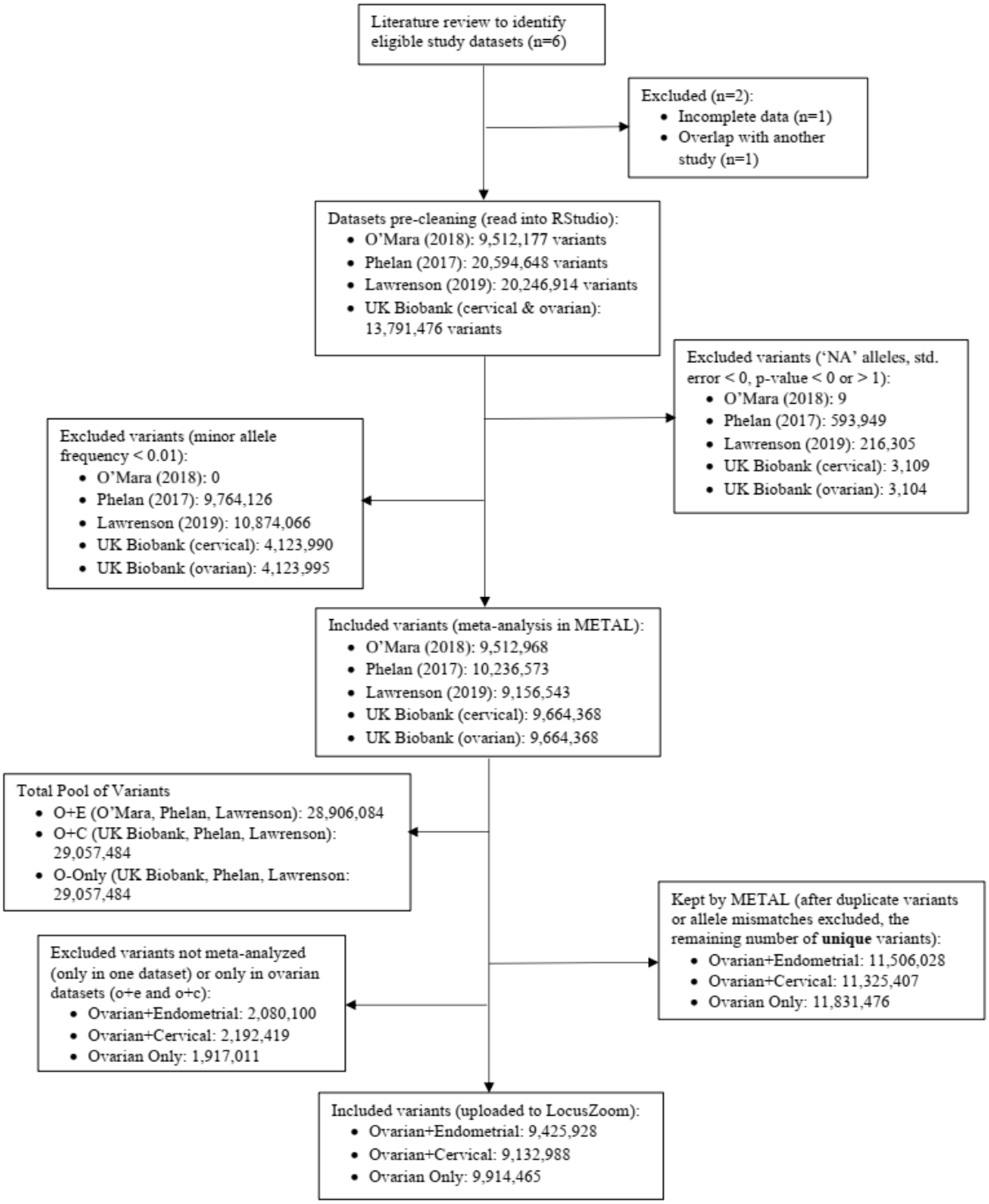


Figure 1: Flowchart of data from identification to assessment

2.3 Analysis Steps

Studies were meta-analyzed using the command line tool METAL (version 2011-03-25). METAL utilizes standard p-value-based or inverse variance-based methods for genetic meta-analysis. For this analysis, the inverse variance method was employed to preserve the interpretation of the estimated effect sizes. The equations that are used by METAL to compute the pooled effect estimates for inverse variance method are shown below (Willer et al., 2010). Scripts were written and used by METAL in accordance with the documentation for selecting the standard error scheme. Information included in each script about each dataset and dataset inputs included SNP marker name, reference/baseline and effect/minor allele, p-value, beta or log odds ratio, separator of .txt file, and standard error. Resulting files, per chromosome from METAL runs, were then read into RStudio and combined into results files per meta-analysis.

2.4 Fixed Effects Inverse Variance Meta-Analysis

The METAL software uses four mathematical steps to accomplish fixed effects meta-analysis under the inverse variance paradigm. The below equations represent the modeling performed in fixed effects methods.

(1.1) $Z = \beta / se$

where

(1.2) $\beta = \sum_i \beta_i w_i / \sum_i w_i$

and

(1.3) $w_i = 1/se_i^2$

and

(1.4) $se = \sqrt{1/\sum_i w_i}$

For each study contributing to the meta-analysis, Z-scores are calculated by dividing the pooled effect size, β, by the pooled standard error, se (1.1). In the remaining equations, the term i represents the studies included. To that end, if a variant was present only in one study, it was processed by METAL (with no pooling but using that variant's standard error) and excluded after reading the files back into RStudio, because that variant is not informative under the meta-analysis. The pooled effect size is calculated by summing the product of the per-study effect size, β_i with the per-study weight, w_i, then dividing it by the sum of the per-study weights (1.2). The per-study weight is calculated by taking the reciprocal of the corresponding study squared standard error (1.3). The pooled standard error is represented in equation 1.4, where the reciprocal of the per-study weight is summed and evaluated under a square root. Because the inverse variance method does not adjust for between-study variance, this method falls under a fixed effects method. Fixed effects methods are based on outcomes or phenotypes wherein there is no significant heterogeneity across all studies; that is, when combined under meta-analysis, each trial is assumed to share a common effect, such that the overall effect size is a true estimate of the common effect (Zhu et al, 2018; Li and Zhu, 2017). However, when there is considerable heterogeneity between studies, a random effect is generally used, which accounts for heterogeneity between studies. Random effects inverse variance models incorporate a second source of variance, between-dataset variance, τ^2, as shown below in Equation 1.5, which reflects the difference in effect sizes across input studies based on the scale of effect (Borenstein et al., 2010).

(1.5) $w_i = 1/(se_i^2 + \tau^2)$

In this analysis, a fixed effect inverse variance paradigm was selected, though each method confers both benefits and disadvantages. First, compared to random effects, fixed effects methods optimize the discovery rate (at the consequence of allowing higher false discovery rates, i.e., Type I error) for cross-phenotype SNP associations, as random effects models are much more statistically conservative and have limited power (Evangelou and Ioannidis, 2013; Pereira et al., 2009; Zhu et al., 2018). Because some of these analyses examined cross-phenotype genomic associations across ancestry groups, maximum power was desired, so that any possible SNPs associated with the cancer types could be elucidated, and then investigated for real genetic basis for association with both phenotypes. It has been indicated that fixed effects models may be more powerful than random effects models, even in the presence of considerable heterogeneity (Zhu et al., 2018; Pereira et al., 2009). This leads to the second point: because of the conservative statistical significance threshold in GWAS analysis, the number of loci that are significantly related to the outcome are generally low (Zhu et al., 2018). Further, random effects methods suffer from unreliable results (due to unreliable between-study variance when pooling effects) when the number of input studies is low (Guolo and Varin, 2017; Lin et al., 2020). As such, given the acceptance of some false positives (Type I error) in the goal of large-scale exploration and discovery, and the application of fixed effects methods to cross-phenotype data (with the assumption that the effect is consistent across phenotypes), this method was adequately suited for this analysis.

2.5 Additional Steps and Considerations

METAL also calculates heterogeneity within SNPs across datasets, shown below. While this is not incorporated into the models themselves, it is still useful to assess the between-dataset heterogeneity, per SNP, to quantify the variance between datasets. This is conducted through the I^2 statistic, which translates Cochran's Q (chi-squared statistic) into a percentage of heterogeneity between studies, as shown in Equation 2.1, where Q is the statistic following a chi-square distribution with (n – 1) degrees of freedom based on the number of datasets included in the models (Higgins et al., 2021).

(2.1) $I^2 = ((Q - df)/Q) * 100\%$

In contrast to τ^2, the I^2 coefficient is the proportion of observed variation relative to differences in effect sizes (Borenstein et al., 2009). The percentages themselves do not necessarily represent objective heterogeneity, for several reasons, but particularly because chi-squared tests are not very powerful for small sample meta-analyses. Generally, 0% to 40% heterogeneity may not be important, 30% to 90% may represent moderate to substantial heterogeneity, and 75% to 100% may represent significant heterogeneity. Though fixed effects methods assume no significant heterogeneity between studies, reporting significant variants with significant heterogeneity is important because it may inform further follow-up analysis, as it is possible that identified significant variants are linked to true causal variants but exhibit heterogeneity, which can be explored by techniques such as fine mapping (Ioannidis et al., 2007). Fine mapping involves the identification of causal variants by considering genes related to the outcome and examining

variants that are linked to those found significant by genome-wide association, but themselves do not reach genome-wide significance (Ioannidis et al., 2007).

Resulting files from METAL were uploaded to LocusZoom (my.locuszoom.org), an online service for visualization of GWAS results. Each meta-analysis, as well as each dataset that was entered into a model, was separately uploaded and stored. For each, Manhattan plots, QQ plots, and regional association plots were generated. Manhattan plots present the -log10 p-value for each SNP in chromosomal order, while regional association plots demonstrate the same information but for a narrow interval of the genome and are also used to examine linkage disequilibrium and recombination rate.

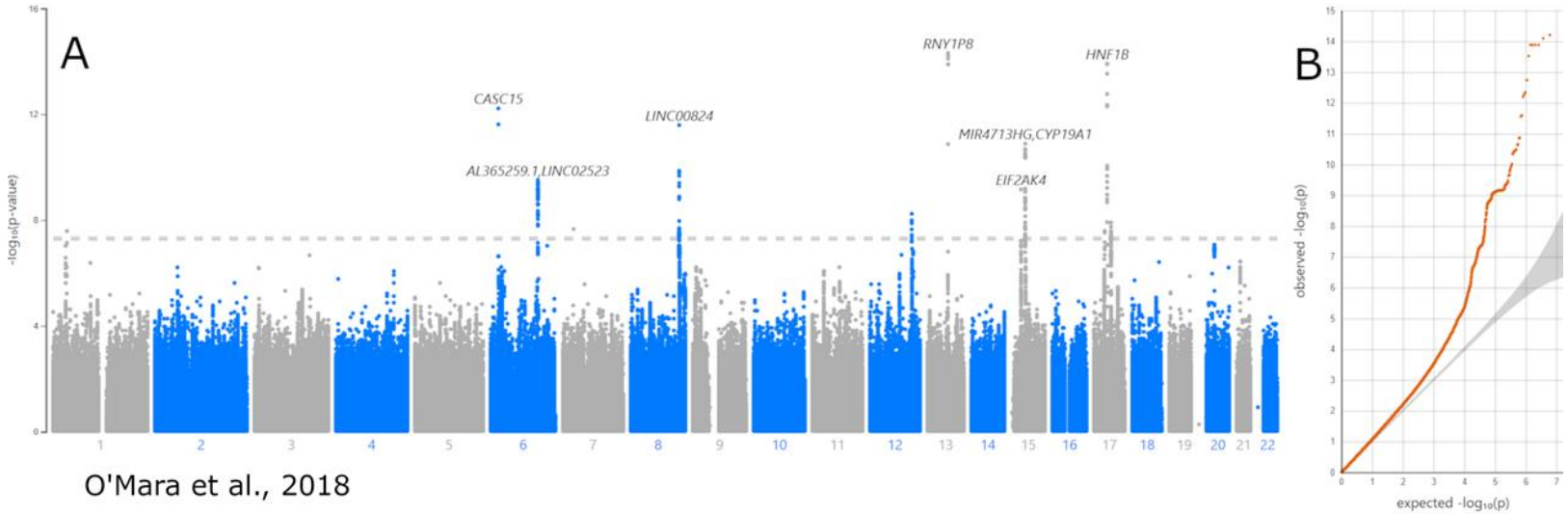
3.0 Results

3.1 Individual Study Results

In total, there were between 9 and 20 million SNPs per each of the data sources, and approximately 9 million SNPs per meta-analysis uploaded to LocusZoom after filtering (Figure 1). Many of the variants were lost to having minor allele frequency < 0.01, as well as during meta-analysis, as several markers were either duplicates of other markers (in name) or were excluded based on having different effect and/or baseline alleles compared to other studies with the same marker. It is important to reiterate that each meta-analysis contained one individual study as well as the Phelan (2017) and Lawrenson (2019) studies when considering the number of variants that were meta-analyzed.

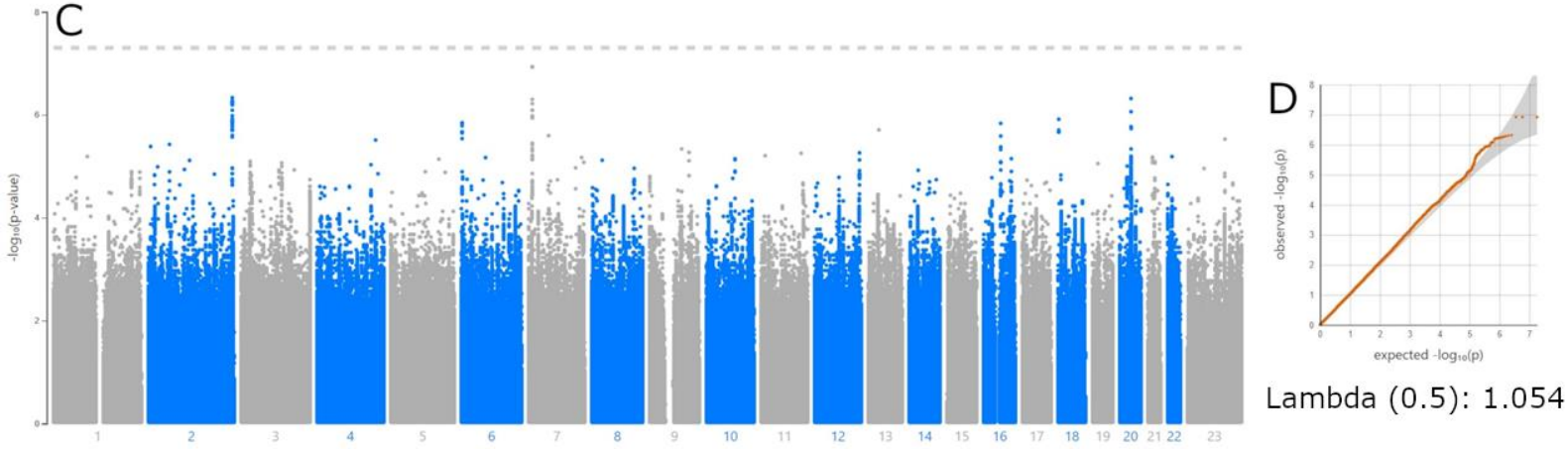
Graphical results, including Manhattan and quantile-quantile plots, for the individual studies uploaded to the LocusZoom tool are shown in Figures 2A-2J; these uploaded datasets were the filtered versions used in the meta-analyses. Nearest genes to SNPs with the highest negative log-p-value per significant locus are annotated on the Manhattan plots by LocusZoom. Overall, the results of individual studies reflect those presented in the respective publications. For instance, the O'Mara (2018) Manhattan plot and the uploaded data in Figure 2A appear visually similar, with a few discrepancies. Specifically, both exhibit the nearly the same pattern of significant SNPs along chromosomes 6, 8, 15, and 17, however, the results presented in Figure 2A do not show significant SNPs on chromosomes 2 or 11 (O'Mara et al., 2018; Figure 2A). This can be explained by the filtering steps prior to using this data in the meta-analysis. While the Phelan (2017) and

Lawrenson (2019) studies did not report overall Manhattan plots, the results in Figures 2C and 2E are consistent with the results reported by the respective authors, again considering the results presented here were post-filtering. At present, there are no published GWAS results for cervical and ovarian cancers from the UK Biobank data repository. Genomic control lambdas for the quantile-quantile plots were, at the median for each set, 1.087 (O'Mara), 1.054 (Lawrenson), 1.054 (Phelan), 1.018 (UK Biobank Cervical), and 1.008 (UK Biobank, Ovarian); the O'Mara, Phelan, and UK Biobank (cervical) datasets exhibited an inflation of p-values, with the Lawrenson dataset showing a slight inflation.



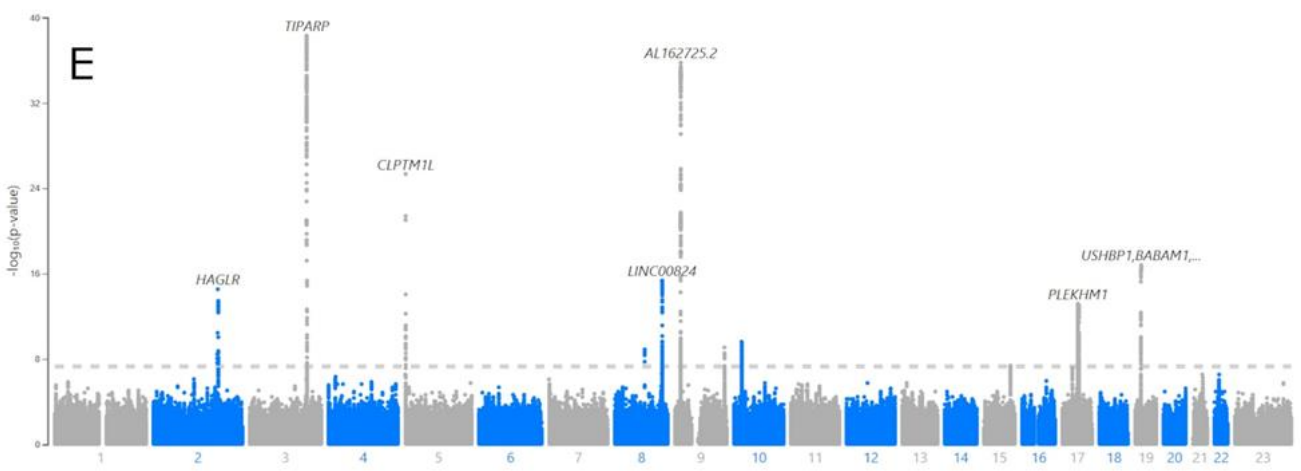
O'Mara et al., 2018

Lambda (0.5): 1.087

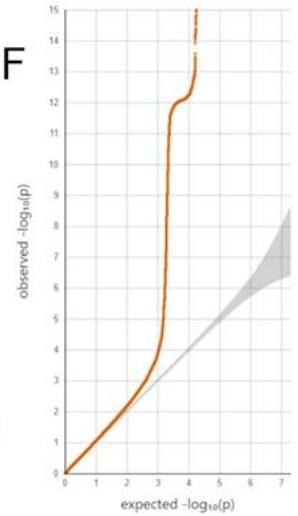


Lawrenson et al., 2019

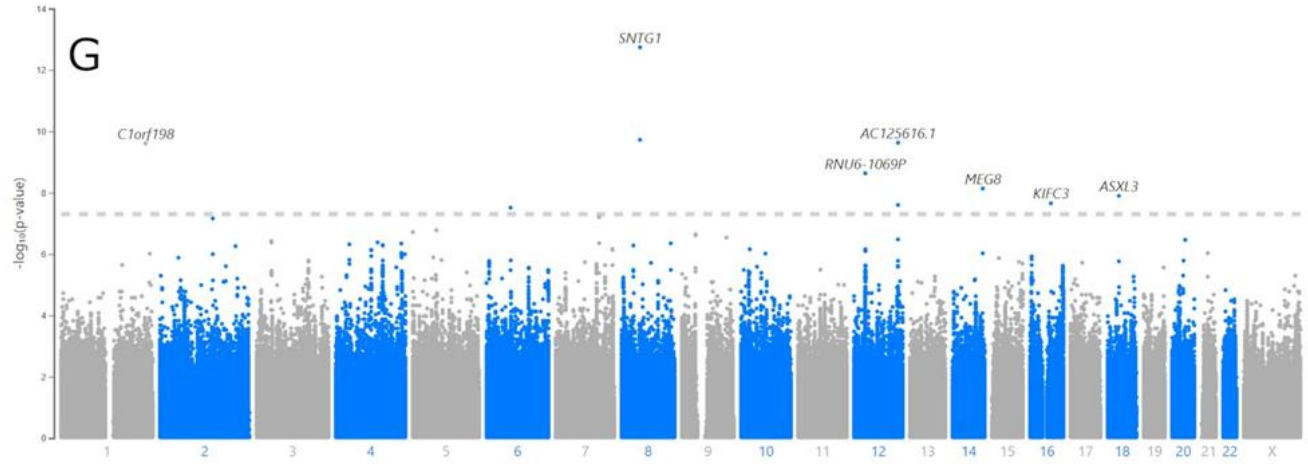
Lambda (0.5): 1.054



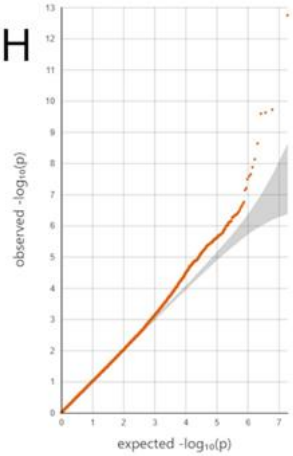
Phelan et al., 2017



Lambda (0.5): 1.054



UK Biobank, Cervical (2018)



Lambda (0.5): 1.018

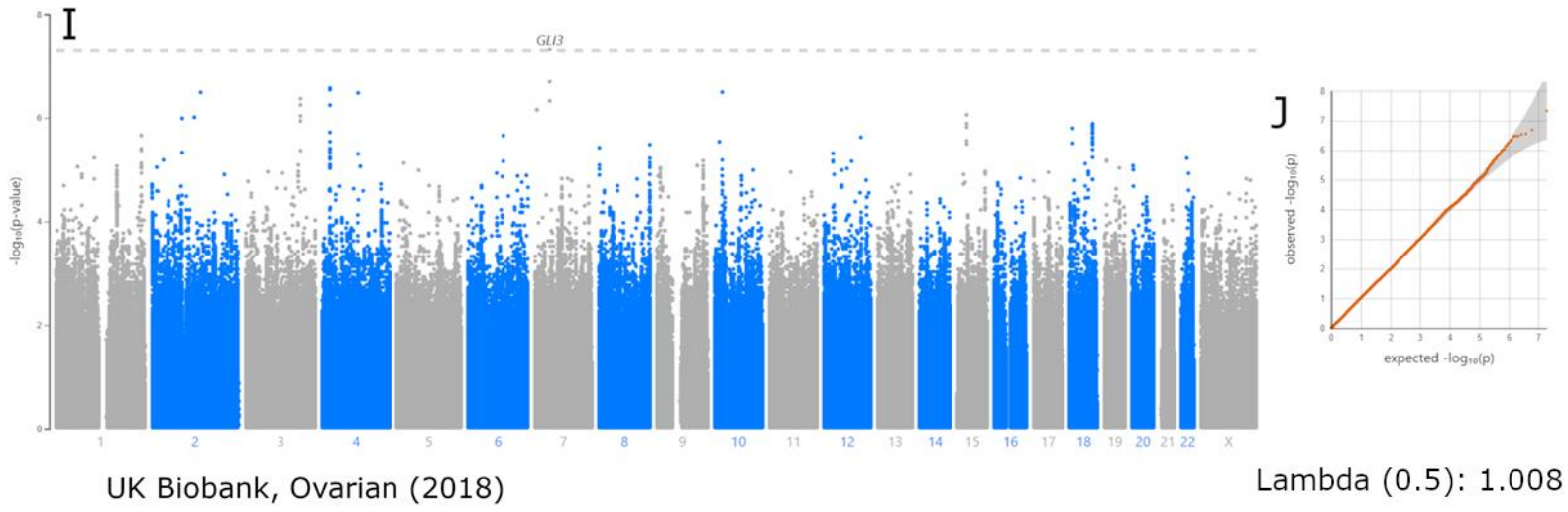


Figure 2: Manhattan and QQ Plots for Individual Data Sources. Data is presented for O'Mara et al., 2018 (2A and 2B), Lawrenson et al., 2019 (2C and 2D), Phelan et al., 2017 (2E and 2F), UK Biobank – cervical cancer (2G and 2H) and UK Biobank – ovarian cancer (2I and 2J).

3.2 Meta-Analysis Results

3.2.1 Ovarian+Endometrial Cancer Meta-Analysis

Abbreviated results for the Ovarian+Endometrial meta-analysis are shown in Table 1. In total, there were approximately 2,155 variants that were found to be significantly associated with the outcomes at the level of $p<0.5x10^{-8}$. As such, variants were selected on the basis of being the 'most significant' variant per locus, per chromosome. Heterogeneity results for selected significant markers is presented in Table 1. Graphical results for the overall meta-analysis are presented in Figure 3A and 3B. Locus plots for the most statistically significant variant per significant locus are presented in Figures 4A through 4H, which also include forest plots for the variant. Locus plots display the regional section along the chromosome for variants that were found to be highly significant, as well as the surrounding variants. Variants in locus plots are colored according to their correlation by non-random population association (linkage disequilibrium) relative to the highly significant variant; in the locus plots, warmer colors correspond to higher correlation, with the highly significant variant (or, in the case of missing linkage disequilibrium, a neighboring variant) represented by a purple diamond. Nearby genes characterized by the GWAS Catalog are shown below locus plots.

Each of the top 12 variants in this meta-analysis were present in both O'Mara et al. (2018) and Phelan et al. (2017), and three of the 12 variants were missing from the Lawrenson et al. (2019) dataset (Table 1). Further, some of the effect estimates were relatively small for selected variants

(e.g., the lowest beta value at variant 2:177037831 corresponds to an odds ratio of approximately 0.94), and the direction of the underlying effects for all selected variants was consistent between the O'Mara et al. (2018) and Phelan et al. (2017) studies, with the exception of 17:44212527 (Table 1). Additionally, heterogeneity tests showed high heterogeneity in underlying studies, as most of the variants showed significant heterogeneity tests (Table 1).

Comparing the Manhattan plots for the O'Mara et al. (2018), Phelan et al. (2017), and Lawrenson et al. (2019) datasets (Figures 2A,2C,2E) with the O+E Manhattan plot (Figure 3A), the meta plot most closely resembles the Phelan data in terms of significant loci. However, both the Phelan et al. and O'Mara et al. datasets overlap a significant locus on chromosome 8, with the nearest annotated gene listed as *LINC00824*. The *TIPARP* locus on chromosome 3, on the other hand, appears in the Phelan et al. (2017) ovarian cancer data, but there were no significant variants for chromosome 3 in the O'Mara et al. (2018) endometrial cancer data. The quantile plot (Figure 3B) exhibits a non-uniform p-value distribution, similar to both the O'Mara et al. (2018) and Phelan et al. (2017) quantile plots (Figures 2B and 2F), contrasting the relatively non-deviant quantile plot of the Lawrenson et al. (2019) data (Figure 2D).

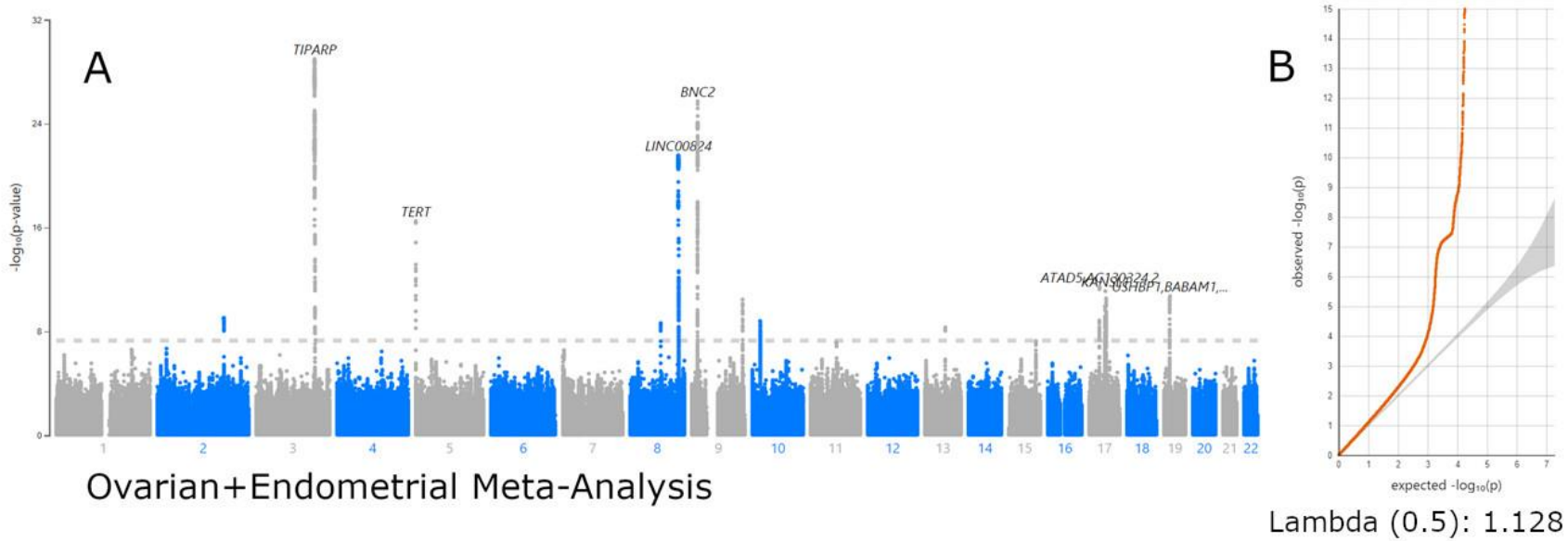
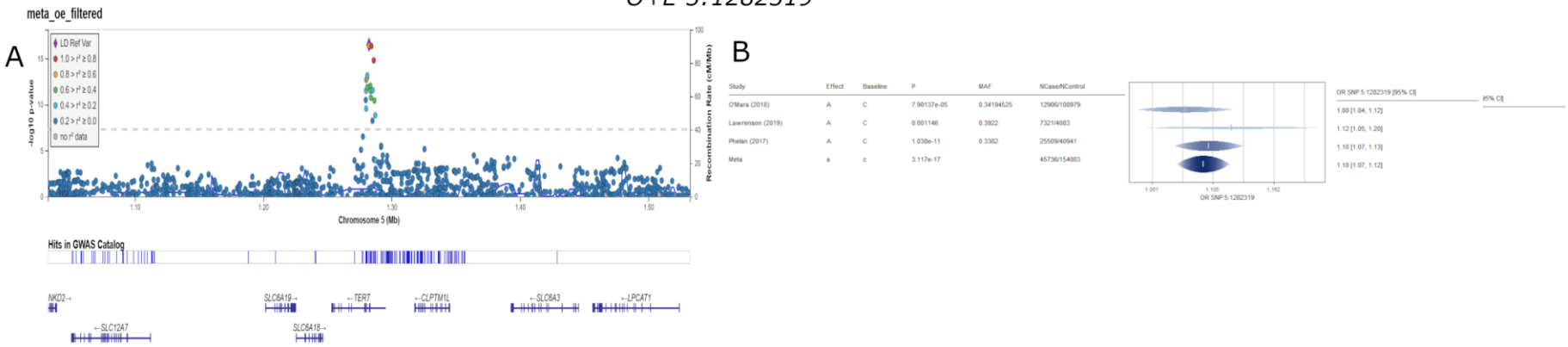
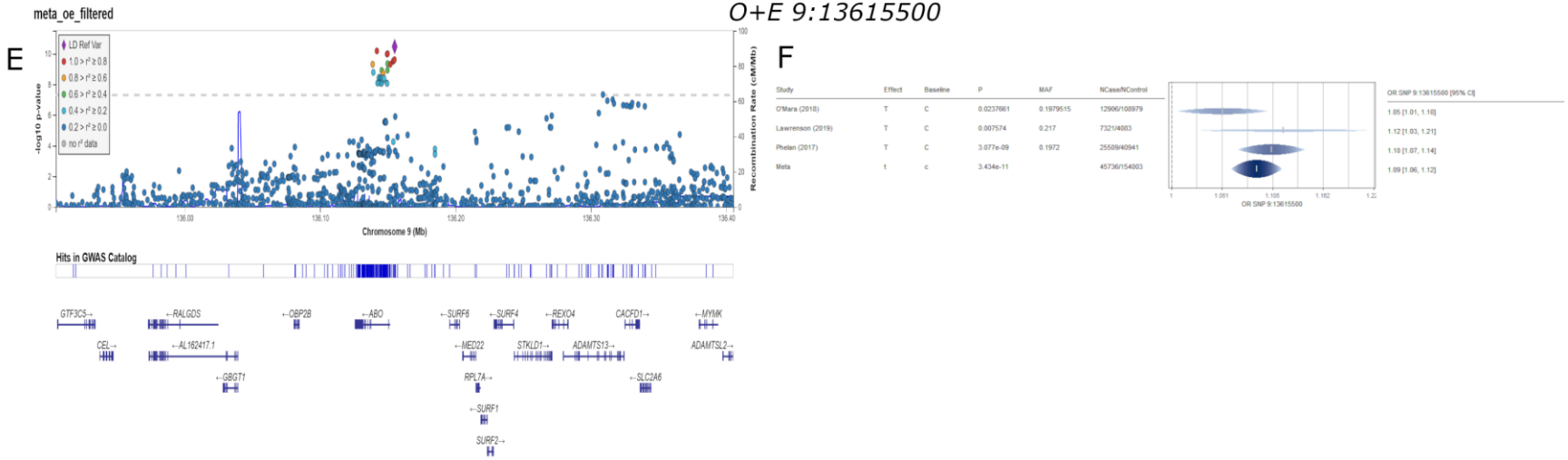


Figure 3: Manhattan and QQ Plots for O+E Meta-Analysis

O+E 5:1282319



O+E 9:13615500



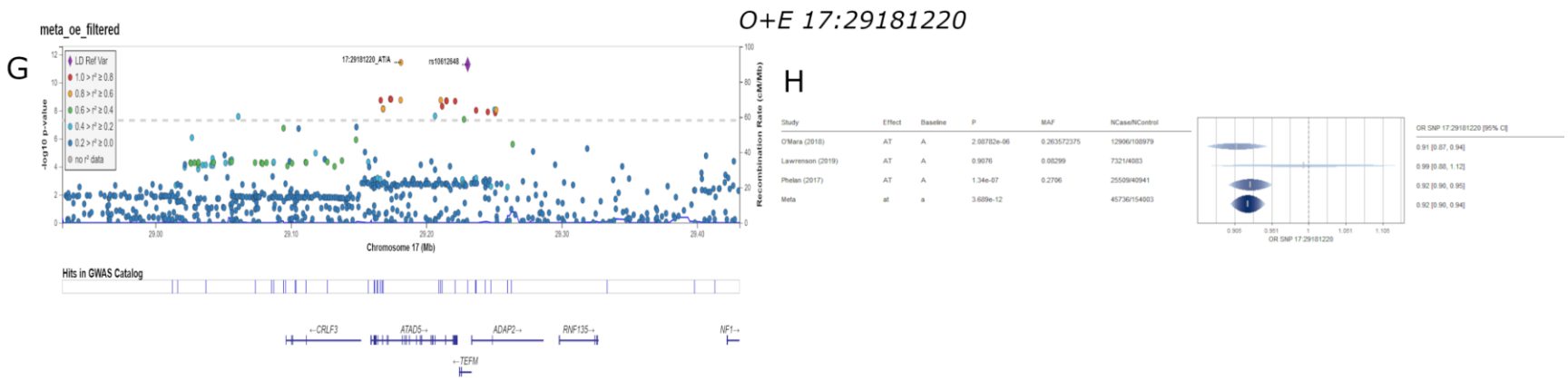


Figure 4: Locus and Forest Plots for High Significant Variants in O+E. Locus and forest plots are shown for O+E variant 5:1282319 (4A and 4B), 8:129541931 (4C and 4D), 9:13615500 (4E and 4F), and 17:29181220 (4G and 4H). For variants without linkage disequilibrium information, a neighbor variant with information was selected; in these cases, both variants are labeled to show the variant of interest and its neighbor.

Table 1: Significant Top Loci Variants for O+E Meta-Analysis

Variant	Chromosome	Position	Effect Allele	Base Allele	Beta	SE	P-Value	Direction*	Het. P-Value
2:177037831	2	177037831	C	G	-0.067	0.011	$9.4x10^{-12}$	---	0.00004
3:156413477	3	156413477	T	C	0.263	0.023	$1.0x10^{-29}$	+?+	$6.1x10^{-11}$
5:1282319	5	1282319	A	C	0.092	0.011	$3.1x10^{-17}$	+++	0.569
8:82653644	8	82653644	A	G	0.122	0.020	$2.3x10^{-9}$	+?+	0.036
8:129541931	8	129541931	A	G	-0.157	0.016	$2.7x10^{-22}$	---	0.870
9:16889937	9	16889937	A	C	-0.132	0.012	$1.9x10^{-26}$	---	$3.3x10^{-10}$
9:136155000	9	136155000	T	C	0.084	0.013	$3.4x10^{-11}$	+++	0.175
10:21821274	10	21821274	A	G	0.068	0.011	$1.5x10^{-9}$	+-+	0.043
13:73814891	13	73814891	T	G	-0.070	0.012	$4.8x10^{-9}$	-+-	$5.3x10^{-8}$
17:29181220	17	29181220	A	AT	-0.082	0.012	$3.7x10^{-12}$	---	0.370
17:44212527	17	44212527	A	G	0.106	0.015	$7.8x10^{-12}$	-?+	0.045
19:17390917	19	17390917	C	G	-0.069	0.010	$1.9x10^{-11}$	---	0.0002

***'Direction' column represents sign of the effect estimate in (left to right): O'Mara et al. (2018), Lawrenson et al. (2019), Phelan et al. (2017). The '?' indicates a missing marker from the input study.**

The four markers that did not demonstrate significant heterogeneity were the variants with significant associations in the O'Mara et al. (2018) study and one or both of the Phelan et al. (2017) and Lawrenson et al. (2019) studies. As such, these variants are explored here. Forest plots for these SNPs (Figures 4B,D,F,H) demonstrate that for the four selected significant meta effect estimates, the underlying Phelan et al. (2017) effect estimates were significantly associated with ovarian cancer in all four variants while the underlying Lawrenson et al. (2019) effect estimates were only found to be significantly associated with ovarian cancer in two of the variants: 9:13615500 and 5:1282319 (Figures 4B and 4F). Similarly, the underlying O'Mara et al. (2018) effect estimates were found to be significantly associated with endometrial cancer in all four variants. Further significant variants are shown in Appendix A. Effect estimates that were significant across underlying individual studies were slight in magnitude, resulting in relatively weak meta estimates where significantly associated with either cancer outcome. Eight of the 12 variants demonstrated significant heterogeneity, indicating these markers may not provide evidence of shared effects.

3.2.2 Ovarian+Cervical Cancer Meta-Analysis Results

There were five variants significantly associated with ovarian+cervical cancer at the threshold of p-value$<5x10^{-8}$, two of which were associated with the same locus. Significant variants and heterogeneity test results are presented in Table 2. Graphical results for the overall analysis and the significant variants are presented in Figures 5A, 5B, and 6A through 6H.

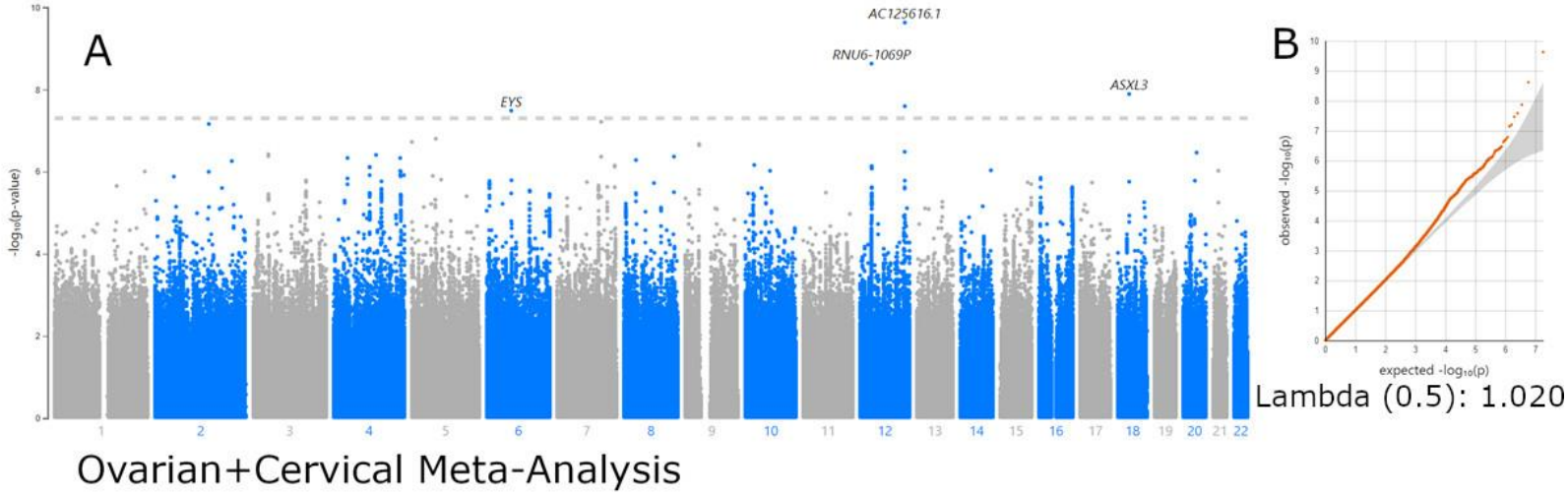
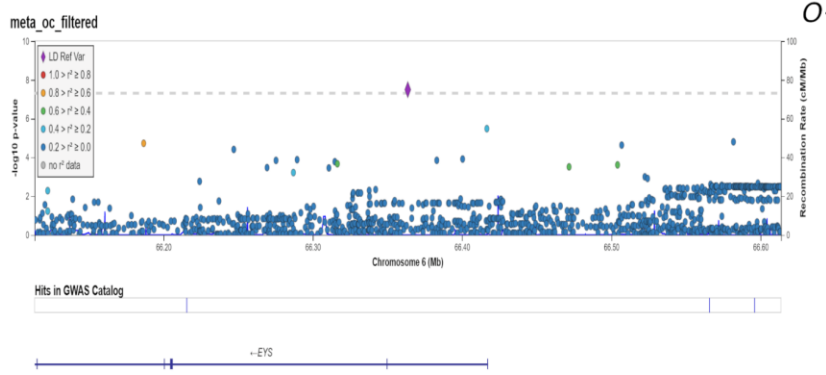


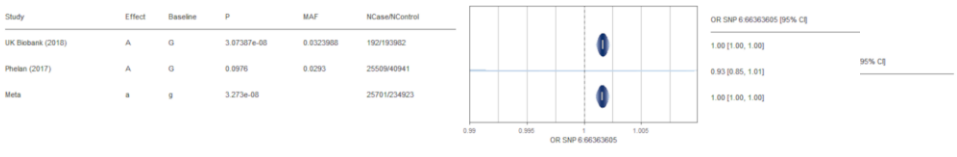
Figure 5: Manhattan and QQ Plots for O+C Meta-Analysis

A

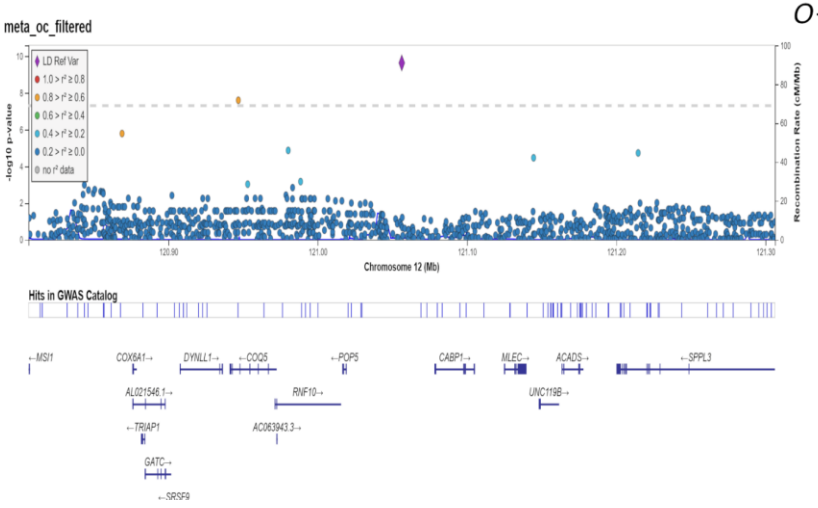


O+C 6:66363605

B

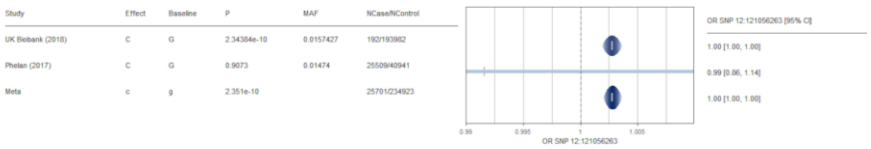


E



O+C 12:121056263

F



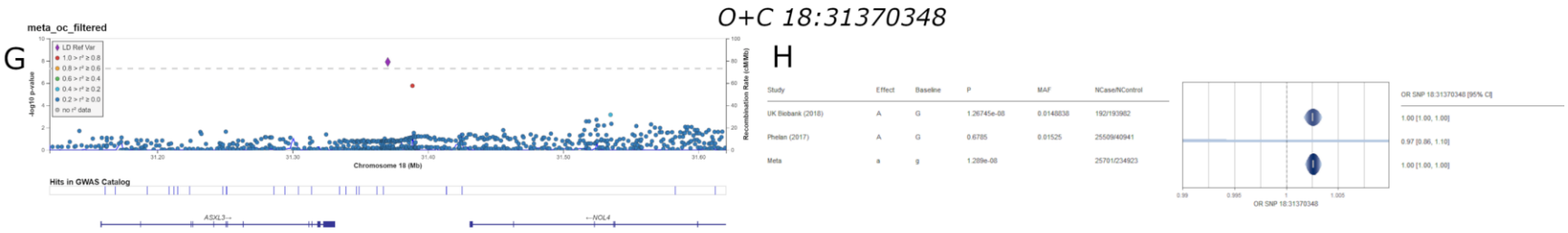


Figure 6: Locus and Forest Plots for High Significant Variants in O+C. Locus and forest plots are shown for O+C variants 6:66363605 (6A and 6B), 12:31954662 (6C and 6D), 12:121056263 (6E and 6F), and 18:31370348 (6G and 6H). For variants without linkage disequilibrium information, a neighbor variant with information was selected; in these cases, both variants are labeled to show the variant of interest and its neighbor.

Table 2: All Significant Variants for O+C Meta-Analysis

Variant	Chromosome	Position	Effect Allele	Base Allele	Beta	SE	P-Value	Direction*	Het. P-Value
6:66363605	6	66363605	A	G	0.002	0.0003	$3.3x10^{-8}$	+?-	0.090
12:31954662	12	31954662	T	G	0.002	0.0004	$2.3x10^{-9}$	+?-	0.791
12:120946783	12	120946783	A	G	0.003	0.0005	$2.5x10^{-8}$	+?-	0.687
12:121056263	12	121056263	C	G	0.003	0.0004	$2.4x10^{-10}$	+?-	0.877
18:31370348	18	31370348	A	G	0.003	0.0004	$1.3x10^{-8}$	+?-	0.649

***'Direction' column represents sign of the effect estimate in (left to right): UK Biobank (cervical) (2018), Lawrenson et al. (2019), Phelan et al. (2017).**

The '?' indicates a missing marker from the input study.

The Manhattan plot (Figure 5A) shows three of the five significant variants of interest were located on chromosome 12. The quantile plot demonstrates that the distribution of p-values was fairly uniform (Figure 5B). Few of the significant variants (or their neighbor, in the case of 12:31954662) showed linkage disequilibrium with nearby variants. All significant variants for this analysis showed opposite direction in underlying effects for the input data sources. Variant 6:66363605 exhibited moderate linkage disequilibrium with a few neighbors (Figure 6A). Variant 12:31954662 showed high linkage disequilibrium with one neighbor, as did variant 18:31370348, and variant 12:121056263 showed low to moderate linkage disequilibrium with several nearby variants (Figures 6C,E,G). Forest plots (along with Table 2) showed that effect estimates and standard errors for meta estimates were extremely weak, such that estimates and confidence intervals in forest plots overlapped by rounding. Heterogeneity tests showed low heterogeneity between individual studies and no tests were significant at the threshold of $p<0.05$ (Table 2); however, the direction of effect is opposite for each of the top variants. As such, there is little between-study heterogeneity, however, there were no top significant variants with significant underlying effect estimates from either the Phelan et al. (2017) or Lawrenson et al. (2019) datasets.

3.2.3 Ovarian-Only Meta-Analysis Results

Overall results for the ovarian-only meta-analysis are shown in the Manhattan and quantile plots, Figures 7A and B. Results for other top significant variants (i.e., those with only one underlying significant effect estimate from either the Phelan et al. (2017) or Lawrenson et al. (2019) datasets) are presented in Table 3. No significant results were generated with all three datasets included (specifically, the UK Biobank ovarian cancer dataset was absent in the 39 total significant variants), likely the result of genotyping array differences (Table 3, Appendix A).

The majority of significant SNPs were located on chromosome 17, however, there were a few significant SNPs in chromosomes 2, 3, 9, and 19 (Figure 7A). The quantile plot showed that the distribution of p-values was not uniform (Figure 7B). Comparing the Manhattan plot with that of the Lawrenson et al. (2019) data in Figure 2E, the cluster of significant SNPs on chromosome 17 seems to be shared between the two; that is, the meta estimates were weighted in favor of the significant SNPs and less on the Lawrenson et al. (2019) data (Figure 2C).

Since the underlying Lawrenson et al. (2019) effect estimates were not significant for most of the selected significant meta estimates (except in the case of 3:156481585, Figures 8C and 8D), the meta effect estimates were generally highly similar to the Phelan et al. (2017) estimates (Appendix A). This variant showed few other significant variants surrounding it, and there was low linkage disequilibrium in the region. Variants for which one underlying data source was significant and for which both data sources were consistent in effect direction are presented in Figures 8A, 8B, and 8E through 8J. Many of these variants followed a trend similar to the variant in Figure 8C, in that there were few other significant variants in their region, there was generally low linkage disequilibrium, etc.

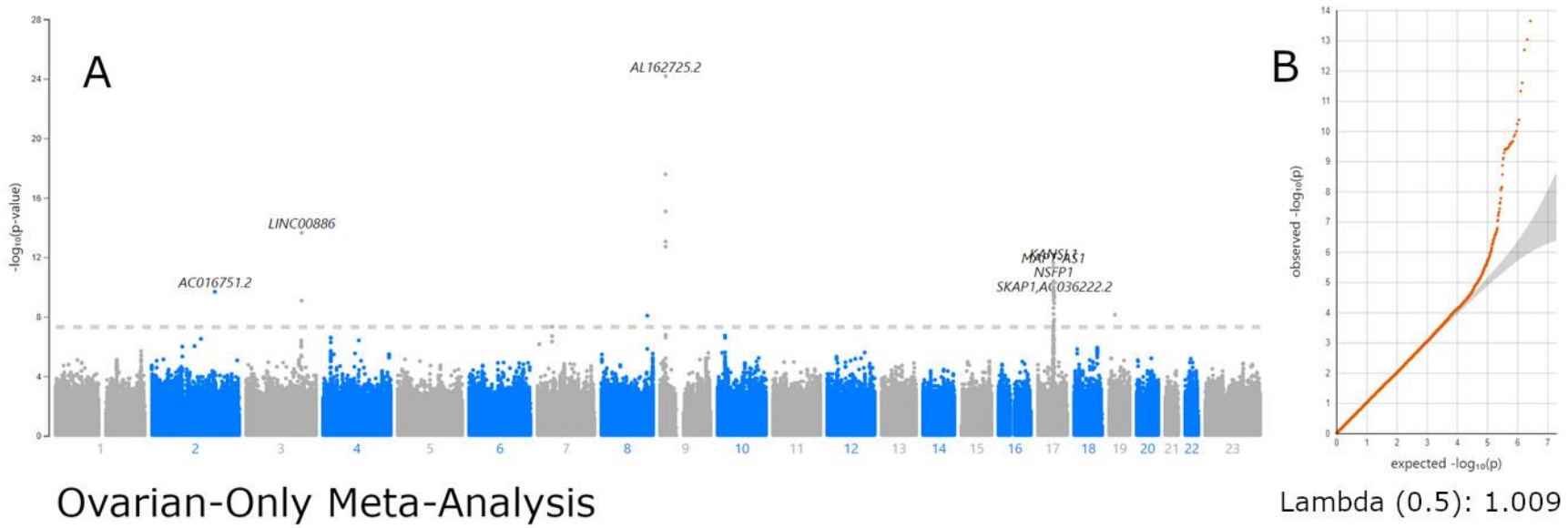
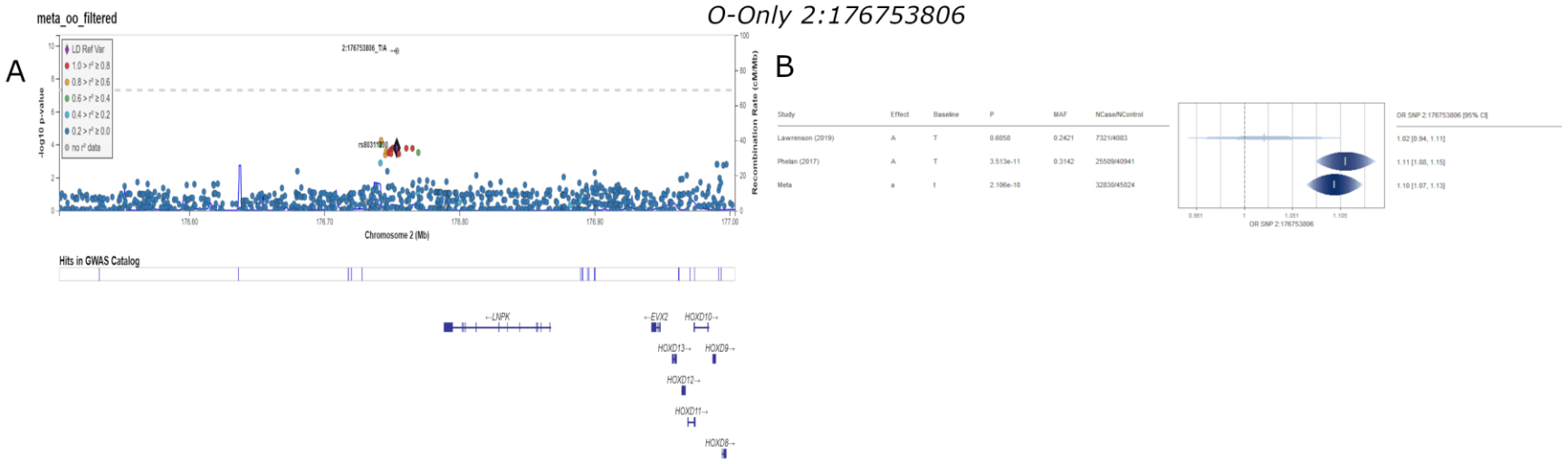
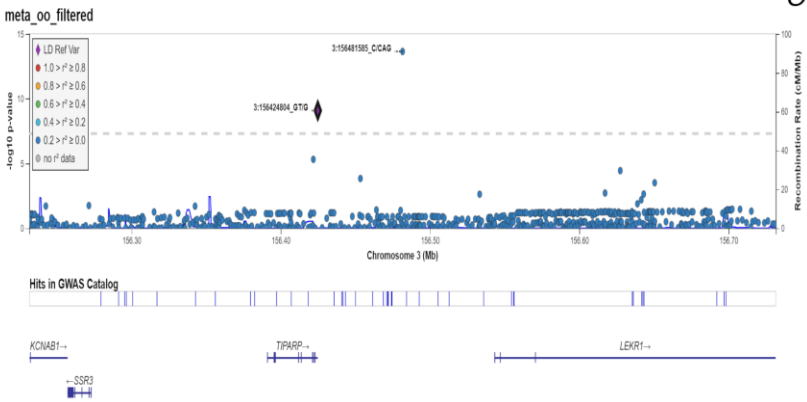


Figure 7: Manhattan and QQ Plot for Ovarian-Only Meta-Analysis



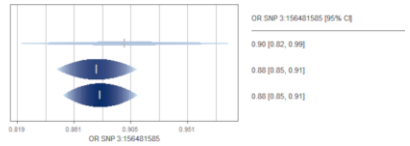
O-Only 3:156481585

C



D

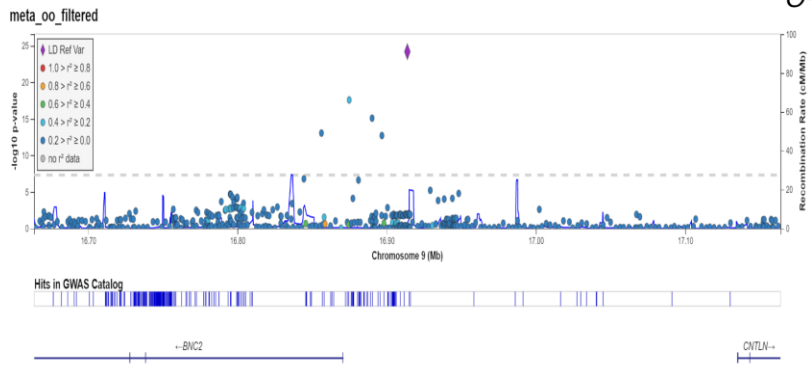
study	Effect	Baseline	P.value	MAF	ncc
Lawrenson (2019)	C	CAG	0.02298	0.1534	7321/4083
Phelan (2017)	C	CAG	2.868e-13	0.1588	25509/40941
Meta	c	cag	2.288e-14		32830/45024



OR SNP 3:156481585 [95% CI]
0.90 [0.82, 0.99]
0.88 [0.85, 0.91]
0.88 [0.85, 0.91]

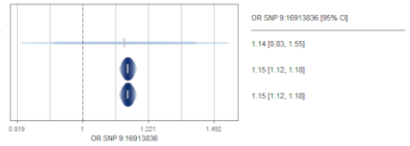
O-Only 9:16913836

E



F

Study	Effect	Baseline	P	MAF	NCase/NControl
Lawrenson (2019)	A	T	0.4302	0.01221	7321/4083
Phelan (2017)	A	T	7.55e-25	0.4212	25509/40941
Meta	a	t	6.596e-25		32830/45024



OR SNP 9:16913836 [95% CI]
1.14 [0.83, 1.55]
1.15 [1.12, 1.18]
1.15 [1.12, 1.18]

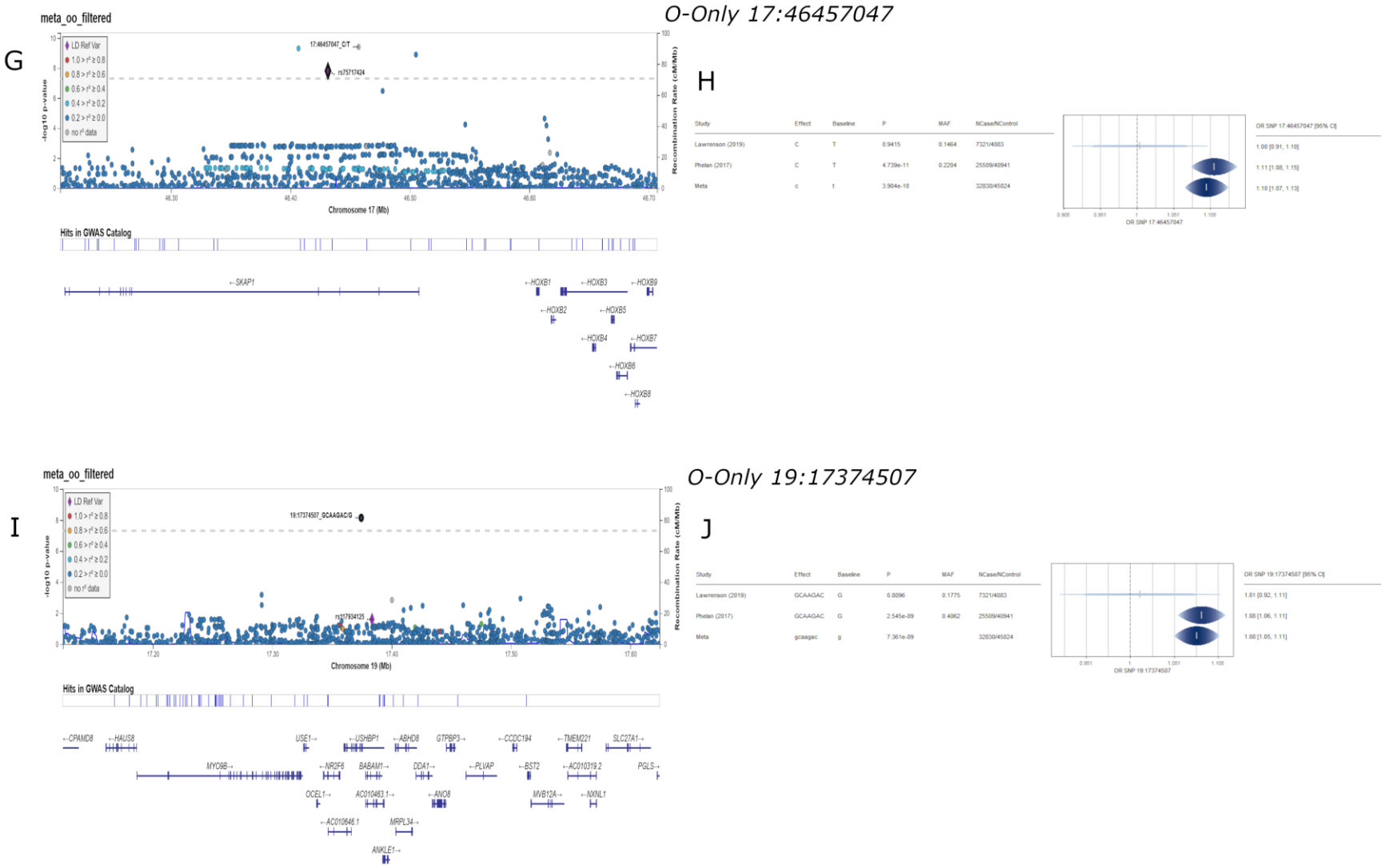


Figure 8: Locus and Forest Plots of the High Significant Variants for Ovarian-Only Meta-Analysis Locus and forest plots are shown for O-Only variants 2:176753806 (8A and 8B), 3:156481585 (8C and 8D), 9:16913836 (8E and 8F), 17:46457047 (8G and 8H), and 19:17374507 (8I and 8J). For the

variants without linkage disequilibrium information, a neighbor variant with information was selected; in these cases, both variants are labeled to show the variant of interest and its neighbor.

Table 3: Significant Top Loci Variants for Ovarian-Only Meta-Analysis

Variant	Chromosome	Position	Effect Allele	Base Allele	Beta	SE	P-Value	Direction	Het. P-Value
2:176753806	2	176753806	A	T	0.094	0.015	$2.1x10^{-10}$	?++	0.053
3:156481585	3	156481585	CAG	C	-0.127	0.017	$2.3x10^{-14}$	?--	0.624
8:129480083	8	129480083	G	GT	-0.113	0.020	$8.5x10^{-9}$	?+-	0.137
9:16913836	9	16913836	A	T	0.138	0.013	$6.6x10^{-25}$	?++	0.946
17:43956139	17	43956139	AAAAAT	AAAAATAAAT	0.113	0.016	$4.7x10^{-12}$	?-+	0.018
17:44260644	17	44260644	A	AT	0.111	0.016	$2.5x10^{-12}$	?-+	0.151
17:44517766	17	44517766	A	T	0.120	0.018	$5.5x10^{-11}$	?-+	0.103
17:46457047	17	46457047	T	C	0.095	0.015	$3.9x10^{-10}$	?++	0.041
19:17374507	19	17374507	G	GCAAGAC	0.076	0.013	$7.4x10^{-9}$	?++	0.150

***'Direction' column represents sign of the effect estimate in (left to right): UK Biobank (ovarian) (2018), Lawrenson et al. (2019), Phelan et al. (2017).**

The '?' indicates a missing marker from the input study.

4.0 Discussion

The ovarian+endometrial meta-analysis overall showed the highest number of variants significant at p<5x10-8, and when comparing the Manhattan plots between this meta-analysis and its individual studies, the Phelan et al. (2017) results seem to be responsible for most of the peaks in the pooled results. However, both the Phelan (2017) and O'Mara (2018) show a strong peak in chromosome 8 (*LINC00824* peak), and both show peaks in chromosome 17. At least one variant in the *LINC00824* region has been associated with endometrial cancer, specifically, 8q24.21 (rs4733613-C), also implicated in telomeric processes (Cheng et al., 2016). None of the underlying variants (Phelan, O'Mara, Lawrenson datasets) were significant for the peak at 17:29181220 (Figures 4G, 4H), nor any of the nearby variants, yet the pooled estimates show a significant group of SNPs in this region. Of note, some previously studied shared genetic factors between ovarian and endometrial cancers were not observed in this analysis, specifically, *ANXA4*, *UGT1A1*, *FXYD2*, and *GLRX* identified by Zorn et al. (2005).

The ovarian+endometrial meta-analysis showed several loci associated with genes that have been studied in relation to both ovarian and endometrial cancers. The locus with variant at 5:1282319 (Figures 4A and 4B) was significantly associated with the outcomes across all the input studies, and was located on the *TERT* gene region, which primarily functions in regulation of telomere length, but has been implicated in the development of ovarian cancer (Dratwa et al., 2020). Further, there are several *SLC* genes surrounding this locus, and generally, solute transport genes have been reported as potential tumor suppressive (Bhutia et al., 2016). Lastly, this locus was also close to the *CLPTM1L* gene, responsible for a transmembrane protein that has been

associated with cancers including lung, pancreas, skin, breast, ovary, and cervix (Li et al., 2020). Further, the *TERT-CLPTM1L* region has been found to host several individual variants associated with increased endometrial cancer risk (Carvajal-Carmona et al., 2015).

The variant at 8:129541931 (Figures 4C and 4D) did not generate nearby gene hits in the GWAS catalog, indicating it may fall in a non-coding region. The 9:13615500 (Figures 4E and 4F) variant showed several genes upstream and downstream of the locus, which itself was located on the *ABO* region. *ABO* has been correlated with both ovarian and endometrial cancers, specifically in women with type A blood groups (Wang-Hong et al., 2011; Gates et al., 2011). There were a number of surfeit genes found surrounding the locus. There is limited evidence suggesting *SURF* genes may be related in the development of cancers – specifically, one study found increased expression in *SURF4* in tumor tissue (Kim et al., 2018). Additionally, the nearby *GBGT1* region has been associated with ovarian cancer, but has not specifically been implicated in endometrial cancers (Jacob et al., 2014). Another regional gene, *REXO4*, exhibits exonuclease activity that may be related to upregulation and DNA alteration in breast cancers (Krishnamurthy et al., 2011). Similarly, increased expression of the *CACFD1* gene, another nearby gene to this locus, has been linked to increased tumor growth in breast cancers (Madan et al., 2019). Additionally, the *BNC2* gene, identified by Cesaratto et al. (2016) as a protective factor against ovarian cancer, was also identified in this analysis (variant 9:16889937), however, this variant was only significant for the underlying ovarian estimates, and not the endometrial estimate (Appendix A). Given the number of significant results identified in this analysis, there are a number of variants and genes that may be potentially shared across ovarian and endometrial cancer genesis and development.

The variant at 17:29181220 (Figures 4G and 4H) was located in the *ATAD5* region, a gene which has been associated with both endometrial and ovarian cancers (Kostovska et al., 2016; Kuchenbaecker et al., 2015). One study suggested that increased expression of *TEFM*, a gene close by this locus, was related to elevated growth and metastasis of hepatocellular carcinomas (Wan et al., 2021). Another nearby gene, *ADAP2*, was recognized in one study by weighted gene co-expression network, to be a hub gene for acute myeloid leukemia (Yu et al., 2020). Lastly, another nearby gene to this locus, *RNF135*, was thought to be downregulated in association with attenuated glioblastoma proliferation (Liu et al., 2016).

Regarding overall ovarian+endometrial results, one study reported the rearrangement and amplification of *KANSL1* (variant 17:44212527, Appendix A) in ovarian cancer, while another described its role more generally in tumors of the breast, prostate, lung, and in glioblastomas and lymphomas, though the latter study suggested the role is dependent on ancestry (positive for European, negative for Asian and African) (Fejzo et al., 2021; Zhou et al., 2017). *SKAP1* (variant 17:44212527, Appendix A) has been implicated in both ovarian and breast cancers, which has been thought to function in the regulation of T-cells, though could potentially affect tumorigenesis regardless of cell-type (Marcotte et al., 2012). Some genes near the 19:17390917 locus (Appendix A) have been studied in relation to gynecologic cancers, specifically, *ANKLE1* and *ABHD8* have both been associated with ovarian cancer, and *MRPL34* with breast cancer. Increased expression of *ABHD8* was found to be associated with reduced breast and ovarian cancer cell activity, while *ANKLE1* demonstrated increased growth and proliferation of both breast and ovarian cancer cells with increased expression (Lawrenson et al., 2016). *MRPL34* was found to have an association

with higher risk and lower survival rate in breast cancer when it was recorded at lower expression levels (Baxter et al., 2018).

The ovarian+cervical meta-analysis was overall unsuccessful at identifying compelling significant variants. That is, the few significant variants were only significant for one underlying dataset. Further, these variants were 'localized', in that they did not show other significant variants in their vicinity. As such, it is difficult to conclude shared genetic factors from this specific analysis. It is possible that cervical cancers share fewer genetic factors with ovarian cancers than endometrial cancers, for example, as it is predominantly related to HPV (with minor relation to Peutz-Jeghers syndrome), while the other gynecologic cancers may have a greater number of risk factors (Ledford and Lockwood, 2019; Constantinou and Tischkowitz, 2017).

Regarding overall results for the ovarian+cervical meta-analysis, the *EYS* gene (variant 6:66363605, Figures 6A and 6B) has not been widely implicated in association with cancers. The *RNU6-1069P* (variant 12:31954662, Figures 6C and 6D) and *AC125616.1* (variant 12:121056263, Figures 6E and 6F) genes have also not been studied in relation to tumors of gynecologic or other origin. The *ASXL3* gene (variant 18:31370348, Figures 6G and 6H), on the other hand, has been assessed for its role in respiratory cancer in humans in one study which found its high expression to be associated with increased copy number in cancer cells, as well as that it targeted and inhibited malignant growth (Shukla et al., 2017). Therefore, these may provide opportunities for follow-up assessment to determine any potential relationship between these variants and either ovarian and/or cervical cancers.

Results from the ovarian-only meta-analysis were somewhat mixed. Most of the highly significant variants per locus were only significant in the underlying Phelan et al. (2017) data, and none of the significant variants were present in the filtered ovarian cancer UK Biobank (2018) data. This is most likely the result of combining genotyped data from different arrays. That is, the UK Biobank incorporated data genotyped on an array purposed for a wide variety of endpoints, where the other three data sources used cancer-specific genotyping arrays (Bycroft et al., 2018; Phelan et al., 2017; Lawrenson et al., 2019; O'mara et al., 2018). As such, some alleles may have been coded differently for the same variants across platforms.

The *LNPK* gene has not been consistently associated with ovarian cancer (variant 2:176753806, Figures 8A and 8B). While *AL162725.2* (variant 9:16913836, Figure 8E) has not been thoroughly studied in relation with cancers, there are other variants that are significant nearby on chromosome 9 that overlap *BNC2* (Figure 8E). *BNC2* has been associated as a protective gene against ovarian cancer (Cesaratto et al., 2016). *KANSL1* and *SKAP1* (chromosome 17 variants, Appendix A and Figures 8G and 8H) have, as well, been implicated in ovarian cancer (Fejzo et al., 2021; Zhou et al., 2017; Marcotte et al., 2012). Further, the variant at 19:17374507 (Figure 8I) was very close to the genes *ABHD8, ANKLE1,* and *MRPL34*, which were identified as ovarian cancer factors in the ovarian+endometrial analysis. Pan et al. (2018) examined the role of *MAPT-AS1* (chromosome 17 variants, Appendix A) in relation to breast cancer cells, and found that it was correlated negatively with tumor cell growth and expansion, and theorized that its overexpression may protect *MAPT* transcripts (Pan et al., 2018).

The variant at 3:156481585 did not show linkage disequilibrium with its neighbor variants (Figure 8C). Heterogeneity tests for the remaining variants were largely not significant (Table 3). The heterogeneity test for this variant was not significant, indicating little heterogeneity between the Lawrenson et al. (2019) and Phelan et al. (2017) studies for this variant. The selected variant did not appear to be associated with a specific gene region (Figure 8C). However, it had a significant neighbor variant that was associated with the *TIPARP* gene. This gene has been associated with both breast cancer as well as ovarian cancer, with reduced expression being associated with greater ovarian tumor development (Cheng et al., 2019; Goode et al., 2010). Permuth et al. (2016) demonstrated that another nearby gene region, *LEKR1*, was associated with epithelial ovarian cancer. Additionally, Ramakrishna et al. (2010) demonstrated that *KCNAB1*, *TIPARP*, and *SSR3*, close genes to this locus, were associated with promoted growth of ovarian cancer cells by increased expression.

The ovarian-only meta-analysis showed some significant genes that were also identified by the ovarian+endometrial meta-analysis, chiefly, *BNC2, KANSL1* and *SKAP1*, all of which have been found to be associated in some manner with ovarian cancers (Cesaratto et al., 2016; Fejzo et al., 2021; Zhou et al., 2017; Marcotte et al., 2012). This analysis additionally identified the *TIPARP*, *ABHD8, ANKLE1*, and *MRPL34* genes, also found in the ovarian+endometrial analysis. As such, these findings may underscore the importance of these genes in ovarian cancers, with potential importance to endometrial cancer.

The majority of genes that were found to be significant in these analyses that were associated with a gynecologic cancer were associated with ovarian cancer, specifically. This might

be logical given that ovarian cancer datasets were included in all analyses. However, this does not necessarily indicate that those genes found related to ovarian cancer are only related to ovarian cancer. Those genes may be associated with ovarian or the other gynecologic cancer involved in the model, respectively. Further, some of the genes that have not been documented in association with any cancer, may truly be associated with one or more of the gynecologic cancer sub-types. There were variants in gene groups identified in this analysis that may be influential for more than one gynecologic cancer sub-type, particularly those that have not been heavily researched in connection with gynecologic cancers like the *SLC* and *SURF* gene groups. These potential factors supplement previous findings of shared genetic variants, like those reported on by Wang et al. (2018) (e.g., *MCM2*, *MMP2*, *COLIA1*, *JUN*). As such, these results may add to the existing evidence that gynecologic cancers share some genetic variants and gene groups.

4.1 Limitations

There are several limitations that are imperative to discuss relative to these analyses. Primarily, there were very few datasets to incorporate. This can lead to some individual results exerting high influence over the pooled estimates, which occurred in some of these analyses. That is, results are not very informative if they are driven primarily by one input dataset, with missing variants in other datasets. This was particularly the case evident in the ovarian+cervical cancer dataset, where few significant results were seen, and those that were significant were not able to be assessed in the Lawrenson et al. (2019) data. For these significant variants, the minor allele frequencies in Asian ancestry individuals is small and would not have been included at the level of minor allele frequency greater than 0.01. While there was data from the UK Biobank cervical

cancer dataset present in the five significant variants, there was no data assessed in the significant results from the UK Biobank in the ovarian-only analysis, perhaps due to genotyping performed on different arrays, limiting power and generalizability of these results.

Further, there were no vaginal cancer or vulvar cancer datasets, and due to some sample overlap, each analysis contained the same two ovarian cancer datasets; therefore, there is little to conclude about the breadth of gynecologic cancer outcomes.

Heterogeneity was also an issue for these analyses. Some variants that were found significant in the models showed significant heterogeneity tests, indicating that those variants are not good representations of true shared genetic architecture. In a broader sense, the liberties taken with the homogeneity assumption of meta-analysis means that variants can only be suggested for effect, and that no conclusions can be drawn about the magnitude of association for any of the cancer outcomes (with the exception of the ovarian-only meta-analysis). Further, combining datasets across ethnicity groups may impact the combinability of some variants, since ancestry greatly influences allelic frequencies and common polymorphisms. Since the Lawrenson et al. (2019) data was performed on individuals of Asian ancestry, this may explain some of the missingness in significant results (specifically, in the ovarian+cervical analysis), as well as the generalizability of results, as causal variants and genetic risk factors may not be consistent across ancestry groups. The strengths and weaknesses of applying the fixed effects paradigm to a cross-trait analysis are discussed in Section 2.2. In essence, the trade-off of discovering as many potentially shared variants as possible (increasing statistical power by using fixed effects) is limited by the inability to conclude a true effect size for multiple phenotypic outcomes.

Additionally, because only summary results were available, no other verifications of underlying genotyping data or follow-up analyses were possible. Further, since per-SNP sample sizes were not included for the majority of datasets, a sample-weighted analysis was not possible. Lastly, sample size imbalances were observed across the studies. Specifically, the Phelan et al. (2017) dataset provided approximately 25,000 cases, whereas the UK Biobank dataset reported only having less than 1,000 cases in both the ovarian and cervical datasets. As such, some of the results were likely dominated by the larger datasets, that is, the results may be biased toward the larger datasets.

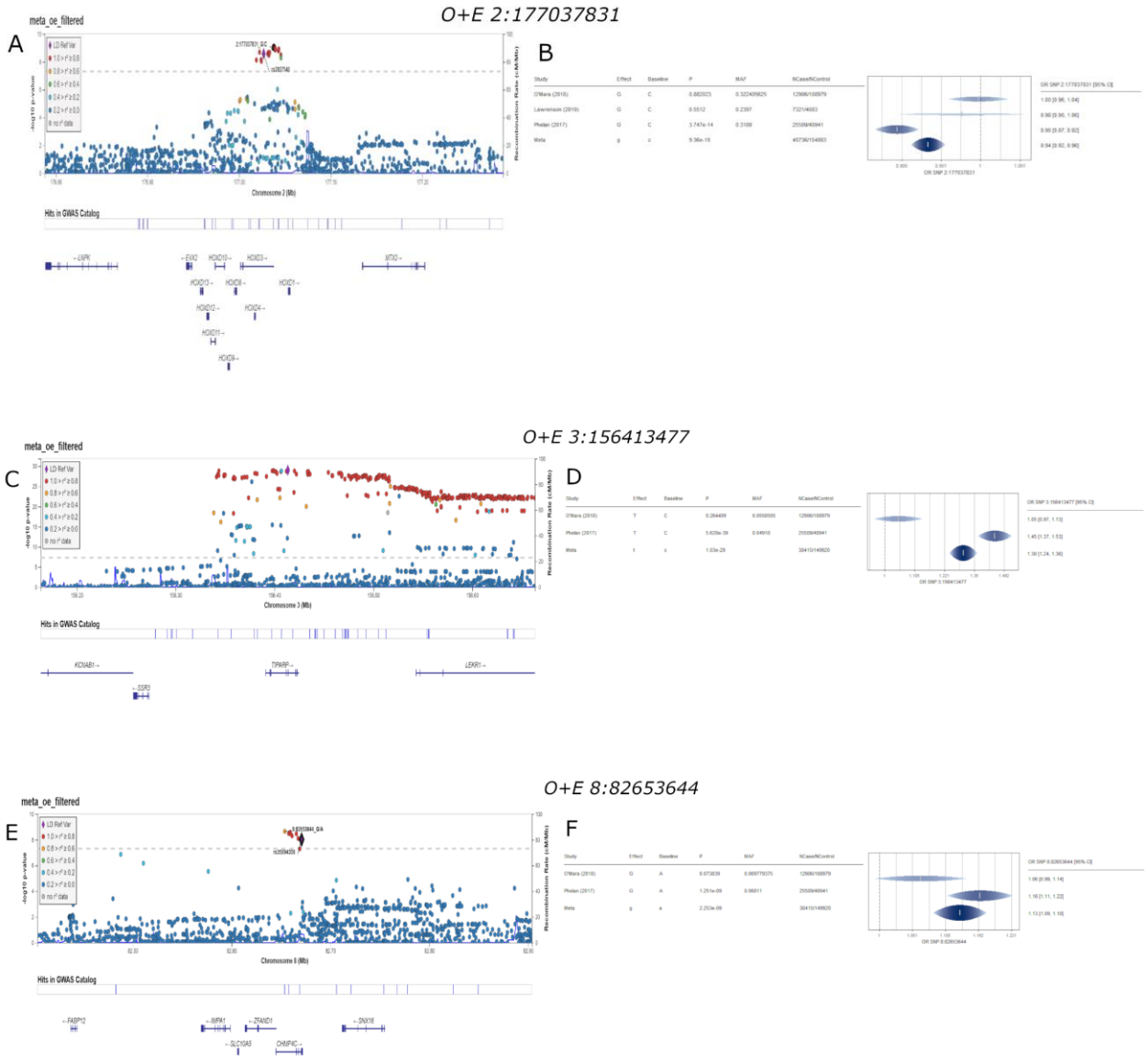
4.2 Conclusion

In conclusion, these analyses summarized pooled estimates across combinations of ovarian cancer datasets with one endometrial and one cervical cancer dataset. Existing shared effects were verified in the ovarian+endometrial analysis specifically, including *TERT*, *CLPTM1L* and *ATAD5*. Several variants were identified that are well-characterized in terms of ovarian cancer association, but warrant further investigation in potential association with the other cancer outcomes. Examples include, for endometrial cancer, *SLC* genes, *SURF* genes, *ABO*, and *BNC2*, and for cervical cancer, *EYS*, *RNU6-1069P*, *AC125616.1*, and *ASXL3* are genes that may warrant further investigation into potential roles in the genesis and growth of these conditions.

As such, one future research direction may include investigating the nearest genes found to the highly significant variants that were not already documented in association with a

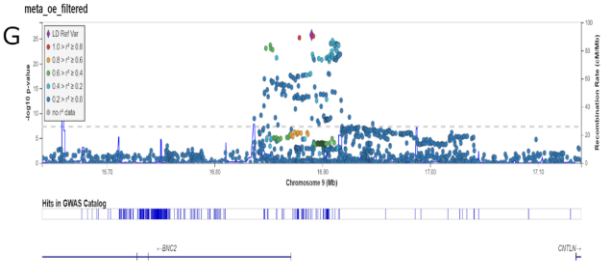
gynecologic cancer. Further research is necessary in the areas of vaginal and vulvar cancer, as well.

Appendix A Other Highly Significant Variants from O+E and O-Only Meta-Analyses

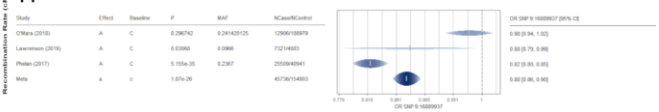


O+E 9:16889937

G

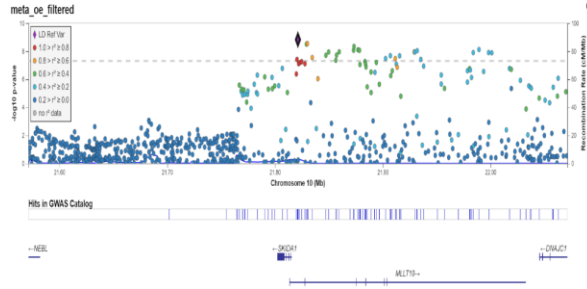


H

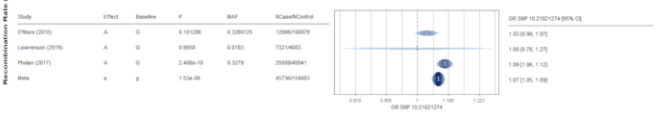


O+E 10:21821274

I

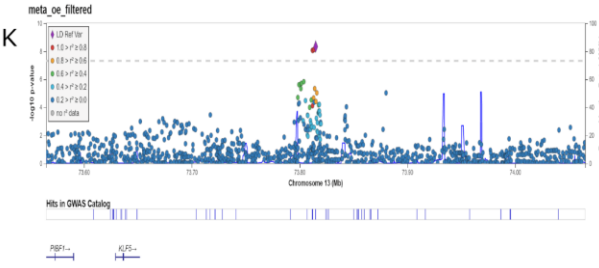


J

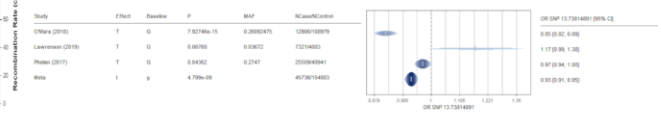


O+E 13:73814891

K

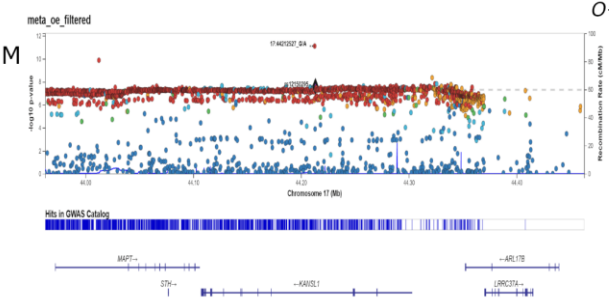


L



O+E 17:44212527

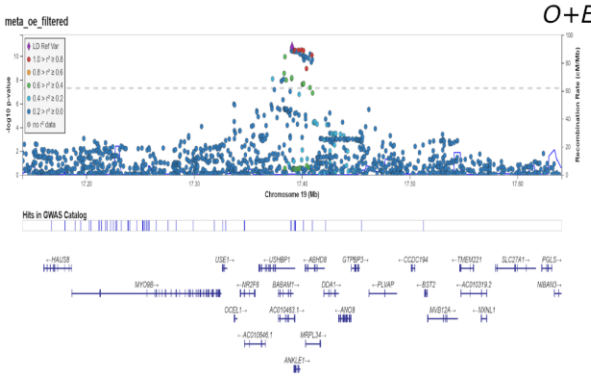
M



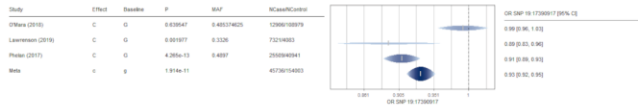
N

O+E 19:17390917

O

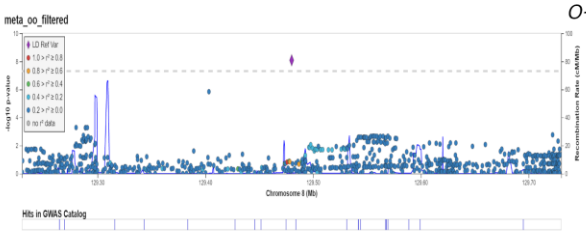


P

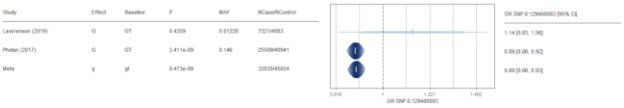


O-Only 8:129480083

Q

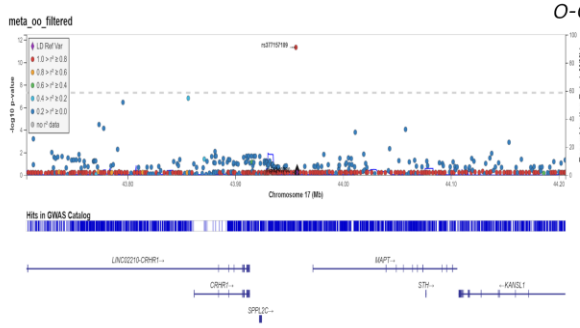


R



O-Only 17:43956139

S



T

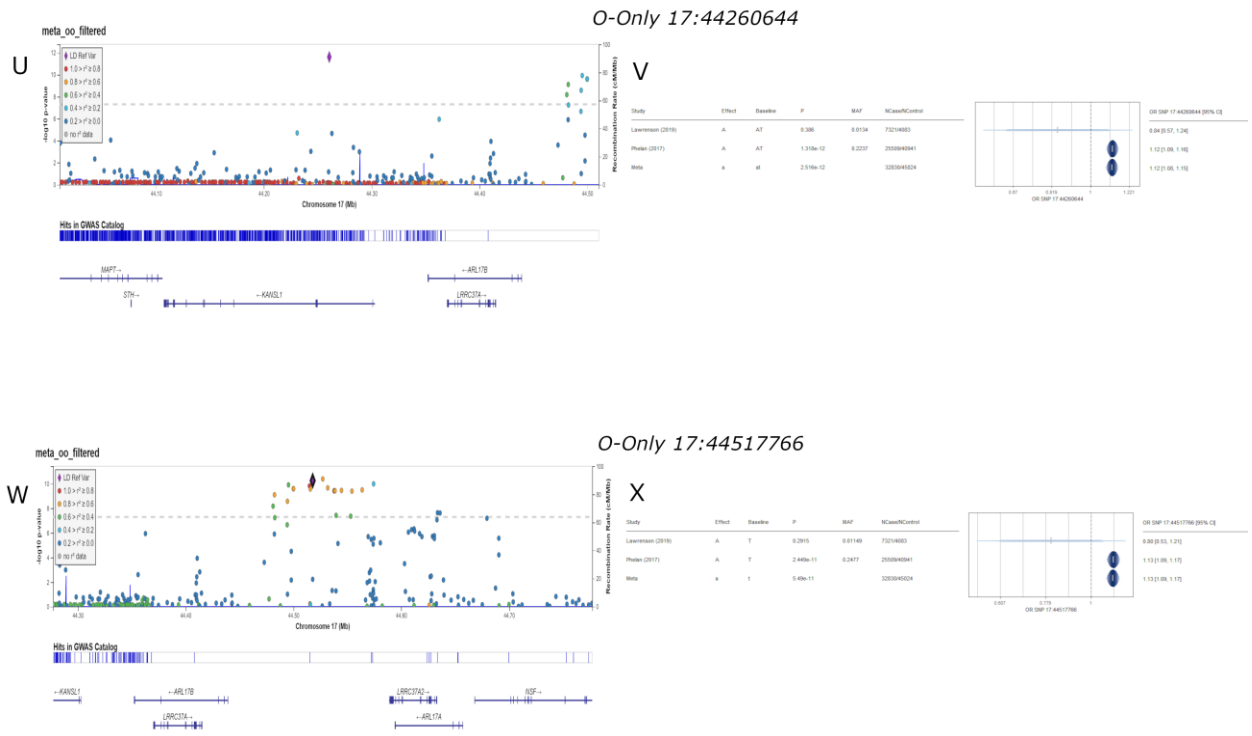


Figure 9: Locus and Forest Plots for Other Highly Significant Variants in the O+E and O-Only Meta-Analyses. Locus and forest plots are shown for O+E variants 2:177037831 (9A and 9B), 3:156413477 (9C and 9D), 8:82653644 (9E and 9F), 9:16889937 (9G and 9H), 10:21821274 (9I and 9J), 13:73814891 (9K and 9L), 17:44212527 (9M and 9N), 19:17390917 (9O and 9P), and O-Only variants 8:129480083 (9Q and 9R), 17:43956139 (9S and 9T), 17:44260644 (9U and 9V), and 17:44517766 (9W and 9X). For the variants without linkage disequilibrium information, a neighbor variant with information was selected; in these cases, both variants are labeled to show the variant of interest and its neighbor.

Appendix B Custom R Code and METAL Scripts

```
#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr1/omara_1.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr1/lawrenson_1.txt
PROCESS /home/Mark/METAL/Chr1/phelan_1.txt

OUTFILE META_Chr1_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr1/ukbb_1.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr1/lawrenson_1.txt
PROCESS /home/Mark/METAL/Chr1/phelan_1.txt

OUTFILE META_Chr1_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr1/ukbb_2_1.txt
```

```
MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr1/lawrenson_1.txt
PROCESS /home/Mark/METAL/Chr1/phelan_1.txt

OUTFILE META_Chr1_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr2/omara_2.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr2/lawrenson_2.txt
PROCESS /home/Mark/METAL/Chr2/phelan_2.txt

OUTFILE META_Chr2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr2/ukbb_2.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
PROCESS /home/Mark/METAL/Chr2/lawrenson_2.txt
PROCESS /home/Mark/METAL/Chr2/phelan_2.txt

OUTFILE META_Chr2_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
```

```
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr2/ukbb_2_2.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr2/lawrenson_2.txt
PROCESS /home/Mark/METAL/Chr2/phelan_2.txt

OUTFILE META_Chr2_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr3/omara_3.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr3/lawrenson_3.txt
PROCESS /home/Mark/METAL/Chr3/phelan_3.txt

OUTFILE META_Chr3_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr3/ukbb_3.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRBABEL overall_SE
PROCESS /home/Mark/METAL/Chr3/lawrenson_3.txt
PROCESS /home/Mark/METAL/Chr3/phelan_3.txt

OUTFILE META_Chr3_c_  .txt
ANALYZE HETEROGENEITY

QUIT
```

```
#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr3/ukbb_2_3.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr3/lawrenson_3.txt
PROCESS /home/Mark/METAL/Chr3/phelan_3.txt

OUTFILE META_Chr3_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr4/omara_4.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr4/lawrenson_4.txt
PROCESS /home/Mark/METAL/Chr4/phelan_4.txt

OUTFILE META_Chr4_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr4/ukbb_4.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr4/lawrenson_4.txt
PROCESS /home/Mark/METAL/Chr4/phelan_4.txt
```

```
OUTFILE META_Chr4_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr4/ukbb_2_4.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr4/lawrenson_4.txt
PROCESS /home/Mark/METAL/Chr4/phelan_4.txt

OUTFILE META_Chr4_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr5/omara_5.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr5/lawrenson_5.txt
PROCESS /home/Mark/METAL/Chr5/phelan_5.txt

OUTFILE META_Chr5_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr5/ukbb_5.txt

MARKER chrpos
ALLELE Effect Baseline
```

```
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr5/lawrenson_5.txt
PROCESS /home/Mark/METAL/Chr5/phelan_5.txt

OUTFILE META_Chr5_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr5/ukbb_2_5.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr5/lawrenson_5.txt
PROCESS /home/Mark/METAL/Chr5/phelan_5.txt

OUTFILE META_Chr5_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr6/omara_6.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr6/lawrenson_6.txt
PROCESS /home/Mark/METAL/Chr6/phelan_6.txt

OUTFILE META_Chr6_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
```

```
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr6/ukbb_6.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr6/lawrenson_6.txt
PROCESS /home/Mark/METAL/Chr6/phelan_6.txt

OUTFILE META_Chr6_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr6/ukbb_2_6.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr6/lawrenson_6.txt
PROCESS /home/Mark/METAL/Chr6/phelan_6.txt

OUTFILE META_Chr6_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr7/omara_7.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr7/lawrenson_7.txt
PROCESS /home/Mark/METAL/Chr7/phelan_7.txt

OUTFILE META_Chr7_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash
```

```
SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr7/ukbb_2_7.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
PROCESS /home/Mark/METAL/Chr7/lawrenson_7.txt
PROCESS /home/Mark/METAL/Chr7/phelan_7.txt

OUTFILE META_Chr7_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr7/ukbb_7.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr7/lawrenson_7.txt
PROCESS /home/Mark/METAL/Chr7/phelan_7.txt

OUTFILE META_Chr7_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr8/omara_8.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr8/lawrenson_8.txt
PROCESS /home/Mark/METAL/Chr8/phelan_8.txt

OUTFILE META_Chr8_  .txt
ANALYZE HETEROGENEITY
```

```
QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr8/ukbb_2_8.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr8/lawrenson_8.txt
PROCESS /home/Mark/METAL/Chr8/phelan_8.txt

OUTFILE META_Chr8_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr8/ukbb_8.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr8/lawrenson_8.txt
PROCESS /home/Mark/METAL/Chr8/phelan_8.txt

OUTFILE META_Chr8_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr9/omara_9.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
```

```
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr9/lawrenson_9.txt
PROCESS /home/Mark/METAL/Chr9/phelan_9.txt

OUTFILE META_Chr9_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr9/ukbb_9.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr9/lawrenson_9.txt
PROCESS /home/Mark/METAL/Chr9/phelan_9.txt

OUTFILE META_Chr9_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr9/ukbb_2_9.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr9/lawrenson_9.txt
PROCESS /home/Mark/METAL/Chr9/phelan_9.txt

OUTFILE META_Chr9_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr10/omara_10.txt
```

```
MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr10/lawrenson_10.txt
PROCESS /home/Mark/METAL/Chr10/phelan_10.txt

OUTFILE META_Chr10_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr10/ukbb_10.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr10/lawrenson_10.txt
PROCESS /home/Mark/METAL/Chr10/phelan_10.txt

OUTFILE META_Chr10_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr10/ukbb_2_10.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr10/lawrenon_10.txt
PROCESS /home/Mark/METAL/Chr10/phelan_10.txt

OUTFILE META_Chr10_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
```

```
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr11/omara_11.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr11/lawrenson_11.txt
PROCESS /home/Mark/METAL/Chr11/phelan_11.txt

OUTFILE META_Chr11_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr11/ukbb_11.txt

MARKER chrpos
ALLELE Effect Baseline
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr11/lawrenson_11.txt
PROCESS /home/Mark/METAL/Chr11/phelan_11.txt

OUTFILE META_Chr11_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr11/ukbb_2_11.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr11/lawrenson_11.txt
PROCESS /home/Mark/METAL/Chr11/phelan_11.txt

OUTFILE META_Chr11_2_  .txt
ANALYZE HETEROGENEITY

QUIT
```

```
#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr12/omara_12.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr12/lawrenson_12.txt
PROCESS /home/Mark/METAL/Chr12/phelan_12.txt

OUTFILE META_Chr12_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr12/ukbb_2_12.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr12/lawrenson_12.txt
PROCESS /home/Mark/METAL/Chr12/phelan_12.txt

OUTFILE META_Chr12_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr12/ukbb_12.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr12/lawrenson_12.txt
PROCESS /home/Mark/METAL/Chr12/phelan_12.txt
```

```
OUTFILE META_Chr12_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr13/omara_13.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr13/lawrenson_13.txt
PROCESS /home/Mark/METAL/Chr13/phelan_13.txt

OUTFILE META_Chr13_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr13/ukbb_13.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr13/lawrenson_13.txt
PROCESS /home/Mark/METAL/Chr13/phelan_13.txt

OUTFILE META_Chr13_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr13/ukbb_2_13.txt

MARKER chrpos
ALLELE Effect Baseline
```

```
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr13/lawrenson_13.txt
PROCESS /home/Mark/METAL/Chr13/phelan_13.txt

OUTFILE META_Chr13_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr14/omara_14.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr14/lawrenson_14.txt
PROCESS /home/Mark/METAL/Chr14/phelan_14.txt

OUTFILE META_Chr14_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr14/ukbb_14.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr14/lawrenson_14.txt
PROCESS /home/Mark/METAL/Chr14/phelan_14.txt

OUTFILE META_Chr14_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
```

```
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr14/ukbb_2_14.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr14/lawrenson_14.txt
PROCESS /home/Mark/METAL/Chr14/phelan_14.txt

OUTFILE META_Chr14_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr15/omara_15.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr15/lawrenson_15.txt
PROCESS /home/Mark/METAL/Chr15/phelan_15.txt

OUTFILE META_Chr15_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr15/ukbb_15.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr15/lawrenson_15.txt
PROCESS /home/Mark/METAL/Chr15/phelan_15.txt

OUTFILE META_Chr15_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash
```

```
SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr15/ukbb_2_15.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr15/lawrenson_15.txt
PROCESS /home/Mark/METAL/Chr15/phelan_15.txt

OUTFILE META_Chr15_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr16/omara_16.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr16/lawrenson_16.txt
PROCESS /home/Mark/METAL/Chr16/phelan_16.txt

OUTFILE META_Chr16_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr16/ukbb_16.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr16/lawrenson_16.txt
PROCESS /home/Mark/METAL/Chr16/phelan_16.txt

OUTFILE META_Chr16_c_  .txt
```

```
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr16/ukbb_2_16.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr16/lawrenson_16.txt
PROCESS /home/Mark/METAL/Chr16/phelan_16.txt

OUTFILE META_Chr16_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr17/omara_17.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr17/lawrenson_17.txt
PROCESS /home/Mark/METAL/Chr17/phelan_17.txt

OUTFILE META_Chr17_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr17/ukbb_17.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
```

```
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr17/lawrenson_17.txt
PROCESS /home/Mark/METAL/Chr17/phelan_17.txt

OUTFILE META_Chr17_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr17/ukbb_2_17.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr17/lawrenson_17.txt
PROCESS /home/Mark/METAL/Chr17/phelan_17.txt

OUTFILE META_Chr17_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr18/omara_18.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr18/lawrenson_18.txt
PROCESS /home/Mark/METAL/Chr18/phelan_18.txt

OUTFILE META_Chr18_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
```

```
PROCESS /home/Mark/METAL/Chr18/ukbb_18.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr18/lawrenson_18.txt
PROCESS /home/Mark/METAL/Chr18/phelan_18.txt

OUTFILE META_Chr18_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr18/ukbb_2_18.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr18/lawrenson_18.txt
PROCESS /home/Mark/METAL/Chr18/phelan_18.txt

OUTFILE META_Chr18_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr19/omara_19.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr19/lawrenson_19.txt
PROCESS /home/Mark/METAL/Chr19/phelan_19.txt

OUTFILE META_Chr19_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR
```

```
MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr19/ukbb_19.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr19/lawrenson_19.txt
PROCESS /home/Mark/METAL/Chr19/phelan_19.txt

OUTFILE META_Chr19_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr19/ukbb_2_19.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr19/lawrenson_19.txt
PROCESS /home/Mark/METAL/Chr19/phelan_19.txt

OUTFILE META_Chr19_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr20/omara_20.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr20/lawrenson_20.txt
PROCESS /home/Mark/METAL/Chr20/phelan_20.txt

OUTFILE META_Chr20_  .txt
ANALYZE HETEROGENEITY
```

```
QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr20/ukbb_20.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr20/lawrenson_20.txt
PROCESS /home/Mark/METAL/Chr20/phelan_20.txt

OUTFILE META_Chr20_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr20/ukbb_2_20.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr20/lawrenson_20.txt
PROCESS /home/Mark/METAL/Chr20/phelan_20.txt

OUTFILE META_Chr20_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr21/omara_21.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
```

```
PROCESS /home/Mark/METAL/Chr21/lawrenson_21.txt
PROCESS /home/Mark/METAL/Chr21/phelan_21.txt

OUTFILE META_Chr21_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr21/ukbb_21.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr21/lawrenson_21.txt
PROCESS /home/Mark/METAL/Chr21/phelan_21.txt

OUTFILE META_Chr21_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr21/ukbb_2_21.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr21/lawrenson_21.txt
PROCESS /home/Mark/METAL/Chr21/phelan_21.txt

OUTFILE META_Chr21_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE effect_allele other_allele
PVALUE p_value
EFFECT beta
SEPARATOR COMMA
STDERRLABEL standard_error
PROCESS /home/Mark/METAL/Chr22/omara_22.txt
```

```
MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr22/lawrenson_22.txt
PROCESS /home/Mark/METAL/Chr22/phelan_22.txt

OUTFILE META_Chr22_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr22/ukbb_22.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr22/lawrenson_22.txt
PROCESS /home/Mark/METAL/Chr22/phelan_22.txt

OUTFILE META_Chr22_c_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE alt ref
PVALUE pval
EFFECT beta
SEPARATOR COMMA
STDERRLABEL se
PROCESS /home/Mark/METAL/Chr22/ukbb_2_22.txt

MARKER chrpos
ALLELE Effect Baseline
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr22/lawrenson_22.txt
PROCESS /home/Mark/METAL/Chr22/phelan_22.txt

OUTFILE META_Chr22_2_  .txt
ANALYZE HETEROGENEITY

QUIT

#!bin/bash

SCHEME STDERR

MARKER chrpos
ALLELE Effect Baseline
```

```
PVALUE overall_pvalue
EFFECT overall_OR
SEPARATOR COMMA
STDERRLABEL overall_SE
PROCESS /home/Mark/METAL/Chr23/lawrenson_23.txt
PROCESS /home/Mark/METAL/Chr23/phelan_23.txt

OUTFILE META_Chr23_  .txt
ANALYZE HETEROGENEITY

QUIT
```

```
library(data.table)
library(tidyverse)

#Dataframes with combined results in one file:

#Cervical Cancer:
ukbb <- as.data.frame(fread("C53.gwas.imputed_v3.female.tsv"))

#ukbbcont <- as.data.frame(fread("ukbbcont.tsv"))
#^^^ C53 has 192 cases and 193982 controls.
#^^^ C56 has 691 cases and 193483 controls.

#Endometrial Cancer:
omara <- as.data.frame(fread("harmonised.qc.tsv"))

omara_bed <- omara[,c(15,14,13)]
omara_bed <- arrange(omara_bed, omara_bed$chromosome, omara_bed$base_pair_location)
omara_bed$chromStart <- omara_bed$base_pair_location - 1
omara_bed$chromEnd <- omara_bed$base_pair_location
omara_bed <- omara_bed[,c(1,4,5,3)]
omara_bed$chromosome <- paste("chr", omara_bed$chromosome, sep="")
#write.table(omara_bed, sep="\t",file="C:\\Users\\Mark\\Documents\\omara_bed.txt", row.names = FALSE, col.names = FALSE, quote = FALSE)
omara_liftover <- read.table("hglft_results.txt", sep="\t")

omara_liftover <- omara_liftover %>% rename(base_pair_location=V3, variant_id=V4)

omara_merge <- merge(omara, omara_liftover[,c(3,4)], by="variant_id")

omara <- omara_merge
rm(omara_merge)
rm(omara_liftover)
rm(omara_bed)

omara$chrpos <- paste(omara$chromosome, omara$base_pair_location.y, sep=":")
omara <- omara %>% rename(base_pair_location=base_pair_location.y)
#Mapping UKBB rsIDs:
ukvariants <- as.data.frame(fread("variants.tsv"))
ukvariants$chrpos <- paste(ukvariants$chr, ukvariants$pos, sep=":")
ukbb <- merge(x=ukbb, y=ukvariants[,c("variant", "rsid", "chr", "pos", "chrpos", "ref", "alt")], by="variant", all.x=TRUE)

rm(ukvariants)

#Filter MAF < 0.01, nonsense p-val,se:
ukbb <- ukbb %>% filter(minor_AF > 0.01) %>% filter(se > 0) %>%
        filter(pval < 1 & pval > 0)
#```{r, include=FALSE} #Creating pre-test dataframes for the studies:

law1 <- read.table("SummaryResults_Asian_chr1.txt", header=TRUE,sep=",") law2 <- read.table("SummaryResults_Asian_chr2.txt", header=TRUE,sep=",") law3 <- read.table("SummaryResults_Asian_chr3.txt", header=TRUE,sep=",") law4 <- read.table("SummaryResults_Asian_chr4.txt", header=TRUE,sep=",") law5 <- read.table("SummaryResults_Asian_chr5.txt", header=TRUE,sep=",") law6 <- read.table("SummaryResults_Asian_chr6.txt", header=TRUE,sep=",") law7 <- read.table("SummaryResults_Asian_chr7.txt", header=TRUE,sep=",") law8 <- read.table("SummaryResults_Asian_chr8.txt", header=TRUE,sep=",") law9 <- read.table("SummaryResults_Asian_chr9.txt", header=TRUE,sep=",") law10 <- read.table("SummaryResults_Asian_chr10.txt", header=TRUE,sep=",") law11 <- read.table("SummaryResults_Asian_chr11.txt", header=TRUE,sep=",") law12 <- read.table("SummaryResults_Asian_chr12.txt", header=TRUE,sep=",") law13 <- read.table("SummaryResults_Asian_chr13.txt", header=TRUE,sep=",") law14 <- read.table("SummaryResults_Asian_chr14.txt", header=TRUE,sep=",") law15 <- read.table("SummaryResults_Asian_chr15.txt", header=TRUE,sep=",") law16 <- read.table("SummaryResults_Asian_chr16.txt", header=TRUE,sep=",") law17 <- read.table("SummaryResults_Asian_chr17.txt", header=TRUE,sep=",") law18 <- read.table("SummaryResults_Asian_chr18.txt",
```

```
header=TRUE,sep=“,”) law19 <- read.table(“SummaryResults_Asian_chr19.txt”, header=TRUE,sep=“,”) law20 <- read.table(“SummaryResults_Asian_chr20.txt”, header=TRUE,sep=“,”) law21 <- read.table(“SummaryResults_Asian_chr21.txt”, header=TRUE,sep=“,”) law22 <- read.table(“SummaryResults_Asian_chr22.txt”, header=TRUE,sep=“,”) law23 <- read.table(“SummaryResults_Asian_chr23.txt”, header=TRUE,sep=“,”) law <- rbind(law1,law2,law3,law4,law5,law6,law7,law8,law9,law10,law11,law12,
law13,law14,law15,law16,law17,law18,law19,law20,law21,law22,law23) lawchrpos<−paste(lawChromosome, law$Position, sep=“:”) law <- law %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>% filter(overall_pvalue > 0 & overall_pvalue < 1) law <- law[,c(1:12,41)]

        write.table(law, sep=“,file=”C:\Users\Mark\Documents\law_pretest.txt", row.names = FALSE, col.names = TRUE, quote = FALSE) rm(law) rm(list = c(‘law1’,‘law2’,‘law3’,‘law4’,‘law5’,‘law6’,‘law7’,‘law8’,‘law9’, ‘law10’,‘law11’,‘law12’,‘law13’,‘law14’,‘law15’,‘law16’,‘law17’, ‘law18’,‘law19’,‘law20’,‘law21’,‘law22’,‘law23’))

        phe1 <- read.table(“Summary_chr1.txt”, header=TRUE, sep=“) phe2 <- read.table(”Summary_chr2.txt“, header=TRUE, sep=”,“) phe3 <- read.table(”Summary_chr3.txt“, header=TRUE, sep=”,“) phe4 <- read.table(”Summary_chr4.txt“, header=TRUE, sep=”,“) phe5 <- read.table(”Summary_chr5.txt“, header=TRUE, sep=”,“) phe6 <- read.table(”Summary_chr6.txt“, header=TRUE, sep=”,“) phe7 <- read.table(”Summary_chr7.txt“, header=TRUE, sep=”) phe8 <- read.table(“Summary_chr8.txt”, header=TRUE, sep=“,”) phe9 <- read.table(“Summary_chr9.txt”, header=TRUE, sep=“,”) phe10 <- read.table(“Summary_chr10.txt”, header=TRUE, sep=“,”) phe11 <- read.table(“Summary_chr11.txt”, header=TRUE, sep=“,”) phe12 <- read.table(“Summary_chr12.txt”, header=TRUE, sep=“,”) phe13 <- read.table(“Summary_chr13.txt”, header=TRUE, sep=“,”) phe14 <- read.table(“Summary_chr14.txt”, header=TRUE, sep=“,”) phe15 <- read.table(“Summary_chr15.txt”, header=TRUE, sep=“,”) phe16 <- read.table(“Summary_chr16.txt”, header=TRUE, sep=“,”) phe17 <- read.table(“Summary_chr17.txt”, header=TRUE, sep=“,”) phe18 <- read.table(“Summary_chr18.txt”, header=TRUE, sep=“,”) phe19 <- read.table(“Summary_chr19.txt”, header=TRUE, sep=“,”) phe20 <- read.table(“Summary_chr20.txt”, header=TRUE, sep=“,”) phe21 <- read.table(“Summary_chr21.txt”, header=TRUE, sep=“,”) phe22 <- read.table(“Summary_chr22.txt”, header=TRUE, sep=“,”) phe23 <- read.table(“Summary_chr23.txt”, header=TRUE, sep=“,”)

        #Subsetting first because overall vector is too big: phe1 <- phe1[,c(1:12)] phe2 <- phe2[,c(1:12)] phe3 <- phe3[,c(1:12)] phe4 <- phe4[,c(1:12)] phe5 <- phe5[,c(1:12)] phe6 <- phe6[,c(1:12)] phe7 <- phe7[,c(1:12)] phe8 <- phe8[,c(1:12)] phe9 <- phe9[,c(1:12)] phe10 <- phe10[,c(1:12)] phe11 <- phe11[,c(1:12)] phe12 <- phe12[,c(1:12)] phe13 <- phe13[,c(1:12)] phe14 <- phe14[,c(1:12)] phe15 <- phe15[,c(1:12)] phe16 <- phe16[,c(1:12)] phe17 <- phe17[,c(1:12)] phe18 <- phe18[,c(1:12)] phe19 <- phe19[,c(1:12)] phe20 <- phe20[,c(1:12)] phe21 <- phe21[,c(1:12)] phe22 <- phe22[,c(1:12)] phe23 <- phe23[,c(1:12)]

        phe <- rbind(phe1,phe2,phe3,phe4,phe5,phe6,phe7,phe8,phe9,phe10,phe11,phe12, phe13,phe14,phe15,phe16,phe17,phe18,phe19,phe20,phe21,phe22,phe23) phechrpos<−paste(pheChromosome, phe$Position, sep=“:”) phe <- phe %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>% filter(overall_pvalue > 0 & overall_pvalue < 1)

        phe <- arrange(phe, Chromosome, Position)

        write.table(phe, sep=“,file=”C:\Users\Mark\Documents\phe_pretest.txt", row.names = FALSE, col.names = TRUE, quote = FALSE) rm(phe) rm(list = c(‘phe1’,‘phe2’,‘phe3’,‘phe4’,‘phe5’,‘phe6’,‘phe7’,‘phe8’,‘phe9’, ‘phe10’,‘phe11’,‘phe12’,‘phe13’,‘phe14’,‘phe15’,‘phe16’,‘phe17’, ‘phe18’,‘phe19’,‘phe20’,‘phe21’,‘phe22’,‘phe23’))

        #```

        #Create and format pre-test dataframes for OMARA & UKBB:

        omara <- omara %>% filter(effect_allele_frequency > 0.01 & effect_allele_frequency < 0.99) %>% filter(standard_error > 0) %>%
                filter(p_value > 0 & p_value < 1)
        omara <- arrange(omara, chromosome, base_pair_location)

        #write.table(omara, sep="\t", file="C:\\Users\\Mark\\Documents\\omara_pretest.txt", row.names=FALSE, col.names = TRUE, quote = FALSE)

        ukbb <- arrange(ukbb, chr, pos)

        #write.table(ukbb, sep="\t", file="C:\\Users\\Mark\\Documents\\ukbb_pretest.txt", row.names = FALSE, col.names = TRUE, quote = FALSE)
        #Chr Template:
        library(tidyverse)
```

```
library(data.table)
omara_1 <- omara %>% filter(chromosome == 1)
phelan_1 <- read.table("Summary_chr1.txt", header=TRUE, sep="\t")
lawrenson_1 <- read.table("SummaryResults_Asian_chr1.txt", header=TRUE, sep=",")
ukbb_1 <- filter(ukbb, grepl("^1:", chrpos))
#Clean SNP naming convention - chr:pos

phelan_1$chrpos <- paste(phelan_1$Chromosome, phelan_1$Position, sep=":")
lawrenson_1$chrpos <- paste(lawrenson_1$Chromosome, lawrenson_1$Position, sep=":")


#Prune NA entries

omara_1 <- filter(omara_1, !grepl("NA", chrpos))
phelan_1 <- filter(phelan_1, !grepl("NA", chrpos))
lawrenson_1 <- filter(lawrenson_1, !grepl("NA", chrpos))
ukbb_1 <- filter(ukbb_1, !grepl("NA", chrpos))

phelan_1 <- phelan_1 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_1 <- lawrenson_1 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
#Organize files for analysis:

omara_1 <- omara_1[,c(26,17,18,19,20,21)]
#^ SNP,  other allele & effect allele, beta, SE, p-val.

lawrenson_1 <- lawrenson_1[,c(41,6,5,9,10,12)]
#^ SNP, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_1 <- phelan_1[,c(73,6,5,9,10,12)]
#^ SNP, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_1 <- ukbb_1[,c(16,17,18,9,10,12)]
#^ SNP, baseline allele, effect allele, beta, SE, overall p-val.
#Export files for METAL format:

write.table(omara_1, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr1\\omara_1.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_1, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr1\\lawrenson_1.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_1, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr1\\phelan_1.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_1, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr1\\ukbb_1.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_1','lawrenson_1','phelan_1','ukbb_1'))
#2:
omara_2 <- omara %>% filter(chromosome == 2)
phelan_2 <- read.table("Summary_chr2.txt", header=TRUE, sep=",")
lawrenson_2 <- read.table("SummaryResults_Asian_chr2.txt", header=TRUE, sep=",")
ukbb_2 <- filter(ukbb, grepl("^2:", chrpos))
phelan_2$chrpos <- paste(phelan_2$Chromosome, phelan_2$Position, sep=":")
lawrenson_2$chrpos <- paste(lawrenson_2$Chromosome, lawrenson_2$Position, sep=":")

omara_2 <- filter(omara_2, !grepl("NA", chrpos))
phelan_2 <- filter(phelan_2, !grepl("NA", chrpos))
lawrenson_2 <- filter(lawrenson_2, !grepl("NA", chrpos))
ukbb_2 <- filter(ukbb_2, !grepl("NA", chrpos))
```

```
phelan_2 <- phelan_2 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_2 <- lawrenson_2 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_2 <- omara_2[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_2 <- lawrenson_2[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_2 <- phelan_2[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_2 <- ukbb_2[,c(16,17,18,9,10,12)]
#^ SNP, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_2, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr2\\omara_2.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_2, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr2\\lawrenson_2.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_2, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr2\\phelan_2.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_2, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr2\\ukbb_2.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_2','lawrenson_2','phelan_2','ukbb_2'))
#3:
omara_3 <- omara %>% filter(chromosome == 3)
phelan_3 <- read.table("Summary_chr3.txt", header=TRUE, sep=",")
lawrenson_3 <- read.table("SummaryResults_Asian_chr3.txt", header=TRUE, sep=",")
ukbb_3 <- filter(ukbb, grepl("^3:", chrpos))
phelan_3$chrpos <- paste(phelan_3$Chromosome, phelan_3$Position, sep=":")
lawrenson_3$chrpos <- paste(lawrenson_3$Chromosome, lawrenson_3$Position, sep=":")

omara_3 <- filter(omara_3, !grepl("NA", chrpos))
phelan_3 <- filter(phelan_3, !grepl("NA", chrpos))
lawrenson_3 <- filter(lawrenson_3, !grepl("NA", chrpos))
ukbb_3 <- filter(ukbb_3, !grepl("NA", chrpos))

phelan_3 <- phelan_3 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_3 <- lawrenson_3 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_3 <- omara_3[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_3 <- lawrenson_3[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_3 <- phelan_3[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_3 <- ukbb_3[,c(16,17,18,9,10,12)]
#^ SNP, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_3, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr3\\omara_3.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_3, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr3\\lawrenson_3.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)
```

```
write.table(phelan_3, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr3\\phelan_3.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_3, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr3\\ukbb_3.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_3','lawrenson_3','phelan_3','ukbb_3'))
#4:
omara_4 <- omara %>% filter(chromosome == 4)
phelan_4 <- read.table("Summary_chr4.txt", header=TRUE, sep=",")
lawrenson_4 <- read.table("SummaryResults_Asian_chr4.txt", header=TRUE, sep=",")
ukbb_4 <- filter(ukbb, grepl("^4:", chrpos))
phelan_4$chrpos <- paste(phelan_4$Chromosome, phelan_4$Position, sep=":")
lawrenson_4$chrpos <- paste(lawrenson_4$Chromosome, lawrenson_4$Position, sep=":")

omara_4 <- filter(omara_4, !grepl("NA", chrpos))
phelan_4 <- filter(phelan_4, !grepl("NA", chrpos))
lawrenson_4 <- filter(lawrenson_4, !grepl("NA", chrpos))
ukbb_4 <- filter(ukbb_4, !grepl("NA", chrpos))

phelan_4 <- phelan_4 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_4 <- lawrenson_4 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_4 <- omara_4[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_4 <- lawrenson_4[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_4 <- phelan_4[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_4 <- ukbb_4[,c(16,17,18,9,10,12)]
#^ SNP, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_4, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr4\\omara_4.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_4, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr4\\lawrenson_4.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_4, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr4\\phelan_4.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_4, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr4\\ukbb_4.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_4','lawrenson_4','phelan_4','ukbb_4'))
#5:
omara_5 <- omara %>% filter(chromosome == 5)
phelan_5 <- read.table("Summary_chr5.txt", header=TRUE, sep=",")
lawrenson_5 <- read.table("SummaryResults_Asian_chr5.txt", header=TRUE, sep=",")
ukbb_5 <- filter(ukbb, grepl("^5:", chrpos))
phelan_5$chrpos <- paste(phelan_5$Chromosome, phelan_5$Position, sep=":")
lawrenson_5$chrpos <- paste(lawrenson_5$Chromosome, lawrenson_5$Position, sep=":")

omara_5 <- filter(omara_5, !grepl("NA", chrpos))
phelan_5 <- filter(phelan_5, !grepl("NA", chrpos))
lawrenson_5 <- filter(lawrenson_5, !grepl("NA", chrpos))
ukbb_5 <- filter(ukbb_5, !grepl("NA", chrpos))
```

```
phelan_5 <- phelan_5 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_5 <- lawrenson_5 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_5 <- omara_5[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_5 <- lawrenson_5[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_5 <- phelan_5[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_5 <- ukbb_5[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_5, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr5\\omara_5.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_5, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr5\\lawrenson_5.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_5, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr5\\phelan_5.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_5, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr5\\ukbb_5.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_5','lawrenson_5','phelan_5','ukbb_5'))
#6:
omara_6 <- omara %>% filter(chromosome == 6)
phelan_6 <- read.table("Summary_chr6.txt", header=TRUE, sep=",")
lawrenson_6 <- read.table("SummaryResults_Asian_chr6.txt", header=TRUE, sep=",")
ukbb_6 <- filter(ukbb, grepl("^6:", chrpos))
phelan_6$chrpos <- paste(phelan_6$Chromosome, phelan_6$Position, sep=":")
lawrenson_6$chrpos <- paste(lawrenson_6$Chromosome, lawrenson_6$Position, sep=":")

omara_6 <- filter(omara_6, !grepl("NA", chrpos))
phelan_6 <- filter(phelan_6, !grepl("NA", chrpos))
lawrenson_6 <- filter(lawrenson_6, !grepl("NA", chrpos))
ukbb_6 <- filter(ukbb_6, !grepl("NA", chrpos))

phelan_6 <- phelan_6 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_6 <- lawrenson_6 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_6 <- omara_6[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_6 <- lawrenson_6[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_6 <- phelan_6[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_6 <- ukbb_6[,c(16,17,18,9,10,12)]
#^ SNP, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_6, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr6\\omara_6.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_6, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr6\\lawrenson_6.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)
```

```
write.table(phelan_6, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr6\\phelan_6.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_6, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr6\\ukbb_6.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_6','lawrenson_6','phelan_6','ukbb_6'))
#7:
omara_7 <- omara %>% filter(chromosome == 7)
phelan_7 <- read.table("Summary_chr7.txt", header=TRUE, sep="\t")
lawrenson_7 <- read.table("SummaryResults_Asian_chr7.txt", header=TRUE, sep=",")
ukbb_7 <- filter(ukbb, grepl("^7:", chrpos))
phelan_7$chrpos <- paste(phelan_7$Chromosome, phelan_7$Position, sep=":")
lawrenson_7$chrpos <- paste(lawrenson_7$Chromosome, lawrenson_7$Position, sep=":")

omara_7 <- filter(omara_7, !grepl("NA", chrpos))
phelan_7 <- filter(phelan_7, !grepl("NA", chrpos))
lawrenson_7 <- filter(lawrenson_7, !grepl("NA", chrpos))
ukbb_7 <- filter(ukbb_7, !grepl("NA", chrpos))

phelan_7 <- phelan_7 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_7 <- lawrenson_7 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_7 <- omara_7[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_7 <- lawrenson_7[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_7 <- phelan_7[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_7 <- ukbb_7[,c(16,17,18,9,10,12)]
#^ SNP, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_7, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr7\\omara_7.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_7, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr7\\lawrenson_7.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_7, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr7\\phelan_7.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_7, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr7\\ukbb_7.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_7','lawrenson_7','phelan_7','ukbb_7'))
#8:
omara_8 <- omara %>% filter(chromosome == 8)
phelan_8 <- read.table("Summary_chr8.txt", header=TRUE, sep=",")
lawrenson_8 <- read.table("SummaryResults_Asian_chr8.txt", header=TRUE, sep=",")
ukbb_8 <- filter(ukbb, grepl("^8:", chrpos))
phelan_8$chrpos <- paste(phelan_8$Chromosome, phelan_8$Position, sep=":")
lawrenson_8$chrpos <- paste(lawrenson_8$Chromosome, lawrenson_8$Position, sep=":")

omara_8 <- filter(omara_8, !grepl("NA", chrpos))
phelan_8 <- filter(phelan_8, !grepl("NA", chrpos))
lawrenson_8 <- filter(lawrenson_8, !grepl("NA", chrpos))
ukbb_8 <- filter(ukbb_8, !grepl("NA", chrpos))

phelan_8 <- phelan_8 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
```

```
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_8 <- lawrenson_8 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_8 <- omara_8[,c(26,17,18,19,20,21)]
#^ SNP, rs, harmonised other allele & effect allele, beta, SE, p-val.

lawrenson_8 <- lawrenson_8[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_8 <- phelan_8[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_8 <- ukbb_8[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_8, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr8\\omara_8.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_8, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr8\\lawrenson_8.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_8, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr8\\phelan_8.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_8, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr8\\ukbb_8.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_8','lawrenson_8','phelan_8','ukbb_8'))
#9:
omara_9 <- omara %>% filter(chromosome == 9)
phelan_9 <- read.table("Summary_chr9.txt", header=TRUE, sep=",")
lawrenson_9 <- read.table("SummaryResults_Asian_chr9.txt", header=TRUE, sep=",")
ukbb_9 <- filter(ukbb, grepl("^9:", chrpos))
phelan_9$chrpos <- paste(phelan_9$Chromosome, phelan_9$Position, sep=":")
lawrenson_9$chrpos <- paste(lawrenson_9$Chromosome, lawrenson_9$Position, sep=":")

omara_9 <- filter(omara_9, !grepl("NA", chrpos))
phelan_9 <- filter(phelan_9, !grepl("NA", chrpos))
lawrenson_9 <- filter(lawrenson_9, !grepl("NA", chrpos))
ukbb_9 <- filter(ukbb_9, !grepl("NA", chrpos))

phelan_9 <- phelan_9 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_9 <- lawrenson_9 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_9 <- omara_9[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_9 <- lawrenson_9[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_9 <- phelan_9[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_9 <- ukbb_9[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_9, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr9\\omara_9.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_9, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr9\\lawrenson_9.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)
```

```
write.table(phelan_9, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr9\\phelan_9.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_9, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr9\\ukbb_9.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_9','lawrenson_9','phelan_9','ukbb_9'))
#10:
omara_10 <- omara %>% filter(chromosome == 10)
phelan_10 <- read.table("Summary_chr10.txt", header=TRUE, sep=",")
lawrenson_10 <- read.table("SummaryResults_Asian_chr10.txt", header=TRUE, sep=",")
ukbb_10 <- filter(ukbb, grepl("^10:", chrpos))
phelan_10$chrpos <- paste(phelan_10$Chromosome, phelan_10$Position, sep=":")
lawrenson_10$chrpos <- paste(lawrenson_10$Chromosome, lawrenson_10$Position, sep=":")

omara_10 <- filter(omara_10, !grepl("NA", chrpos))
phelan_10 <- filter(phelan_10, !grepl("NA", chrpos))
lawrenson_10 <- filter(lawrenson_10, !grepl("NA", chrpos))
ukbb_10 <- filter(ukbb_10, !grepl("NA", chrpos))

phelan_10 <- phelan_10 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_10 <- lawrenson_10 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_10 <- omara_10[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_10 <- lawrenson_10[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_10 <- phelan_10[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_10 <- ukbb_10[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_10, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr10\\omara_10.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_10, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr10\\lawrenson_10.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_10, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr10\\phelan_10.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_10, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr10\\ukbb_10.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_10','lawrenson_10','phelan_10','ukbb_10'))
#11:
omara_11 <- omara %>% filter(chromosome == 11)
phelan_11 <- read.table("Summary_chr11.txt", header=TRUE, sep=",")
lawrenson_11 <- read.table("SummaryResults_Asian_chr11.txt", header=TRUE, sep=",")
ukbb_11 <- filter(ukbb, grepl("^11:", chrpos))
phelan_11$chrpos <- paste(phelan_11$Chromosome, phelan_11$Position, sep=":")
lawrenson_11$chrpos <- paste(lawrenson_11$Chromosome, lawrenson_11$Position, sep=":")

omara_11 <- filter(omara_11, !grepl("NA", chrpos))
phelan_11 <- filter(phelan_11, !grepl("NA", chrpos))
lawrenson_11 <- filter(lawrenson_11, !grepl("NA", chrpos))
ukbb_11 <- filter(ukbb_11, !grepl("NA", chrpos))

phelan_11 <- phelan_11 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
```

```
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_11 <- lawrenson_11 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_11 <- omara_11[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_11 <- lawrenson_11[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_11 <- phelan_11[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_11 <- ukbb_11[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_11, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr11\\omara_11.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_11, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr11\\lawrenson_11.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_11, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr11\\phelan_11.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_11, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr11\\ukbb_11.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_11','lawrenson_11','phelan_11','ukbb_11'))
#12:
omara_12 <- omara %>% filter(chromosome == 12)
phelan_12 <- read.table("Summary_chr12.txt", header=TRUE, sep=",")
lawrenson_12 <- read.table("SummaryResults_Asian_chr12.txt", header=TRUE, sep=",")
ukbb_12 <- filter(ukbb, grepl("^12:", chrpos))
phelan_12$chrpos <- paste(phelan_12$Chromosome, phelan_12$Position, sep=":")
lawrenson_12$chrpos <- paste(lawrenson_12$Chromosome, lawrenson_12$Position, sep=":")

omara_12 <- filter(omara_12, !grepl("NA", chrpos))
phelan_12 <- filter(phelan_12, !grepl("NA", chrpos))
lawrenson_12 <- filter(lawrenson_12, !grepl("NA", chrpos))
ukbb_12 <- filter(ukbb_12, !grepl("NA", chrpos))

phelan_12 <- phelan_12 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_12 <- lawrenson_12 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_12 <- omara_12[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_12 <- lawrenson_12[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_12 <- phelan_12[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_12 <- ukbb_12[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_12, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr12\\omara_12.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_12, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr12\\lawrenson_12.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)
```

```
write.table(phelan_12, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr12\\phelan_12.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_12, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr12\\ukbb_12.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_12','lawrenson_12','phelan_12','ukbb_12'))
#13:
omara_13 <- omara %>% filter(chromosome == 13)
phelan_13 <- read.table("Summary_chr13.txt", header=TRUE, sep=",")
lawrenson_13 <- read.table("SummaryResults_Asian_chr13.txt", header=TRUE, sep=",")
ukbb_13 <- filter(ukbb, grepl("^13:", chrpos))
phelan_13$chrpos <- paste(phelan_13$Chromosome, phelan_13$Position, sep=":")
lawrenson_13$chrpos <- paste(lawrenson_13$Chromosome, lawrenson_13$Position, sep=":")

omara_13 <- filter(omara_13, !grepl("NA", chrpos))
phelan_13 <- filter(phelan_13, !grepl("NA", chrpos))
lawrenson_13 <- filter(lawrenson_13, !grepl("NA", chrpos))
ukbb_13 <- filter(ukbb_13, !grepl("NA", chrpos))

phelan_13 <- phelan_13 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_13 <- lawrenson_13 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_13 <- omara_13[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_13 <- lawrenson_13[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_13 <- phelan_13[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_13 <- ukbb_13[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_13, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr13\\omara_13.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_13, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr13\\lawrenson_13.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_13, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr13\\phelan_13.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_13, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr13\\ukbb_13.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_13','lawrenson_13','phelan_13','ukbb_13'))
#14:
omara_14 <- omara %>% filter(chromosome == 14)
phelan_14 <- read.table("Summary_chr14.txt", header=TRUE, sep=",")
lawrenson_14 <- read.table("SummaryResults_Asian_chr14.txt", header=TRUE, sep=",")
ukbb_14 <- filter(ukbb, grepl("^14:", chrpos))
phelan_14$chrpos <- paste(phelan_14$Chromosome, phelan_14$Position, sep=":")
lawrenson_14$chrpos <- paste(lawrenson_14$Chromosome, lawrenson_14$Position, sep=":")

omara_14 <- filter(omara_14, !grepl("NA", chrpos))
phelan_14 <- filter(phelan_14, !grepl("NA", chrpos))
lawrenson_14 <- filter(lawrenson_14, !grepl("NA", chrpos))
ukbb_14 <- filter(ukbb_14, !grepl("NA", chrpos))

phelan_14 <- phelan_14 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
```

```
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_14 <- lawrenson_14 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_14 <- omara_14[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_14 <- lawrenson_14[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_14 <- phelan_14[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_14 <- ukbb_14[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_14, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr14\\omara_14.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_14, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr14\\lawrenson_14.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_14, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr14\\phelan_14.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_14, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr14\\ukbb_14.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_14','lawrenson_14','phelan_14','ukbb_14'))
#15:
omara_15 <- omara %>% filter(chromosome == 15)
phelan_15 <- read.table("Summary_chr15.txt", header=TRUE, sep=",")
lawrenson_15 <- read.table("SummaryResults_Asian_chr15.txt", header=TRUE, sep=",")
ukbb_15 <- filter(ukbb, grepl("^15:", chrpos))
phelan_15$chrpos <- paste(phelan_15$Chromosome, phelan_15$Position, sep=":")
lawrenson_15$chrpos <- paste(lawrenson_15$Chromosome, lawrenson_15$Position, sep=":")

omara_15 <- filter(omara_15, !grepl("NA", chrpos))
phelan_15 <- filter(phelan_15, !grepl("NA", chrpos))
lawrenson_15 <- filter(lawrenson_15, !grepl("NA", chrpos))
ukbb_15 <- filter(ukbb_15, !grepl("NA", chrpos))

phelan_15 <- phelan_15 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_15 <- lawrenson_15 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_15 <- omara_15[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_15 <- lawrenson_15[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_15 <- phelan_15[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_15 <- ukbb_15[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_15, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr15\\omara_15.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_15, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr15\\lawrenson_15.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)
```

```
write.table(phelan_15, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr15\\phelan_15.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_15, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr15\\ukbb_15.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_15','lawrenson_15','phelan_15','ukbb_15'))
#16:
omara_16 <- omara %>% filter(chromosome == 16)
phelan_16 <- read.table("Summary_chr16.txt", header=TRUE, sep=",")
lawrenson_16 <- read.table("SummaryResults_Asian_chr16.txt", header=TRUE, sep=",")
ukbb_16 <- filter(ukbb, grepl("^16:", chrpos))
phelan_16$chrpos <- paste(phelan_16$Chromosome, phelan_16$Position, sep=":")
lawrenson_16$chrpos <- paste(lawrenson_16$Chromosome, lawrenson_16$Position, sep=":")

omara_16 <- filter(omara_16, !grepl("NA", chrpos))
phelan_16 <- filter(phelan_16, !grepl("NA", chrpos))
lawrenson_16 <- filter(lawrenson_16, !grepl("NA", chrpos))
ukbb_16 <- filter(ukbb_16, !grepl("NA", chrpos))

phelan_16 <- phelan_16 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_16 <- lawrenson_16 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_16 <- omara_16[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_16 <- lawrenson_16[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_16 <- phelan_16[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_16 <- ukbb_16[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_16, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr16\\omara_16.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_16, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr16\\lawrenson_16.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_16, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr16\\phelan_16.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_16, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr16\\ukbb_16.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_16','lawrenson_16','phelan_16','ukbb_16'))
#17:
omara_17 <- omara %>% filter(chromosome == 17)
phelan_17 <- read.table("Summary_chr17.txt", header=TRUE, sep=",")
lawrenson_17 <- read.table("SummaryResults_Asian_chr17.txt", header=TRUE, sep=",")
ukbb_17 <- filter(ukbb, grepl("^17:", chrpos))
phelan_17$chrpos <- paste(phelan_17$Chromosome, phelan_17$Position, sep=":")
lawrenson_17$chrpos <- paste(lawrenson_17$Chromosome, lawrenson_17$Position, sep=":")

omara_17 <- filter(omara_17, !grepl("NA", chrpos))
phelan_17 <- filter(phelan_17, !grepl("NA", chrpos))
lawrenson_17 <- filter(lawrenson_17, !grepl("NA", chrpos))
ukbb_17 <- filter(ukbb_17, !grepl("NA", chrpos))

phelan_17 <- phelan_17 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
```

```
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_17 <- lawrenson_17 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_17 <- omara_17[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_17 <- lawrenson_17[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_17 <- phelan_17[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_17 <- ukbb_17[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_17, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr17\\omara_17.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_17, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr17\\lawrenson_17.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_17, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr17\\phelan_17.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_17, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr17\\ukbb_17.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_17','lawrenson_17','phelan_17','ukbb_17'))
#18:
omara_18 <- omara %>% filter(chromosome == 18)
phelan_18 <- read.table("Summary_chr18.txt", header=TRUE, sep=",")
lawrenson_18 <- read.table("SummaryResults_Asian_chr18.txt", header=TRUE, sep=",")
ukbb_18 <- filter(ukbb, grepl("^18:", chrpos))
phelan_18$chrpos <- paste(phelan_18$Chromosome, phelan_18$Position, sep=":")
lawrenson_18$chrpos <- paste(lawrenson_18$Chromosome, lawrenson_18$Position, sep=":")

omara_18 <- filter(omara_18, !grepl("NA", chrpos))
phelan_18 <- filter(phelan_18, !grepl("NA", chrpos))
lawrenson_18 <- filter(lawrenson_18, !grepl("NA", chrpos))
ukbb_18 <- filter(ukbb_18, !grepl("NA", chrpos))

phelan_18 <- phelan_18 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_18 <- lawrenson_18 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_18 <- omara_18[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_18 <- lawrenson_18[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_18 <- phelan_18[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_18 <- ukbb_18[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_18, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr18\\omara_18.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_18, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr18\\lawrenson_18.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)
```

```
write.table(phelan_18, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr18\\phelan_18.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_18, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr18\\ukbb_18.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_18','lawrenson_18','phelan_18','ukbb_18'))
#19:
omara_19 <- omara %>% filter(chromosome == 19)
phelan_19 <- read.table("Summary_chr19.txt", header=TRUE, sep=",")
lawrenson_19 <- read.table("SummaryResults_Asian_chr19.txt", header=TRUE, sep=",")
ukbb_19 <- filter(ukbb, grepl("^19:", chrpos))
phelan_19$chrpos <- paste(phelan_19$Chromosome, phelan_19$Position, sep=":")
lawrenson_19$chrpos <- paste(lawrenson_19$Chromosome, lawrenson_19$Position, sep=":")

omara_19 <- filter(omara_19, !grepl("NA", chrpos))
phelan_19 <- filter(phelan_19, !grepl("NA", chrpos))
lawrenson_19 <- filter(lawrenson_19, !grepl("NA", chrpos))
ukbb_19 <- filter(ukbb_19, !grepl("NA", chrpos))

phelan_19 <- phelan_19 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_19 <- lawrenson_19 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_19 <- omara_19[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_19 <- lawrenson_19[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_19 <- phelan_19[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_19 <- ukbb_19[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_19, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr19\\omara_19.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_19, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr19\\lawrenson_19.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_19, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr19\\phelan_19.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_19, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr19\\ukbb_19.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_19','lawrenson_19','phelan_19','ukbb_19'))
#20:
omara_20 <- omara %>% filter(chromosome == 20)
phelan_20 <- read.table("Summary_chr20.txt", header=TRUE, sep=",")
lawrenson_20 <- read.table("SummaryResults_Asian_chr20.txt", header=TRUE, sep=",")
ukbb_20 <- filter(ukbb, grepl("^20:", chrpos))
phelan_20$chrpos <- paste(phelan_20$Chromosome, phelan_20$Position, sep=":")
lawrenson_20$chrpos <- paste(lawrenson_20$Chromosome, lawrenson_20$Position, sep=":")

omara_20 <- filter(omara_20, !grepl("NA", chrpos))
phelan_20 <- filter(phelan_20, !grepl("NA", chrpos))
lawrenson_20 <- filter(lawrenson_20, !grepl("NA", chrpos))
ukbb_20 <- filter(ukbb_20, !grepl("NA", chrpos))

phelan_20 <- phelan_20 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
```

```
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_20 <- lawrenson_20 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_20 <- omara_20[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_20 <- lawrenson_20[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_20 <- phelan_20[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_20 <- ukbb_20[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_20, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr20\\omara_20.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_20, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr20\\lawrenson_20.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_20, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr20\\phelan_20.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_20, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr20\\ukbb_20.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_20','lawrenson_20','phelan_20','ukbb_20'))
#21:
omara_21 <- omara %>% filter(chromosome == 21)
phelan_21 <- read.table("Summary_chr21.txt", header=TRUE, sep=",")
lawrenson_21 <- read.table("SummaryResults_Asian_chr21.txt", header=TRUE, sep=",")
ukbb_21 <- filter(ukbb, grepl("^21:", chrpos))
phelan_21$chrpos <- paste(phelan_21$Chromosome, phelan_21$Position, sep=":")
lawrenson_21$chrpos <- paste(lawrenson_21$Chromosome, lawrenson_21$Position, sep=":")

omara_21 <- filter(omara_21, !grepl("NA", chrpos))
phelan_21 <- filter(phelan_21, !grepl("NA", chrpos))
lawrenson_21 <- filter(lawrenson_21, !grepl("NA", chrpos))
ukbb_21 <- filter(ukbb_21, !grepl("NA", chrpos))

phelan_21 <- phelan_21 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_21 <- lawrenson_21 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_21 <- omara_21[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_21 <- lawrenson_21[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_21 <- phelan_21[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_21 <- ukbb_21[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_21, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr21\\omara_21.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_21, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr21\\lawrenson_21.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)
```

```
write.table(phelan_21, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr21\\phelan_21.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_21, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr21\\ukbb_21.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_21','lawrenson_21','phelan_21','ukbb_21'))
#22:
omara_22 <- omara %>% filter(chromosome == 22)
phelan_22 <- read.table("Summary_chr22.txt", header=TRUE, sep=",")
lawrenson_22 <- read.table("SummaryResults_Asian_chr22.txt", header=TRUE, sep=",")
ukbb_22 <- filter(ukbb, grepl("^22:", chrpos))
phelan_22$chrpos <- paste(phelan_22$Chromosome, phelan_22$Position, sep=":")
lawrenson_22$chrpos <- paste(lawrenson_22$Chromosome, lawrenson_22$Position, sep=":")

omara_22 <- filter(omara_22, !grepl("NA", chrpos))
phelan_22 <- filter(phelan_22, !grepl("NA", chrpos))
lawrenson_22 <- filter(lawrenson_22, !grepl("NA", chrpos))
ukbb_22 <- filter(ukbb_22, !grepl("NA", chrpos))

phelan_22 <- phelan_22 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_22 <- lawrenson_22 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
omara_22 <- omara_22[,c(26,17,18,19,20,21)]
#^ SNP, other allele & effect allele, beta, SE, p-val.

lawrenson_22 <- lawrenson_22[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_22 <- phelan_22[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

ukbb_22 <- ukbb_22[,c(16,17,18,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, beta, SE, overall p-val.
write.table(omara_22, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr22\\omara_22.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(lawrenson_22, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr22\\lawrenson_22.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_22, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr22\\phelan_22.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(ukbb_22, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr22\\ukbb_22.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(list = c('omara_22','lawrenson_22','phelan_22','ukbb_22'))
#23:

#omara_23 <- omara %>% filter(chromosome == 23) #NO data
phelan_23 <- read.table("Summary_chr23.txt", header=TRUE, sep=",")
lawrenson_23 <- read.table("SummaryResults_Asian_chr23.txt", header=TRUE, sep=",")
#ukbb_23 <- filter(ukbb, grepl("^23:", chrpos)) #NO data
phelan_23$chrpos <- paste(phelan_23$Chromosome, phelan_23$Position, sep=":")
lawrenson_23$chrpos <- paste(lawrenson_23$Chromosome, lawrenson_23$Position, sep=":")

phelan_23 <- filter(phelan_23, !grepl("NA", chrpos))
lawrenson_23 <- filter(lawrenson_23, !grepl("NA", chrpos))

phelan_23 <- phelan_23 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
```

```
lawrenson_23 <- lawrenson_23 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
        filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_23 <- lawrenson_23[,c(41,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.

phelan_23 <- phelan_23[,c(73,1,6,5,9,10,12)]
#^ SNP, rs, baseline allele, effect allele, overall OR, overall SE, overall p-val.
write.table(lawrenson_23, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr23\\lawrenson_23.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

write.table(phelan_23, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr23\\phelan_23.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

rm(lawrenson_23)
rm(phelan_23)
rm(omara)
rm(ukbb)
-Results Compiling:

Ovarian + Endometrial META ####

METAr1 <- read.table("META_Chr1_1.txt", header=TRUE, sep="\t")
METAr1$chromosome <- 1

METAr2 <- read.table("META_Chr2_1.txt", header=TRUE, sep="\t")
METAr2$chromosome <- 2

METAr3 <- read.table("META_Chr3_1.txt", header=TRUE, sep="\t")
METAr3$chromosome <- 3

METAr4 <- read.table("META_Chr4_1.txt", header=TRUE, sep="\t")
METAr4$chromosome <- 4

METAr5 <- read.table("META_Chr5_1.txt", header=TRUE, sep="\t")
METAr5$chromosome <- 5

METAr6 <- read.table("META_Chr6_1.txt", header=TRUE, sep="\t")
METAr6$chromosome <- 6

METAr7 <- read.table("META_Chr7_1.txt", header=TRUE, sep="\t")
METAr7$chromosome <- 7

METAr8 <- read.table("META_Chr8_1.txt", header=TRUE, sep="\t")
METAr8$chromosome <- 8

METAr9 <- read.table("META_Chr9_1.txt", header=TRUE, sep="\t")
METAr9$chromosome <- 9

METAr10 <- read.table("META_Chr10_1.txt", header=TRUE, sep="\t")
METAr10$chromosome <- 10

METAr11 <- read.table("META_Chr11_1.txt", header=TRUE, sep="\t")
METAr11$chromosome <- 11

METAr12 <- read.table("META_Chr12_1.txt", header=TRUE, sep="\t")
METAr12$chromosome <- 12

METAr13 <- read.table("META_Chr13_1.txt", header=TRUE, sep="\t")
METAr13$chromosome <- 13

METAr14 <- read.table("META_Chr14_1.txt", header=TRUE, sep="\t")
METAr14$chromosome <- 14
```

```
METAr15 <- read.table("META_Chr15_1.txt", header=TRUE, sep="\t")
METAr15$chromosome <- 15

METAr16 <- read.table("META_Chr16_1.txt", header=TRUE, sep="\t")
METAr16$chromosome <- 16

METAr17 <- read.table("META_Chr17_1.txt", header=TRUE, sep="\t")
METAr17$chromosome <- 17

METAr18 <- read.table("META_Chr18_1.txt", header=TRUE, sep="\t")
METAr18$chromosome <- 18

METAr19 <- read.table("META_Chr19_1.txt", header=TRUE, sep="\t")
METAr19$chromosome <- 19

METAr20 <- read.table("META_Chr20_1.txt", header=TRUE, sep="\t")
METAr20$chromosome <- 20

METAr21 <- read.table("META_Chr21_1.txt", header=TRUE, sep="\t")
METAr21$chromosome <- 21

METAr22 <- read.table("META_Chr22_1.txt", header=TRUE, sep="\t")
METAr22$chromosome <- 22
METAr <- rbind(METAr1, METAr2, METAr3, METAr4, METAr5, METAr6, METAr7, METAr8, METAr9,
METAr10, METAr11, METAr12, METAr13,
        METAr14, METAr15, METAr16, METAr17, METAr18, METAr19, METAr20, METAr21, METAr22)
rm(list = c('METAr1','METAr2','METAr3','METAr4','METAr5','METAr6',
        'METAr7','METAr8','METAr9','METAr10','METAr11','METAr12',
        'METAr13','METAr14','METAr15','METAr16','METAr17','METAr18',
        'METAr19','METAr20','METAr21','METAr22'))
METAr$Position <- gsub("^.*:", "", METAr$MarkerName)

METAr$chromosome <- as.numeric(METAr$chromosome)
METAr$Position <- as.numeric(METAr$Position)
METAr <- arrange(METAr, chromosome, Position)

METAr$chrposal <- paste(METAr$MarkerName, METAr$Allele2, sep="_")
METAr$chrposal <- paste(METAr$chrposal, METAr$Allele1, sep="/" )

#Assessing number of missing underlying studies by marker:
#Cut down SNPs to just those in at least 2 underlying studies (NO ovarian only):

METAr$num_miss <- str_count(as.character(METAr$Direction), "\\?")
METAr$oma_miss <- substr(METAr$Direction, 1, 1)
METAr$rowID <- row.names(METAr)
#Below line finds only ovarian estimates, that is, only non-ovarian SNP was missing:
METAr_filtered <- METAr %>% filter(oma_miss == '?' & num_miss == 1)
METAr <- METAr[!(METAr$rowID %in% METAr_filtered$rowID),]
METAr <- METAr %>% filter(num_miss < 2) #can't be estimates from only one study

#write.table(METAr, sep="\t", file="C:\\Users\\Mark\\Documents\\META_oe_filtered.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)


#rm(METAr)
METAr_sig <- filter(METAr, P.value <= 0.00000005)
METAr_sig <- METAr_sig[,c(12:13,1:11)]


omara_liftover <- read.table("hglft_results.txt", sep="\t")
```

```
omara_liftover <- omara_liftover %>% rename(base_pair_location=V3, variant_id=V4)

omara <- as.data.frame(fread("harmonised.qc.tsv"))

omara_merge <- merge(omara, omara_liftover[,c(3,4)], by="variant_id")

omara <- omara_merge
rm(omara_merge)
rm(omara_liftover)

omara$chrpos <- paste(omara$chromosome, omara$base_pair_location.y, sep=":")
omara <- omara %>% rename(base_pair_location=base_pair_location.y)
library(metaviz)
#SNP 9:16889937

phelan_9 <- read.table("Summary_chr9.txt", header=TRUE, sep=",")
phelan_9$chrpos <- paste(phelan_9$Chromosome, phelan_9$Position, sep=":")
lawrenson_9 <- read.table("SummaryResults_Asian_chr9.txt", header=TRUE, sep=",")
lawrenson_9$chrpos <- paste(lawrenson_9$Chromosome, lawrenson_9$Position, sep=":")

phe9_fp1 <- filter(phelan_9, grepl("9:16889937", chrpos))
law9_fp1 <- filter(lawrenson_9, grepl("9:16889937", chrpos))
oma9_fp1 <- filter(omara, grepl("9:16889937", chrpos))

oma9_fp1 <- oma9_fp1[,c(26,15,25,17:18,16,19:21)]
oma9_fp1$study <- "O'Mara (2018)"
oma9_fp1$N.cases <- 12906
oma9_fp1$N.controls <- 108979
phe9_fp1 <- phe9_fp1[,c(73,3:7,9:10,12)]
phe9_fp1$study <- "Phelan (2017)"
phe9_fp1$N.cases <- 25509
phe9_fp1$N.controls <- 40941
law9_fp1 <- law9_fp1[,c(41,3:7,9:10,12)]
law9_fp1$study <- "Lawrenson (2019)"
law9_fp1$N.cases <- 7321
law9_fp1$N.controls <- 4083

oma9_fp1 <- oma9_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma9_fp1, phe9_fp1)
combined1 <- rbind(combined1, law9_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_9_fp1 <- filter(METAr, chromosome == 9)
METAr_9_fp1 <- filter(METAr_9_fp1, grepl("9:16889937", MarkerName))
METAr_9_fp1$study <- "Meta"
METAr_9_fp1$N.cases <- oma9_fp1$N.cases + law9_fp1$N.cases + phe9_fp1$N.cases
METAr_9_fp1$N.controls <- oma9_fp1$N.controls + law9_fp1$N.controls + phe9_fp1$N.controls
METAr_9_fp1 <- METAr_9_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_9_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
```

```
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 9:16889937", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 3:156413477

phelan_3 <- read.table("Summary_chr3.txt", header=TRUE, sep=",")
phelan_3$chrpos <- paste(phelan_3$Chromosome, phelan_3$Position, sep=":")

phe3_fp1 <- filter(phelan_3, grepl("3:156413477", chrpos))
oma3_fp1 <- filter(omara, grepl("3:156413477", chrpos))

oma3_fp1 <- oma3_fp1[,c(26,15,25,17:18,16,19:21)]
oma3_fp1$study <- "O'Mara (2018)"
oma3_fp1$N.cases <- 12906
oma3_fp1$N.controls <- 108979
phe3_fp1 <- phe3_fp1[,c(73,3:7,9:10,12)]
phe3_fp1$study <- "Phelan (2017)"
phe3_fp1$N.cases <- 25509
phe3_fp1$N.controls <- 40941

oma3_fp1 <- oma3_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma3_fp1, phe3_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_3_fp1 <- filter(METAr, chromosome == 3)
METAr_3_fp1 <- filter(METAr_3_fp1, grepl("3:156413477", MarkerName))
METAr_3_fp1$study <- "Meta"
METAr_3_fp1$N.cases <- oma3_fp1$N.cases + phe3_fp1$N.cases
METAr_3_fp1$N.controls <- oma3_fp1$N.controls + phe3_fp1$N.controls
METAr_3_fp1 <- METAr_3_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                               stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_3_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 3:156413477", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 8:129541931
```

```
phelan_8 <- read.table("Summary_chr8.txt", header=TRUE, sep=",")
phelan_8$chrpos <- paste(phelan_8$Chromosome, phelan_8$Position, sep=":")
lawrenson_8 <- read.table("SummaryResults_Asian_chr8.txt", header=TRUE, sep=",")
lawrenson_8$chrpos <- paste(lawrenson_8$Chromosome, lawrenson_8$Position, sep=":")

phe8_fp1 <- filter(phelan_8, grepl("8:129541931", chrpos))
law8_fp1 <- filter(lawrenson_8, grepl("8:129541931", chrpos))
oma8_fp1 <- filter(omara, grepl("8:129541931", chrpos))

oma8_fp1 <- oma8_fp1[,c(26,15,25,17:18,16,19:21)]
oma8_fp1$study <- "O'Mara (2018)"
oma8_fp1$N.cases <- 12906
oma8_fp1$N.controls <- 108979
phe8_fp1 <- phe8_fp1[,c(73,3:7,9:10,12)]
phe8_fp1$study <- "Phelan (2017)"
phe8_fp1$N.cases <- 25509
phe8_fp1$N.controls <- 40941
law8_fp1 <- law8_fp1[,c(41,3:7,9:10,12)]
law8_fp1$study <- "Lawrenson (2019)"
law8_fp1$N.cases <- 7321
law8_fp1$N.controls <- 4083

oma8_fp1 <- oma8_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma8_fp1, phe8_fp1)
combined1 <- rbind(combined1, law8_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_8_fp1 <- filter(METAr, chromosome == 8)
METAr_8_fp1 <- filter(METAr_8_fp1, grepl("8:129541931", MarkerName))
METAr_8_fp1$study <- "Meta"
METAr_8_fp1$N.cases <- oma8_fp1$N.cases + law8_fp1$N.cases + phe8_fp1$N.cases
METAr_8_fp1$N.controls <- oma8_fp1$N.controls + law8_fp1$N.controls + phe8_fp1$N.controls
METAr_8_fp1 <- METAr_8_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_8_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 8:129541931", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 5:1282319

phelan_5 <- read.table("Summary_chr5.txt", header=TRUE, sep=",")
phelan_5$chrpos <- paste(phelan_5$Chromosome, phelan_5$Position, sep=":")
```

```
lawrenson_5 <- read.table("SummaryResults_Asian_chr5.txt", header=TRUE, sep=",")
lawrenson_5$chrpos <- paste(lawrenson_5$Chromosome, lawrenson_5$Position, sep=":")

phe5_fp1 <- filter(phelan_5, grepl("5:1282319", chrpos))
law5_fp1 <- filter(lawrenson_5, grepl("5:1282319", chrpos))
oma5_fp1 <- filter(omara, grepl("5:1282319", chrpos))

oma5_fp1 <- oma5_fp1[,c(26,15,25,17:18,16,19:21)]
oma5_fp1$study <- "O'Mara (2018)"
oma5_fp1$N.cases <- 12906
oma5_fp1$N.controls <- 108979
phe5_fp1 <- phe5_fp1[,c(73,3:7,9:10,12)]
phe5_fp1$study <- "Phelan (2017)"
phe5_fp1$N.cases <- 25509
phe5_fp1$N.controls <- 40941
law5_fp1 <- law5_fp1[,c(41,3:7,9:10,12)]
law5_fp1$study <- "Lawrenson (2019)"
law5_fp1$N.cases <- 7321
law5_fp1$N.controls <- 4083

oma5_fp1 <- oma5_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)
#Two entries containing phrase "5:1282319", subsetting:
phe5_fp1 <- phe5_fp1[!(phe5_fp1$Baseline == "G"),]
law5_fp1 <- law5_fp1[!(law5_fp1$Baseline == "G"),]
oma5_fp1 <- oma5_fp1[!(oma5_fp1$Baseline == "G"),]

combined1 <- rbind(oma5_fp1, phe5_fp1)
combined1 <- rbind(combined1, law5_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_5_fp1 <- filter(METAr, chromosome == 5)
METAr_5_fp1 <- filter(METAr_5_fp1, grepl("5:1282319", MarkerName))

METAr_5_fp1 <- METAr_5_fp1[!(METAr_5_fp1$Allele2 == "g"),]

METAr_5_fp1$study <- "Meta"
METAr_5_fp1$N.cases <- oma5_fp1$N.cases + law5_fp1$N.cases + phe5_fp1$N.cases
METAr_5_fp1$N.controls <- oma5_fp1$N.controls + law5_fp1$N.controls + phe5_fp1$N.controls
METAr_5_fp1 <- METAr_5_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_5_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 5:1282319", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
```

```
#SNP 2:177037831


phelan_2 <- read.table("Summary_chr2.txt", header=TRUE, sep=",")
phelan_2$chrpos <- paste(phelan_2$Chromosome, phelan_2$Position, sep=":")
lawrenson_2 <- read.table("SummaryResults_Asian_chr2.txt", header=TRUE, sep=",")
lawrenson_2$chrpos <- paste(lawrenson_2$Chromosome, lawrenson_2$Position, sep=":")

phe2_fp1 <- filter(phelan_2, grepl("2:177037831", chrpos))
law2_fp1 <- filter(lawrenson_2, grepl("2:177037831", chrpos))
oma2_fp1 <- filter(omara, grepl("2:177037831", chrpos))

oma2_fp1 <- oma2_fp1[,c(26,15,25,17:18,16,19:21)]
oma2_fp1$study <- "O'Mara (2018)"
oma2_fp1$N.cases <- 12906
oma2_fp1$N.controls <- 108979
phe2_fp1 <- phe2_fp1[,c(73,3:7,9:10,12)]
phe2_fp1$study <- "Phelan (2017)"
phe2_fp1$N.cases <- 25509
phe2_fp1$N.controls <- 40941
law2_fp1 <- law2_fp1[,c(41,3:7,9:10,12)]
law2_fp1$study <- "Lawrenson (2019)"
law2_fp1$N.cases <- 7321
law2_fp1$N.controls <- 4083

oma2_fp1 <- oma2_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma2_fp1, phe2_fp1)
combined1 <- rbind(combined1, law2_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_2_fp1 <- filter(METAr, chromosome == 2)
METAr_2_fp1 <- filter(METAr_2_fp1, grepl("2:177037831", MarkerName))
METAr_2_fp1$Effect <- -(as.numeric(METAr_2_fp1$Effect)) #Reversing direction
#for Allele swap.

METAr_2_fp1$study <- "Meta"
METAr_2_fp1$N.cases <- oma2_fp1$N.cases + law2_fp1$N.cases + phe2_fp1$N.cases
METAr_2_fp1$N.controls <- oma2_fp1$N.controls + law2_fp1$N.controls + phe2_fp1$N.controls
METAr_2_fp1 <- METAr_2_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele2, Baseline=Allele1)
                                      #Account for allele swap^
combined2 <- merge(combined1, METAr_2_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 2:177037831", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
```

```
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 8:82653644

phelan_8 <- read.table("Summary_chr8.txt", header=TRUE, sep=",")
phelan_8$chrpos <- paste(phelan_8$Chromosome, phelan_8$Position, sep=":")

phe8_fp2 <- filter(phelan_8, grepl("8:82653644", chrpos))
oma8_fp2 <- filter(omara, grepl("8:82653644", chrpos))

oma8_fp2 <- oma8_fp2[,c(26,15,25,17:18,16,19:21)]
oma8_fp2$study <- "O'Mara (2018)"
oma8_fp2$N.cases <- 12906
oma8_fp2$N.controls <- 108979
phe8_fp2 <- phe8_fp2[,c(73,3:7,9:10,12)]
phe8_fp2$study <- "Phelan (2017)"
phe8_fp2$N.cases <- 25509
phe8_fp2$N.controls <- 40941

oma8_fp2 <- oma8_fp2 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma8_fp2, phe8_fp2)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_8_fp2 <- filter(METAr, chromosome == 8)
METAr_8_fp2 <- filter(METAr_8_fp2, grepl("8:82653644", MarkerName))
METAr_8_fp2$Effect <- -(as.numeric(METAr_8_fp2$Effect)) #Reversing direction
#for Allele swap.

METAr_8_fp2$study <- "Meta"
METAr_8_fp2$N.cases <- oma8_fp2$N.cases + phe8_fp2$N.cases
METAr_8_fp2$N.controls <- oma8_fp2$N.controls + phe8_fp2$N.controls
METAr_8_fp2 <- METAr_8_fp2 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele2, Baseline=Allele1)
                                      #Account for allele swap^

combined2 <- merge(combined1, METAr_8_fp2, all=TRUE)
ordervec <- c("O'Mara (2018)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 8:82653644", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 9:13615500

phelan_9 <- read.table("Summary_chr9.txt", header=TRUE, sep=",")
phelan_9$chrpos <- paste(phelan_9$Chromosome, phelan_9$Position, sep=":")
lawrenson_9 <- read.table("SummaryResults_Asian_chr9.txt", header=TRUE, sep=",")
lawrenson_9$chrpos <- paste(lawrenson_9$Chromosome, lawrenson_9$Position, sep=":")
```

```
phe9_fp2 <- filter(phelan_9, grepl("9:13615500", chrpos))
law9_fp2 <- filter(lawrenson_9, grepl("9:13615500", chrpos))
oma9_fp2 <- filter(omara, grepl("9:13615500", chrpos))

oma9_fp2 <- oma9_fp2[,c(26,15,25,17:18,16,19:21)]
oma9_fp2$study <- "O'Mara (2018)"
oma9_fp2$N.cases <- 12906
oma9_fp2$N.controls <- 108979
phe9_fp2 <- phe9_fp2[,c(73,3:7,9:10,12)]
phe9_fp2$study <- "Phelan (2017)"
phe9_fp2$N.cases <- 25509
phe9_fp2$N.controls <- 40941
law9_fp2 <- law9_fp2[,c(41,3:7,9:10,12)]
law9_fp2$study <- "Lawrenson (2019)"
law9_fp2$N.cases <- 7321
law9_fp2$N.controls <- 4083

oma9_fp2 <- oma9_fp2 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma9_fp2, phe9_fp2)
combined1 <- rbind(combined1, law9_fp2)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_9_fp2 <- filter(METAr, chromosome == 9)
METAr_9_fp2 <- filter(METAr_9_fp2, grepl("9:13615500", MarkerName))
METAr_9_fp2$study <- "Meta"
METAr_9_fp2$N.cases <- oma9_fp2$N.cases + law9_fp2$N.cases + phe9_fp2$N.cases
METAr_9_fp2$N.controls <- oma9_fp2$N.controls + law9_fp2$N.controls + phe9_fp2$N.controls
METAr_9_fp2 <- METAr_9_fp2 %>% rename(chrpos=MarkerName, log_OR=Effect,
                               stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_9_fp2, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 9:13615500", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 10:21821274

phelan_10 <- read.table("Summary_chr10.txt", header=TRUE, sep=",")
phelan_10$chrpos <- paste(phelan_10$Chromosome, phelan_10$Position, sep=":")
lawrenson_10 <- read.table("SummaryResults_Asian_chr10.txt", header=TRUE, sep=",")
lawrenson_10$chrpos <- paste(lawrenson_10$Chromosome, lawrenson_10$Position, sep=":")

phe10_fp1 <- filter(phelan_10, grepl("10:21821274", chrpos))
law10_fp1 <- filter(lawrenson_10, grepl("10:21821274", chrpos))
```

```
oma10_fp1 <- filter(omara, grepl("10:21821274", chrpos))

oma10_fp1 <- oma10_fp1[,c(26,15,25,17:18,16,19:21)]
oma10_fp1$study <- "O'Mara (2018)"
oma10_fp1$N.cases <- 12906
oma10_fp1$N.controls <- 108979
phe10_fp1 <- phe10_fp1[,c(73,3:7,9:10,12)]
phe10_fp1$study <- "Phelan (2017)"
phe10_fp1$N.cases <- 25509
phe10_fp1$N.controls <- 40941
law10_fp1 <- law10_fp1[,c(41,3:7,9:10,12)]
law10_fp1$study <- "Lawrenson (2019)"
law10_fp1$N.cases <- 7321
law10_fp1$N.controls <- 4083

oma10_fp1 <- oma10_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma10_fp1, phe10_fp1)
combined1 <- rbind(combined1, law10_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_10_fp1 <- filter(METAr, chromosome == 10)
METAr_10_fp1 <- filter(METAr_10_fp1, grepl("10:21821274", MarkerName))
METAr_10_fp1$study <- "Meta"
METAr_10_fp1$N.cases <- oma10_fp1$N.cases + law10_fp1$N.cases + phe10_fp1$N.cases
METAr_10_fp1$N.controls <- oma10_fp1$N.controls + law10_fp1$N.controls + phe10_fp1$N.controls
METAr_10_fp1 <- METAr_10_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_10_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 10:21821274", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 13:73814891

phelan_13 <- read.table("Summary_chr13.txt", header=TRUE, sep=",")
phelan_13$chrpos <- paste(phelan_13$Chromosome, phelan_13$Position, sep=":")
lawrenson_13 <- read.table("SummaryResults_Asian_chr13.txt", header=TRUE, sep=",")
lawrenson_13$chrpos <- paste(lawrenson_13$Chromosome, lawrenson_13$Position, sep=":")

phe13_fp1 <- filter(phelan_13, grepl("13:73814891", chrpos))
law13_fp1 <- filter(lawrenson_13, grepl("13:73814891", chrpos))
oma13_fp1 <- filter(omara, grepl("13:73814891", chrpos))

oma13_fp1 <- oma13_fp1[,c(26,15,25,17:18,16,19:21)]
```

```
oma13_fp1$study <- "O'Mara (2018)"
oma13_fp1$N.cases <- 12906
oma13_fp1$N.controls <- 108979
phe13_fp1 <- phe13_fp1[,c(73,3:7,9:10,12)]
phe13_fp1$study <- "Phelan (2017)"
phe13_fp1$N.cases <- 25509
phe13_fp1$N.controls <- 40941
law13_fp1 <- law13_fp1[,c(41,3:7,9:10,12)]
law13_fp1$study <- "Lawrenson (2019)"
law13_fp1$N.cases <- 7321
law13_fp1$N.controls <- 4083

oma13_fp1 <- oma13_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma13_fp1, phe13_fp1)
combined1 <- rbind(combined1, law13_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_13_fp1 <- filter(METAr, chromosome == 13)
METAr_13_fp1 <- filter(METAr_13_fp1, grepl("13:73814891", MarkerName))
METAr_13_fp1$study <- "Meta"
METAr_13_fp1$N.cases <- oma13_fp1$N.cases + law13_fp1$N.cases + phe13_fp1$N.cases
METAr_13_fp1$N.controls <- oma13_fp1$N.controls + law13_fp1$N.controls + phe13_fp1$N.controls
METAr_13_fp1 <- METAr_13_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_13_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 13:73814891", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 17:29181220

phelan_17 <- read.table("Summary_chr17.txt", header=TRUE, sep=",")
phelan_17$chrpos <- paste(phelan_17$Chromosome, phelan_17$Position, sep=":")
lawrenson_17 <- read.table("SummaryResults_Asian_chr17.txt", header=TRUE, sep=",")
lawrenson_17$chrpos <- paste(lawrenson_17$Chromosome, lawrenson_17$Position, sep=":")

phe17_fp1 <- filter(phelan_17, grepl("17:29181220", chrpos))
law17_fp1 <- filter(lawrenson_17, grepl("17:29181220", chrpos))
oma17_fp1 <- filter(omara, grepl("17:29181220", chrpos))

oma17_fp1 <- oma17_fp1[,c(26,15,25,17:18,16,19:21)]
oma17_fp1$study <- "O'Mara (2018)"
oma17_fp1$N.cases <- 12906
oma17_fp1$N.controls <- 108979
```

```
phe17_fp1 <- phe17_fp1[,c(73,3:7,9:10,12)]
phe17_fp1$study <- "Phelan (2017)"
phe17_fp1$N.cases <- 25509
phe17_fp1$N.controls <- 40941
law17_fp1 <- law17_fp1[,c(41,3:7,9:10,12)]
law17_fp1$study <- "Lawrenson (2019)"
law17_fp1$N.cases <- 7321
law17_fp1$N.controls <- 4083

oma17_fp1 <- oma17_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma17_fp1, phe17_fp1)
combined1 <- rbind(combined1, law17_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_17_fp1 <- filter(METAr, chromosome == 17)
METAr_17_fp1 <- filter(METAr_17_fp1, grepl("17:29181220", MarkerName))
METAr_17_fp1$study <- "Meta"
METAr_17_fp1$Effect <- -(as.numeric(METAr_17_fp1$Effect))

METAr_17_fp1$N.cases <- oma17_fp1$N.cases + law17_fp1$N.cases + phe17_fp1$N.cases
METAr_17_fp1$N.controls <- oma17_fp1$N.controls + law17_fp1$N.controls + phe17_fp1$N.controls
METAr_17_fp1 <- METAr_17_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele2, Baseline=Allele1)

combined2 <- merge(combined1, METAr_17_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 17:29181220", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 17:44212527

phelan_17 <- read.table("Summary_chr17.txt", header=TRUE, sep=",")
phelan_17$chrpos <- paste(phelan_17$Chromosome, phelan_17$Position, sep=":")

phe17_fp2 <- filter(phelan_17, grepl("17:44212527", chrpos))
oma17_fp2 <- filter(omara, grepl("17:44212527", chrpos))
#Two different results, METAL used first instance:
oma17_fp2 <- filter(oma17_fp2, oma17_fp2$beta == -0.0276902)
#Two identical rows, keeping first:
oma17_fp2 <- oma17_fp2[1,]

oma17_fp2 <- oma17_fp2[,c(26,15,25,17:18,16,19:21)]
oma17_fp2$study <- "O'Mara (2018)"
oma17_fp2$N.cases <- 12906
oma17_fp2$N.controls <- 108979
```

```
phe17_fp2 <- phe17_fp2[,c(73,3:7,9:10,12)]
phe17_fp2$study <- "Phelan (2017)"
phe17_fp2$N.cases <- 25509
phe17_fp2$N.controls <- 40941

oma17_fp2 <- oma17_fp2 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma17_fp2, phe17_fp2)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_17_fp2 <- filter(METAr, chromosome == 17)
METAr_17_fp2 <- filter(METAr_17_fp2, grepl("17:44212527", MarkerName))
METAr_17_fp2$Effect <- -(as.numeric(METAr_17_fp2$Effect)) #Reversing direction
#for Allele swap.
METAr_17_fp2$study <- "Meta"
METAr_17_fp2$N.cases <- oma17_fp2$N.cases + phe17_fp2$N.cases
METAr_17_fp2$N.controls <- oma17_fp2$N.controls + phe17_fp2$N.controls
METAr_17_fp2 <- METAr_17_fp2 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele2, Baseline=Allele1)
                                      #Account for allele swap^
combined2 <- merge(combined1, METAr_17_fp2, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 17:44212527", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 19:17390917

phelan_19 <- read.table("Summary_chr19.txt", header=TRUE, sep=",")
phelan_19$chrpos <- paste(phelan_19$Chromosome, phelan_19$Position, sep=":")
lawrenson_19 <- read.table("SummaryResults_Asian_chr19.txt", header=TRUE, sep=",")
lawrenson_19$chrpos <- paste(lawrenson_19$Chromosome, lawrenson_19$Position, sep=":")

phe19_fp1 <- filter(phelan_19, grepl("19:17390917", chrpos))
law19_fp1 <- filter(lawrenson_19, grepl("19:17390917", chrpos))
oma19_fp1 <- filter(omara, grepl("19:17390917", chrpos))

oma19_fp1 <- oma19_fp1[,c(26,15,25,17:18,16,19:21)]
oma19_fp1$study <- "O'Mara (2018)"
oma19_fp1$N.cases <- 12906
oma19_fp1$N.controls <- 108979
phe19_fp1 <- phe19_fp1[,c(73,3:7,9:10,12)]
phe19_fp1$study <- "Phelan (2017)"
phe19_fp1$N.cases <- 25509
phe19_fp1$N.controls <- 40941
law19_fp1 <- law19_fp1[,c(41,3:7,9:10,12)]
law19_fp1$study <- "Lawrenson (2019)"
```

```
law19_fp1$N.cases <- 7321
law19_fp1$N.controls <- 4083

oma19_fp1 <- oma19_fp1 %>% rename(Chromosome=chromosome, Position=base_pair_location,
                        Effect=effect_allele, Baseline=other_allele,
                        EAF=effect_allele_frequency, overall_OR = beta,
                        overall_SE=standard_error, overall_pvalue=p_value)

combined1 <- rbind(oma19_fp1, phe19_fp1)
combined1 <- rbind(combined1, law19_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_19_fp1 <- filter(METAr, chromosome == 19)
METAr_19_fp1 <- filter(METAr_19_fp1, grepl("19:17390917", MarkerName))
METAr_19_fp1$study <- "Meta"
METAr_19_fp1$N.cases <- oma19_fp1$N.cases + law19_fp1$N.cases + phe19_fp1$N.cases
METAr_19_fp1$N.controls <- oma19_fp1$N.controls + law19_fp1$N.controls + phe19_fp1$N.controls
METAr_19_fp1 <- METAr_19_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_19_fp1, all=TRUE)
ordervec <- c("O'Mara (2018)", "Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 19:17390917", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#Make tables for significant results, alongside input alleles:
#^ Filtering 2,155 observations just to select SNPs:
alleleframe <- METAr_sig[c(METAr_sig$MarkerName %in% c("2:177037831", "3:156413477", "5:1282319",
                                "8:82653644", "8:129541931", "9:136155000",
                                 "9:16889937", "10:21821274", "13:73814891",
                                "17:29181220", "17:44212527", "19:17390917")),]
omara_alleles <- filter(omara, omara$chrpos %in% alleleframe$MarkerName)
omara_alleles <- omara_alleles[,c(26,17:18)]
alleleframe <- merge(alleleframe, omara_alleles, by.x="MarkerName", by.y="chrpos", all.x=TRUE)
alleleframe <- alleleframe[,c(1:3,14:15,4:13)]


alleleframe <- alleleframe %>% rename(SNP=MarkerName, `Input Effect Allele` = effect_allele,
                            `Input Base Allele` = other_allele, `Output Effect Allele` = Allele1,
                            `Output Base Allele` = Allele2, Beta = Effect, SE=StdErr,
                            `P-Value` = P.value, `Het. I-Squared` = HetISq, `Het. Chi-Squared` = HetChiSq,
                            `Het. DF` = HetDf, `Het. P-Value` = HetPVal, Chromosome = chromosome)
alleleframe <- arrange(alleleframe, alleleframe$Chromosome, alleleframe$Position)
alleleframe$`Output Effect Allele` <- toupper(alleleframe$`Output Effect Allele`)
alleleframe$`Output Base Allele` <- toupper(alleleframe$`Output Base Allele`)

alleleframe_res <- alleleframe[,c(1:11)]
alleleframe_het <- alleleframe[,c(1:3,11:15)]
```

```
	library(stargazer)
	invisible(stargazer(alleleframe_res, type="html", summary=FALSE, rownames=FALSE,
out="ovarendo_meta_results_table.doc",
		title="Ovarian+Endometrial Meta-Analysis Significant Results"))
	invisible(stargazer(alleleframe_het, type="html", summary=FALSE, rownames=FALSE,
out="ovarendo_meta_results_table_2.doc",
		title="Ovarian+Endometrial Meta-Analysis Significant Results - Heterogeneity Test"))
	Ovarian + Cervical META ####

	METAr1_c <- read.table("META_Chr1_c_1.txt", header=TRUE, sep="\t")
	METAr1_c$chromosome <- 1

	METAr2_c <- read.table("META_Chr2_c_1.txt", header=TRUE, sep="\t")
	METAr2_c$chromosome <- 2

	METAr3_c <- read.table("META_Chr3_c_1.txt", header=TRUE, sep="\t")
	METAr3_c$chromosome <- 3

	METAr4_c <- read.table("META_Chr4_c_1.txt", header=TRUE, sep="\t")
	METAr4_c$chromosome <- 4

	METAr5_c <- read.table("META_Chr5_c_1.txt", header=TRUE, sep="\t")
	METAr5_c$chromosome <- 5

	METAr6_c <- read.table("META_Chr6_c_1.txt", header=TRUE, sep="\t")
	METAr6_c$chromosome <- 6

	METAr7_c <- read.table("META_Chr7_c_1.txt", header=TRUE, sep="\t")
	METAr7_c$chromosome <- 7

	METAr8_c <- read.table("META_Chr8_c_1.txt", header=TRUE, sep="\t")
	METAr8_c$chromosome <- 8

	METAr9_c <- read.table("META_Chr9_c_1.txt", header=TRUE, sep="\t")
	METAr9_c$chromosome <- 9

	METAr10_c <- read.table("META_Chr10_c_1.txt", header=TRUE, sep="\t")
	METAr10_c$chromosome <- 10

	METAr11_c <- read.table("META_Chr11_c_1.txt", header=TRUE, sep="\t")
	METAr11_c$chromosome <- 11

	METAr12_c <- read.table("META_Chr12_c_1.txt", header=TRUE, sep="\t")
	METAr12_c$chromosome <- 12

	METAr13_c <- read.table("META_Chr13_c_1.txt", header=TRUE, sep="\t")
	METAr13_c$chromosome <- 13

	METAr14_c <- read.table("META_Chr14_c_1.txt", header=TRUE, sep="\t")
	METAr14_c$chromosome <- 14

	METAr15_c <- read.table("META_Chr15_c_1.txt", header=TRUE, sep="\t")
	METAr15_c$chromosome <- 15

	METAr16_c <- read.table("META_Chr16_c_1.txt", header=TRUE, sep="\t")
	METAr16_c$chromosome <- 16

	METAr17_c <- read.table("META_Chr17_c_1.txt", header=TRUE, sep="\t")
	METAr17_c$chromosome <- 17

	METAr18_c <- read.table("META_Chr18_c_1.txt", header=TRUE, sep="\t")
```

```
METAr18_c$chromosome <- 18

METAr19_c <- read.table("META_Chr19_c_1.txt", header=TRUE, sep="\t")
METAr19_c$chromosome <- 19

METAr20_c <- read.table("META_Chr20_c_1.txt", header=TRUE, sep="\t")
METAr20_c$chromosome <- 20

METAr21_c <- read.table("META_Chr21_c_1.txt", header=TRUE, sep="\t")
METAr21_c$chromosome <- 21

METAr22_c <- read.table("META_Chr22_c_1.txt", header=TRUE, sep="\t")
METAr22_c$chromosome <- 22
METAr_c <- rbind(METAr1_c, METAr2_c, METAr3_c, METAr4_c, METAr5_c, METAr6_c, METAr7_c, METAr8_c, METAr9_c, METAr10_c, METAr11_c, METAr12_c, METAr13_c, METAr14_c, METAr15_c, METAr16_c, METAr17_c, METAr18_c, METAr19_c, METAr20_c, METAr21_c, METAr22_c)
rm(list = c('METAr1_c','METAr2_c','METAr3_c','METAr4_c','METAr5_c','METAr6_c',
        'METAr7_c','METAr8_c','METAr9_c','METAr10_c','METAr11_c','METAr12_c',
        'METAr13_c','METAr14_c','METAr15_c','METAr16_c','METAr17_c','METAr18_c',
        'METAr19_c','METAr20_c','METAr21_c','METAr22_c'))
METAr_c$Position <- gsub("^.*:", "", METAr_c$MarkerName)

METAr_c$chromosome <- as.numeric(METAr_c$chromosome)
METAr_c$Position <- as.numeric(METAr_c$Position)
METAr_c <- arrange(METAr_c, chromosome, Position)

METAr_c$chrposal <- paste(METAr_c$MarkerName, METAr_c$Allele2, sep="_")
METAr_c$chrposal <- paste(METAr_c$chrposal, METAr_c$Allele1, sep="/" )

#Assessing number of missing underlying studies by marker:
#Cut down SNPs to just those in at least 2 underlying studies (NO ovarian only):

METAr_c$num_miss <- str_count(as.character(METAr_c$Direction), "\\?")
METAr_c$ukb_miss <- substr(METAr_c$Direction, 1, 1)
METAr_c$rowID <- row.names(METAr_c)

METAr_c_filtered <- METAr_c %>% filter(ukb_miss == '?' & num_miss == 1)
METAr_c <- METAr_c[!(METAr_c$rowID %in% METAr_c_filtered$rowID),]
METAr_c <- METAr_c %>% filter(num_miss < 2) #can't be estimates from only one study

#write.table(METAr_c, sep="\t", file="C:\\Users\\Mark\\Documents\\META_oc_filtered.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)


#rm(METAr_c)
METAr_c_sig <- filter(METAr_c, P.value <= 0.00000005)
METAr_c_sig <- METAr_c_sig[,c(12:13,1:11)]
ukbb <- as.data.frame(fread("C53.gwas.imputed_v3.female.tsv"))
#Mapping UKBB rsIDs:
ukvariants <- as.data.frame(fread("variants.tsv"))
ukvariants$chrpos <- paste(ukvariants$chr, ukvariants$pos, sep=":")
ukbb <- merge(x=ukbb, y=ukvariants[,c("variant", "rsid", "chr", "pos", "chrpos", "ref", "alt")], by="variant", all.x=TRUE)

rm(ukvariants)

#Filter MAF < 0.01, nonsense p-val,se:
ukbb <- ukbb %>% filter(minor_AF > 0.01) %>% filter(se > 0) %>%
        filter(pval < 1 & pval > 0)

ukbb <- arrange(ukbb, chr, pos)
library(metaviz)
```

```
#SNP 12:121056263

phelan_12 <- read.table("Summary_chr12.txt", header=TRUE, sep=",")
phelan_12$chrpos <- paste(phelan_12$Chromosome, phelan_12$Position, sep=":")

phe12_fp1 <- filter(phelan_12, grepl("12:121056263", chrpos))
ukb12_fp1 <- filter(ukbb, grepl("12:121056263", chrpos))

ukb12_fp1 <- ukb12_fp1[,c(16,14:15,18,17,3,9:10,12)]
ukb12_fp1$study <- "UK Biobank (2018)"
ukb12_fp1$N.cases <- 192
ukb12_fp1$N.controls <- 193982
phe12_fp1 <- phe12_fp1[,c(73,3:7,9:10,12)]
phe12_fp1$study <- "Phelan (2017)"
phe12_fp1$N.cases <- 25509
phe12_fp1$N.controls <- 40941

ukb12_fp1 <- ukb12_fp1 %>% rename(Chromosome=chr, Position=pos,
                        Effect=alt, Baseline=ref,
                        EAF=minor_AF, overall_OR = beta,
                        overall_SE=se, overall_pvalue=pval)

combined1 <- rbind(ukb12_fp1, phe12_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_c_12_fp1 <- filter(METAr_c, chromosome == 12)
METAr_c_12_fp1 <- filter(METAr_c_12_fp1, grepl("12:121056263", MarkerName))
METAr_c_12_fp1$study <- "Meta"
METAr_c_12_fp1$N.cases <- ukb12_fp1$N.cases + phe12_fp1$N.cases
METAr_c_12_fp1$N.controls <- ukb12_fp1$N.controls + phe12_fp1$N.controls
METAr_c_12_fp1 <- METAr_c_12_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_c_12_fp1, all=TRUE)
ordervec <- c("UK Biobank (2018)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        x_limit = c(-0.01,0.0099),
        xlab="OR SNP 12:121056263", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 12:31954662

phe12_fp1 <- filter(phelan_12, grepl("12:31954662", chrpos))
ukb12_fp1 <- filter(ukbb, grepl("12:31954662", chrpos))

ukb12_fp1 <- ukb12_fp1[,c(16,14:15,18,17,3,9:10,12)]
ukb12_fp1$study <- "UK Biobank (2018)"
ukb12_fp1$N.cases <- 192
ukb12_fp1$N.controls <- 193982
phe12_fp1 <- phe12_fp1[,c(73,3:7,9:10,12)]
```

```
phe12_fp1$study <- "Phelan (2017)"
phe12_fp1$N.cases <- 25509
phe12_fp1$N.controls <- 40941

ukb12_fp1 <- ukb12_fp1 %>% rename(Chromosome=chr, Position=pos,
                        Effect=alt, Baseline=ref,
                        EAF=minor_AF, overall_OR = beta,
                        overall_SE=se, overall_pvalue=pval)

combined1 <- rbind(ukb12_fp1, phe12_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                         P.value=overall_pvalue, MAF=EAF)

METAr_c_12_fp1 <- filter(METAr_c, chromosome == 12)
METAr_c_12_fp1 <- filter(METAr_c_12_fp1, grepl("12:31954662", MarkerName))
METAr_c_12_fp1$Effect <- -(as.numeric(METAr_c_12_fp1$Effect))
METAr_c_12_fp1$study <- "Meta"
METAr_c_12_fp1$N.cases <- ukb12_fp1$N.cases + phe12_fp1$N.cases
METAr_c_12_fp1$N.controls <- ukb12_fp1$N.controls + phe12_fp1$N.controls
METAr_c_12_fp1 <- METAr_c_12_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele2, Baseline=Allele1)

combined2 <- merge(combined1, METAr_c_12_fp1, all=TRUE)
ordervec <- c("UK Biobank (2018)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        x_limit = c(-0.01,0.0099),
        xlab="OR SNP 12:31954662", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 18:31370348

phelan_18 <- read.table("Summary_chr18.txt", header=TRUE, sep=",")
phelan_18$chrpos <- paste(phelan_18$Chromosome, phelan_18$Position, sep=":")

phe18_fp1 <- filter(phelan_18, grepl("18:31370348", chrpos))
ukb18_fp1 <- filter(ukbb, grepl("18:31370348", chrpos))

ukb18_fp1 <- ukb18_fp1[,c(16,14:15,18,17,3,9:10,12)]
ukb18_fp1$study <- "UK Biobank (2018)"
ukb18_fp1$N.cases <- 192
ukb18_fp1$N.controls <- 193982
phe18_fp1 <- phe18_fp1[,c(73,3:7,9:10,12)]
phe18_fp1$study <- "Phelan (2017)"
phe18_fp1$N.cases <- 25509
phe18_fp1$N.controls <- 40941

ukb18_fp1 <- ukb18_fp1 %>% rename(Chromosome=chr, Position=pos,
                        Effect=alt, Baseline=ref,
                        EAF=minor_AF, overall_OR = beta,
                        overall_SE=se, overall_pvalue=pval)
```

```
combined1 <- rbind(ukb18_fp1, phe18_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_c_18_fp1 <- filter(METAr_c, chromosome == 18)
METAr_c_18_fp1 <- filter(METAr_c_18_fp1, grepl("18:31370348", MarkerName))
METAr_c_18_fp1$study <- "Meta"
METAr_c_18_fp1$N.cases <- ukb18_fp1$N.cases + phe18_fp1$N.cases
METAr_c_18_fp1$N.controls <- ukb18_fp1$N.controls + phe18_fp1$N.controls
METAr_c_18_fp1 <- METAr_c_18_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_c_18_fp1, all=TRUE)
ordervec <- c("UK Biobank (2018)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        x_limit=c(-0.01,0.0099),
        xlab="OR SNP 18:31370348", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 6:66363605

phelan_6 <- read.table("Summary_chr6.txt", header=TRUE, sep=",")
phelan_6$chrpos <- paste(phelan_6$Chromosome, phelan_6$Position, sep=":")

phe6_fp1 <- filter(phelan_6, grepl("6:66363605", chrpos))
ukb6_fp1 <- filter(ukbb, grepl("6:66363605", chrpos))

ukb6_fp1 <- ukb6_fp1[,c(16,14:15,18,17,3,9:10,12)]
ukb6_fp1$study <- "UK Biobank (2018)"
ukb6_fp1$N.cases <- 192
ukb6_fp1$N.controls <- 193982
phe6_fp1 <- phe6_fp1[,c(73,3:7,9:10,12)]
phe6_fp1$study <- "Phelan (2017)"
phe6_fp1$N.cases <- 25509
phe6_fp1$N.controls <- 40941

ukb6_fp1 <- ukb6_fp1 %>% rename(Chromosome=chr, Position=pos,
                        Effect=alt, Baseline=ref,
                        EAF=minor_AF, overall_OR = beta,
                        overall_SE=se, overall_pvalue=pval)

combined1 <- rbind(ukb6_fp1, phe6_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_c_6_fp1 <- filter(METAr_c, chromosome == 6)
METAr_c_6_fp1 <- filter(METAr_c_6_fp1, grepl("6:66363605", MarkerName))
METAr_c_6_fp1$study <- "Meta"
METAr_c_6_fp1$N.cases <- ukb6_fp1$N.cases + phe6_fp1$N.cases
METAr_c_6_fp1$N.controls <- ukb6_fp1$N.controls + phe6_fp1$N.controls
```

```
METAr_c_6_fp1 <- METAr_c_6_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_c_6_fp1, all=TRUE)
ordervec <- c("UK Biobank (2018)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,23)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
      variant="rain",
      x_trans_function = exp,
      x_limit = c(-0.01,0.0099),
      xlab="OR SNP 6:66363605", annotate_CI = TRUE,
      study_table = study.table,
      table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
      table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#Make tables for significant results, alongside input alleles:
alleleframe <- METAr_c_sig
ukbb_alleles <- filter(ukbb, ukbb$chrpos %in% alleleframe$MarkerName)
ukbb_alleles <- ukbb_alleles[,c(16:18)]
alleleframe <- merge(alleleframe, ukbb_alleles, by.x="MarkerName", by.y="chrpos", all.x=TRUE)
alleleframe <- alleleframe[,c(1:3,14:15,4:13)]


alleleframe <- alleleframe %>% rename(SNP=MarkerName, `Input Effect Allele` = alt,
                          `Input Base Allele` = ref, `Output Effect Allele` = Allele1,
                          `Output Base Allele` = Allele2, Beta = Effect, SE=StdErr,
                          `P-Value` = P.value, `Het. I-Squared` = HetISq, `Het. Chi-Squared` = HetChiSq,
                          `Het. DF` = HetDf, `Het. P-Value` = HetPVal, Chromosome = chromosome)
alleleframe <- arrange(alleleframe, alleleframe$Chromosome, alleleframe$Position)
alleleframe$`Output Effect Allele` <- toupper(alleleframe$`Output Effect Allele`)
alleleframe$`Output Base Allele` <- toupper(alleleframe$`Output Base Allele`)

alleleframe_res <- alleleframe[,c(1:11)]
alleleframe_het <- alleleframe[,c(1:3,11:15)]


library(stargazer)
stargazer(alleleframe_res, type="html", summary=FALSE, rownames=FALSE, out="ovarcerv_meta_results_table.doc",
      title="Ovarian+Cervical Meta-Analysis Significant Results")
stargazer(alleleframe_het, type="html", summary=FALSE, rownames=FALSE,
out="ovarcerv_meta_results_table_2.doc",
      title="Ovarian+Cervical Meta-Analysis Significant Results - Heterogeneity Test")
Ovarian Only Meta: ######

library(data.table)
library(tidyverse)

ukbb <- as.data.frame(fread("C56.gwas.imputed_v3.female.tsv"))
#Mapping UKBB rsIDs:
ukvariants <- as.data.frame(fread("variants.tsv"))
ukvariants$chrpos <- paste(ukvariants$chr, ukvariants$pos, sep=":")
ukbb <- merge(x=ukbb, y=ukvariants[,c("variant", "rsid", "chr", "pos", "chrpos", "ref", "alt")], by="variant",
all.x=TRUE)

rm(ukvariants)

#Filter MAF < 0.01, nonsense p-val,se:
```

```
ukbb <- ukbb %>% filter(minor_AF > 0.01) %>% filter(se > 0) %>%
        filter(pval < 1 & pval > 0)

ukbb <- arrange(ukbb, chr, pos)
#Write out pretest for ovarian version of UKBB:
#write.table(ukbb, sep="\t", file="C:\\Users\\Mark\\Documents\\ukbb_ovar_pretest.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)
ukbb_1 <- filter(ukbb, grepl("^1:", chrpos))
ukbb_1 <- filter(ukbb_1, !grepl("NA", chrpos))
ukbb_1 <- ukbb_1[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_1, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr1\\ukbb_2_1.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_2 <- filter(ukbb, grepl("^2:", chrpos))
ukbb_2 <- filter(ukbb_2, !grepl("NA", chrpos))
ukbb_2 <- ukbb_2[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_2, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr2\\ukbb_2_2.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_3 <- filter(ukbb, grepl("^3:", chrpos))
ukbb_3 <- filter(ukbb_3, !grepl("NA", chrpos))
ukbb_3 <- ukbb_3[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_3, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr3\\ukbb_2_3.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_4 <- filter(ukbb, grepl("^4:", chrpos))
ukbb_4 <- filter(ukbb_4, !grepl("NA", chrpos))
ukbb_4 <- ukbb_4[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_4, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr4\\ukbb_2_4.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_5 <- filter(ukbb, grepl("^5:", chrpos))
ukbb_5 <- filter(ukbb_5, !grepl("NA", chrpos))
ukbb_5 <- ukbb_5[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_5, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr5\\ukbb_2_5.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_6 <- filter(ukbb, grepl("^6:", chrpos))
ukbb_6 <- filter(ukbb_6, !grepl("NA", chrpos))
ukbb_6 <- ukbb_6[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_6, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr6\\ukbb_2_6.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_7 <- filter(ukbb, grepl("^7:", chrpos))
ukbb_7 <- filter(ukbb_7, !grepl("NA", chrpos))
ukbb_7 <- ukbb_7[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_7, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr7\\ukbb_2_7.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_8 <- filter(ukbb, grepl("^8:", chrpos))
ukbb_8 <- filter(ukbb_8, !grepl("NA", chrpos))
ukbb_8 <- ukbb_8[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_8, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr8\\ukbb_2_8.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_9 <- filter(ukbb, grepl("^9:", chrpos))
ukbb_9 <- filter(ukbb_9, !grepl("NA", chrpos))
ukbb_9 <- ukbb_9[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_9, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr9\\ukbb_2_9.txt", row.names=FALSE,
col.names=TRUE, quote=FALSE)

ukbb_10 <- filter(ukbb, grepl("^10:", chrpos))
```

```
ukbb_10 <- filter(ukbb_10, !grepl("NA", chrpos))
ukbb_10 <- ukbb_10[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_10, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr10\\ukbb_2_10.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_11 <- filter(ukbb, grepl("^11:", chrpos))
ukbb_11 <- filter(ukbb_11, !grepl("NA", chrpos))
ukbb_11 <- ukbb_11[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_11, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr11\\ukbb_2_11.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_12 <- filter(ukbb, grepl("^12:", chrpos))
ukbb_12 <- filter(ukbb_12, !grepl("NA", chrpos))
ukbb_12 <- ukbb_12[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_12, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr12\\ukbb_2_12.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_13 <- filter(ukbb, grepl("^13:", chrpos))
ukbb_13 <- filter(ukbb_13, !grepl("NA", chrpos))
ukbb_13 <- ukbb_13[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_13, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr13\\ukbb_2_13.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_14 <- filter(ukbb, grepl("^14:", chrpos))
ukbb_14 <- filter(ukbb_14, !grepl("NA", chrpos))
ukbb_14 <- ukbb_14[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_14, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr14\\ukbb_2_14.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_15 <- filter(ukbb, grepl("^15:", chrpos))
ukbb_15 <- filter(ukbb_15, !grepl("NA", chrpos))
ukbb_15 <- ukbb_15[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_15, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr15\\ukbb_2_15.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_16 <- filter(ukbb, grepl("^16:", chrpos))
ukbb_16 <- filter(ukbb_16, !grepl("NA", chrpos))
ukbb_16 <- ukbb_16[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_16, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr16\\ukbb_2_16.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_17 <- filter(ukbb, grepl("^17:", chrpos))
ukbb_17 <- filter(ukbb_17, !grepl("NA", chrpos))
ukbb_17 <- ukbb_17[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_17, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr17\\ukbb_2_17.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_18 <- filter(ukbb, grepl("^18:", chrpos))
ukbb_18 <- filter(ukbb_18, !grepl("NA", chrpos))
ukbb_18 <- ukbb_18[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_18, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr18\\ukbb_2_18.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_19 <- filter(ukbb, grepl("^19:", chrpos))
ukbb_19 <- filter(ukbb_19, !grepl("NA", chrpos))
ukbb_19 <- ukbb_19[,c(16,13,17,18,9,10,12)]
#write.table(ukbb_19, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr19\\ukbb_2_19.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

ukbb_20 <- filter(ukbb, grepl("^20:", chrpos))
ukbb_20 <- filter(ukbb_20, !grepl("NA", chrpos))
ukbb_20 <- ukbb_20[,c(16,13,17,18,9,10,12)]
```

```
	#write.table(ukbb_20, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr20\\ukbb_2_20.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

	ukbb_21 <- filter(ukbb, grepl("^21:", chrpos))
	ukbb_21 <- filter(ukbb_21, !grepl("NA", chrpos))
	ukbb_21 <- ukbb_21[,c(16,13,17,18,9,10,12)]
	#write.table(ukbb_21, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr21\\ukbb_2_21.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)

	ukbb_22 <- filter(ukbb, grepl("^22:", chrpos))
	ukbb_22 <- filter(ukbb_22, !grepl("NA", chrpos))
	ukbb_22 <- ukbb_22[,c(16,13,17,18,9,10,12)]
	#write.table(ukbb_22, sep=",", file="C:\\msys64\\home\\Mark\\METAL\\Chr22\\ukbb_2_22.txt", row.names=FALSE, col.names=TRUE, quote=FALSE)
	METAr1_2 <- read.table("META_Chr1_2_1.txt", header=TRUE, sep="\t")
	METAr1_2$chromosome <- 1

	METAr2_2 <- read.table("META_Chr2_2_1.txt", header=TRUE, sep="\t")
	METAr2_2$chromosome <- 2

	METAr3_2 <- read.table("META_Chr3_2_1.txt", header=TRUE, sep="\t")
	METAr3_2$chromosome <- 3

	METAr4_2 <- read.table("META_Chr4_2_1.txt", header=TRUE, sep="\t")
	METAr4_2$chromosome <- 4

	METAr5_2 <- read.table("META_Chr5_2_1.txt", header=TRUE, sep="\t")
	METAr5_2$chromosome <- 5

	METAr6_2 <- read.table("META_Chr6_2_1.txt", header=TRUE, sep="\t")
	METAr6_2$chromosome <- 6

	METAr7_2 <- read.table("META_Chr7_2_1.txt", header=TRUE, sep="\t")
	METAr7_2$chromosome <- 7

	METAr8_2 <- read.table("META_Chr8_2_1.txt", header=TRUE, sep="\t")
	METAr8_2$chromosome <- 8

	METAr9_2 <- read.table("META_Chr9_2_1.txt", header=TRUE, sep="\t")
	METAr9_2$chromosome <- 9

	METAr10_2 <- read.table("META_Chr10_2_1.txt", header=TRUE, sep="\t")
	METAr10_2$chromosome <- 10

	METAr11_2 <- read.table("META_Chr11_2_1.txt", header=TRUE, sep="\t")
	METAr11_2$chromosome <- 11

	METAr12_2 <- read.table("META_Chr12_2_1.txt", header=TRUE, sep="\t")
	METAr12_2$chromosome <- 12

	METAr13_2 <- read.table("META_Chr13_2_1.txt", header=TRUE, sep="\t")
	METAr13_2$chromosome <- 13

	METAr14_2 <- read.table("META_Chr14_2_1.txt", header=TRUE, sep="\t")
	METAr14_2$chromosome <- 14

	METAr15_2 <- read.table("META_Chr15_2_1.txt", header=TRUE, sep="\t")
	METAr15_2$chromosome <- 15

	METAr16_2 <- read.table("META_Chr16_2_1.txt", header=TRUE, sep="\t")
	METAr16_2$chromosome <- 16
```

```
METAr17_2 <- read.table("META_Chr17_2_1.txt", header=TRUE, sep="\t")
METAr17_2$chromosome <- 17

METAr18_2 <- read.table("META_Chr18_2_1.txt", header=TRUE, sep="\t")
METAr18_2$chromosome <- 18

METAr19_2 <- read.table("META_Chr19_2_1.txt", header=TRUE, sep="\t")
METAr19_2$chromosome <- 19

METAr20_2 <- read.table("META_Chr20_2_1.txt", header=TRUE, sep="\t")
METAr20_2$chromosome <- 20

METAr21_2 <- read.table("META_Chr21_2_1.txt", header=TRUE, sep="\t")
METAr21_2$chromosome <- 21

METAr22_2 <- read.table("META_Chr22_2_1.txt", header=TRUE, sep="\t")
METAr22_2$chromosome <- 22

METAr23 <- read.table("META_Chr23_1.txt", header=TRUE, sep="\t")
METAr23$chromosome <- 23
METAr_2 <- rbind(METAr1_2, METAr2_2, METAr3_2, METAr4_2, METAr5_2, METAr6_2, METAr7_2,
                 METAr8_2, METAr9_2, METAr10_2, METAr11_2, METAr12_2, METAr13_2,
                METAr14_2, METAr15_2, METAr16_2, METAr17_2, METAr18_2, METAr19_2,
                METAr20_2, METAr21_2, METAr22_2, METAr23)
rm(list = c('METAr1_2','METAr2_2','METAr3_2','METAr4_2','METAr5_2','METAr6_2',
          'METAr7_2','METAr8_2','METAr9_2','METAr10_2','METAr11_2','METAr12_2',
          'METAr13_2','METAr14_2','METAr15_2','METAr16_2','METAr17_2','METAr18_2',
          'METAr19_2','METAr20_2','METAr21_2','METAr22_2','METAr23'))
METAr_2$Position <- gsub("^.*:", "", METAr_2$MarkerName)

METAr_2$chromosome <- as.numeric(METAr_2$chromosome)
METAr_2$Position <- as.numeric(METAr_2$Position)
METAr_2 <- arrange(METAr_2, chromosome, Position)


#Assessing number of missing underlying studies by marker:
#Cut down SNPs to just those in at least 2 underlying studies:


METAr_2$num_miss <- str_count(as.character(METAr_2$Direction), "\\?")
METAr_2 <- METAr_2 %>% filter(num_miss < 2)

#write.table(METAr_2, sep="\t", file="C:\\Users\\Mark\\Documents\\METAr_2_filtered.txt", row.names = FALSE,
col.names = TRUE, quote = FALSE)


#rm(METAr_2)
#Filtering for significant SNPs:
METAr_2_sig <- filter(METAr_2, P.value <= 0.00000005)
METAr_2_sig <- METAr_2_sig[,c(12:13,1:11)]

#Fix one ambiguous allele - 17:46505002 shows "<ins:me:alu>" as an effect allele,
#which is a generic term for an ALU element insertion. According to dbSNP
#(https://www.ncbi.nlm.nih.gov/snp/?term=46505002%5BPOSITION_GRCH37%5D+AND+17%5BCHR%5D)
#alternate allele is G.

METAr_2_sig$Allele1 <- gsub("<ins:me:alu>", "g", METAr_2_sig$Allele1)
library(metaviz)
#SNP 9:16913836 - (none present in UKBB ovarian data)

phelan_9 <- read.table("Summary_chr9.txt", header=TRUE, sep=",")
lawrenson_9 <- read.table("SummaryResults_Asian_chr9.txt", header=TRUE, sep=",")
```

```
phelan_9$chrpos <- paste(phelan_9$Chromosome, phelan_9$Position, sep=":")
lawrenson_9$chrpos <- paste(lawrenson_9$Chromosome, lawrenson_9$Position, sep=":")

phelan_9 <- phelan_9 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_9 <- lawrenson_9 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)

law9_fp1 <- filter(lawrenson_9, grepl("9:16913836", chrpos))
phe9_fp1 <- filter(phelan_9, grepl("9:16913836", chrpos))

law9_fp1 <- law9_fp1[,c(41,3:7,9,10,12)]
law9_fp1$study <- "Lawrenson (2019)"
law9_fp1$N.cases <- 7321
law9_fp1$N.controls <- 4083
phe9_fp1 <- phe9_fp1[,c(73,3:7,9:10,12)]
phe9_fp1$study <- "Phelan (2017)"
phe9_fp1$N.cases <- 25509
phe9_fp1$N.controls <- 40941

combined1 <- rbind(law9_fp1, phe9_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                        P.value=overall_pvalue, MAF=EAF)

METAr_2_9_fp1 <- filter(METAr_2, chromosome == 9)
METAr_2_9_fp1 <- filter(METAr_2_9_fp1, grepl("9:16913836", MarkerName))
METAr_2_9_fp1$study <- "Meta"
METAr_2_9_fp1$N.cases <- law9_fp1$N.cases + phe9_fp1$N.cases
METAr_2_9_fp1$N.controls <- law9_fp1$N.controls + phe9_fp1$N.controls
METAr_2_9_fp1 <- METAr_2_9_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_2_9_fp1, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]


viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 9:16913836", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 3:156481585

phelan_3 <- read.table("Summary_chr3.txt", header=TRUE, sep=",")
lawrenson_3 <- read.table("SummaryResults_Asian_chr3.txt", header=TRUE, sep=",")

phelan_3$chrpos <- paste(phelan_3$Chromosome, phelan_3$Position, sep=":")
lawrenson_3$chrpos <- paste(lawrenson_3$Chromosome, lawrenson_3$Position, sep=":")

phelan_3 <- phelan_3 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_3 <- lawrenson_3 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)
```

```
law3_fp1 <- filter(lawrenson_3, grepl("3:156481585", chrpos))
phe3_fp1 <- filter(phelan_3, grepl("3:156481585", chrpos))
phe3_fp1 <- phe3_fp1 %>% filter(Effect == "C") #has two matches, one matches alleles to other

law3_fp1 <- law3_fp1[,c(41,3:7,9,10,12)]
law3_fp1$study <- "Lawrenson (2019)"
law3_fp1$N.cases <- 7321
law3_fp1$N.controls <- 4083
phe3_fp1 <- phe3_fp1[,c(73,3:7,9:10,12)]
phe3_fp1$study <- "Phelan (2017)"
phe3_fp1$N.cases <- 25509
phe3_fp1$N.controls <- 40941

combined1 <- rbind(law3_fp1, phe3_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                        P.value=overall_pvalue, MAF=EAF)

METAr_2_3_fp1 <- filter(METAr_2, chromosome == 3)
METAr_2_3_fp1 <- filter(METAr_2_3_fp1, grepl("3:156481585", MarkerName))
METAr_2_3_fp1$Effect <- -(as.numeric(METAr_2_3_fp1$Effect))

METAr_2_3_fp1$study <- "Meta"
METAr_2_3_fp1$N.cases <- law3_fp1$N.cases + phe3_fp1$N.cases
METAr_2_3_fp1$N.controls <- law3_fp1$N.controls + phe3_fp1$N.controls
METAr_2_3_fp1 <- METAr_2_3_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                              stderr=StdErr, Effect=Allele2, Baseline=Allele1)

combined2 <- merge(combined1, METAr_2_3_fp1, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]

viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 3:156481585", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 17:44260644

phelan_17 <- read.table("Summary_chr17.txt", header=TRUE, sep=",")
lawrenson_17 <- read.table("SummaryResults_Asian_chr17.txt", header=TRUE, sep=",")

phelan_17$chrpos <- paste(phelan_17$Chromosome, phelan_17$Position, sep=":")
lawrenson_17$chrpos <- paste(lawrenson_17$Chromosome, lawrenson_17$Position, sep=":")

phelan_17 <- phelan_17 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_17 <- lawrenson_17 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)

law17_fp1 <- filter(lawrenson_17, grepl("17:44260644", chrpos))
phe17_fp1 <- filter(phelan_17, grepl("17:44260644", chrpos))

law17_fp1 <- law17_fp1[,c(41,3:7,9,10,12)]
law17_fp1$study <- "Lawrenson (2019)"
```

```
law17_fp1$N.cases <- 7321
law17_fp1$N.controls <- 4083
phe17_fp1 <- phe17_fp1[,c(73,3:7,9:10,12)]
phe17_fp1$study <- "Phelan (2017)"
phe17_fp1$N.cases <- 25509
phe17_fp1$N.controls <- 40941

combined1 <- rbind(law17_fp1, phe17_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_2_17_fp1 <- filter(METAr_2, chromosome == 17)
METAr_2_17_fp1 <- filter(METAr_2_17_fp1, grepl("17:44260644", MarkerName))
METAr_2_17_fp1$study <- "Meta"
METAr_2_17_fp1$N.cases <- law17_fp1$N.cases + phe17_fp1$N.cases
METAr_2_17_fp1$N.controls <- law17_fp1$N.controls + phe17_fp1$N.controls
METAr_2_17_fp1 <- METAr_2_17_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_2_17_fp1, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]

viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 17:44260644", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 17:43956139

law17_fp2 <- filter(lawrenson_17, grepl("17:43956139", chrpos))
phe17_fp2 <- filter(phelan_17, grepl("17:43956139", chrpos))

law17_fp2 <- law17_fp2[,c(41,3:7,9,10,12)]
law17_fp2$study <- "Lawrenson (2019)"
law17_fp2$N.cases <- 7321
law17_fp2$N.controls <- 4083
phe17_fp2 <- phe17_fp2[,c(73,3:7,9:10,12)]
phe17_fp2$study <- "Phelan (2017)"
phe17_fp2$N.cases <- 25509
phe17_fp2$N.controls <- 40941

combined1 <- rbind(law17_fp2, phe17_fp2)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_2_17_fp2 <- filter(METAr_2, chromosome == 17)
METAr_2_17_fp2 <- filter(METAr_2_17_fp2, grepl("17:43956139", MarkerName))
METAr_2_17_fp2$Effect <- -(as.numeric(METAr_2_17_fp2$Effect))

METAr_2_17_fp2$study <- "Meta"
METAr_2_17_fp2$N.cases <- law17_fp2$N.cases + phe17_fp2$N.cases
METAr_2_17_fp2$N.controls <- law17_fp2$N.controls + phe17_fp2$N.controls
METAr_2_17_fp2 <- METAr_2_17_fp2 %>% rename(chrpos=MarkerName, log_OR=Effect,
```

```
                        stderr=StdErr, Effect=Allele2, Baseline=Allele1)

combined2 <- merge(combined1, METAr_2_17_fp2, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]

viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 17:43956139", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 17:44517766

law17_fp3 <- filter(lawrenson_17, grepl("17:44517766", chrpos))
phe17_fp3 <- filter(phelan_17, grepl("17:44517766", chrpos))

law17_fp3 <- law17_fp3[,c(41,3:7,9,10,12)]
law17_fp3$study <- "Lawrenson (2019)"
law17_fp3$N.cases <- 7321
law17_fp3$N.controls <- 4083
phe17_fp3 <- phe17_fp3[,c(73,3:7,9:10,12)]
phe17_fp3$study <- "Phelan (2017)"
phe17_fp3$N.cases <- 25509
phe17_fp3$N.controls <- 40941

combined1 <- rbind(law17_fp3, phe17_fp3)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                        P.value=overall_pvalue, MAF=EAF)

METAr_2_17_fp3 <- filter(METAr_2, chromosome == 17)
METAr_2_17_fp3 <- filter(METAr_2_17_fp3, grepl("17:44517766", MarkerName))
METAr_2_17_fp3$study <- "Meta"
METAr_2_17_fp3$N.cases <- law17_fp3$N.cases + phe17_fp3$N.cases
METAr_2_17_fp3$N.controls <- law17_fp3$N.controls + phe17_fp3$N.controls
METAr_2_17_fp3 <- METAr_2_17_fp3 %>% rename(chrpos=MarkerName, log_OR=Effect,
                        stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_2_17_fp3, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]

viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 17:44517766", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 2:176753806

phelan_2 <- read.table("Summary_chr2.txt", header=TRUE, sep=",")
```

```
lawrenson_2 <- read.table("SummaryResults_Asian_chr2.txt", header=TRUE, sep=",")

phelan_2$chrpos <- paste(phelan_2$Chromosome, phelan_2$Position, sep=":")
lawrenson_2$chrpos <- paste(lawrenson_2$Chromosome, lawrenson_2$Position, sep=":")

phelan_2 <- phelan_2 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_2 <- lawrenson_2 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)

law2_fp1 <- filter(lawrenson_2, grepl("2:176753806", chrpos))
phe2_fp1 <- filter(phelan_2, grepl("2:176753806", chrpos))

law2_fp1 <- law2_fp1[,c(41,3:7,9,10,12)]
law2_fp1$study <- "Lawrenson (2019)"
law2_fp1$N.cases <- 7321
law2_fp1$N.controls <- 4083
phe2_fp1 <- phe2_fp1[,c(73,3:7,9:10,12)]
phe2_fp1$study <- "Phelan (2017)"
phe2_fp1$N.cases <- 25509
phe2_fp1$N.controls <- 40941

combined1 <- rbind(law2_fp1, phe2_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                    P.value=overall_pvalue, MAF=EAF)

METAr_2_2_fp1 <- filter(METAr_2, chromosome == 2)
METAr_2_2_fp1 <- filter(METAr_2_2_fp1, grepl("2:176753806", MarkerName))
METAr_2_2_fp1$study <- "Meta"
METAr_2_2_fp1$N.cases <- law2_fp1$N.cases + phe2_fp1$N.cases
METAr_2_2_fp1$N.controls <- law2_fp1$N.controls + phe2_fp1$N.controls
METAr_2_2_fp1 <- METAr_2_2_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                        stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_2_2_fp1, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]

viz_forest(summary.table, type="study_only",
       variant="rain",
       x_trans_function = exp,
       xlab="OR SNP 2:176753806", annotate_CI = TRUE,
       study_table = study.table,
       table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
       table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 17:46457047

law17_fp4 <- filter(lawrenson_17, grepl("17:46457047", chrpos))
phe17_fp4 <- filter(phelan_17, grepl("17:46457047", chrpos))

law17_fp4 <- law17_fp4[,c(41,3:7,9,10,12)]
law17_fp4$study <- "Lawrenson (2019)"
law17_fp4$N.cases <- 7321
law17_fp4$N.controls <- 4083
phe17_fp4 <- phe17_fp4[,c(73,3:7,9:10,12)]
phe17_fp4$study <- "Phelan (2017)"
phe17_fp4$N.cases <- 25509
```

```
phe17_fp4$N.controls <- 40941

combined1 <- rbind(law17_fp4, phe17_fp4)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_2_17_fp4 <- filter(METAr_2, chromosome == 17)
METAr_2_17_fp4 <- filter(METAr_2_17_fp4, grepl("17:46457047", MarkerName))
METAr_2_17_fp4$Effect <- -(as.numeric(METAr_2_17_fp4$Effect))

METAr_2_17_fp4$study <- "Meta"
METAr_2_17_fp4$N.cases <- law17_fp4$N.cases + phe17_fp4$N.cases
METAr_2_17_fp4$N.controls <- law17_fp4$N.controls + phe17_fp4$N.controls
METAr_2_17_fp4 <- METAr_2_17_fp4 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele2, Baseline=Allele1)

combined2 <- merge(combined1, METAr_2_17_fp4, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]

viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 17:46457047", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 8:129480083

phelan_8 <- read.table("Summary_chr8.txt", header=TRUE, sep=",")
lawrenson_8 <- read.table("SummaryResults_Asian_chr8.txt", header=TRUE, sep=",")

phelan_8$chrpos <- paste(phelan_8$Chromosome, phelan_8$Position, sep=":")
lawrenson_8$chrpos <- paste(lawrenson_8$Chromosome, lawrenson_8$Position, sep=":")

phelan_8 <- phelan_8 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_8 <- lawrenson_8 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
            filter(overall_pvalue > 0 & overall_pvalue < 1)

law8_fp1 <- filter(lawrenson_8, grepl("8:129480083", chrpos))
phe8_fp1 <- filter(phelan_8, grepl("8:129480083", chrpos))
#Two results, using the SNP used by METAL:
law8_fp1 <- filter(law8_fp1, Effect == "G")
phe8_fp1 <- filter(phe8_fp1, Effect == "G")

law8_fp1 <- law8_fp1[,c(41,3:7,9,10,12)]
law8_fp1$study <- "Lawrenson (2019)"
law8_fp1$N.cases <- 7321
law8_fp1$N.controls <- 4083
phe8_fp1 <- phe8_fp1[,c(73,3:7,9:10,12)]
phe8_fp1$study <- "Phelan (2017)"
phe8_fp1$N.cases <- 25509
phe8_fp1$N.controls <- 40941

combined1 <- rbind(law8_fp1, phe8_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
```

```
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_2_8_fp1 <- filter(METAr_2, chromosome == 8)
METAr_2_8_fp1 <- filter(METAr_2_8_fp1, grepl("8:129480083", MarkerName))
METAr_2_8_fp1$study <- "Meta"
METAr_2_8_fp1$N.cases <- law8_fp1$N.cases + phe8_fp1$N.cases
METAr_2_8_fp1$N.controls <- law8_fp1$N.controls + phe8_fp1$N.controls
METAr_2_8_fp1 <- METAr_2_8_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele1, Baseline=Allele2)

combined2 <- merge(combined1, METAr_2_8_fp1, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]

viz_forest(summary.table, type="study_only",
        variant="rain",
        x_trans_function = exp,
        xlab="OR SNP 8:129480083", annotate_CI = TRUE,
        study_table = study.table,
        table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
        table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#SNP 19:17374507

phelan_19 <- read.table("Summary_chr19.txt", header=TRUE, sep=",")
lawrenson_19 <- read.table("SummaryResults_Asian_chr19.txt", header=TRUE, sep=",")

phelan_19$chrpos <- paste(phelan_19$Chromosome, phelan_19$Position, sep=":")
lawrenson_19$chrpos <- paste(lawrenson_19$Chromosome, lawrenson_19$Position, sep=":")

phelan_19 <- phelan_19 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
             filter(overall_pvalue > 0 & overall_pvalue < 1)
lawrenson_19 <- lawrenson_19 %>% filter(EAF > 0.01 & EAF < 0.99) %>% filter(overall_SE > 0) %>%
             filter(overall_pvalue > 0 & overall_pvalue < 1)

law19_fp1 <- filter(lawrenson_19, grepl("19:17374507", chrpos))
phe19_fp1 <- filter(phelan_19, grepl("19:17374507", chrpos))

law19_fp1 <- law19_fp1[,c(41,3:7,9,10,12)]
law19_fp1$study <- "Lawrenson (2019)"
law19_fp1$N.cases <- 7321
law19_fp1$N.controls <- 4083
phe19_fp1 <- phe19_fp1[,c(73,3:7,9:10,12)]
phe19_fp1$study <- "Phelan (2017)"
phe19_fp1$N.cases <- 25509
phe19_fp1$N.controls <- 40941

combined1 <- rbind(law19_fp1, phe19_fp1)
combined1$EAF <- ifelse(combined1$EAF >= 0.5, 1 - combined1$EAF, combined1$EAF)
combined1 <- combined1 %>% rename(log_OR=overall_OR, stderr=overall_SE,
                          P.value=overall_pvalue, MAF=EAF)

METAr_2_19_fp1 <- filter(METAr_2, chromosome == 19)
METAr_2_19_fp1 <- filter(METAr_2_19_fp1, grepl("19:17374507", MarkerName))
METAr_2_19_fp1$Effect <- -(as.numeric(METAr_2_19_fp1$Effect))

METAr_2_19_fp1$study <- "Meta"
METAr_2_19_fp1$N.cases <- law19_fp1$N.cases + phe19_fp1$N.cases
```

```
METAr_2_19_fp1$N.controls <- law19_fp1$N.controls + phe19_fp1$N.controls
METAr_2_19_fp1 <- METAr_2_19_fp1 %>% rename(chrpos=MarkerName, log_OR=Effect,
                                stderr=StdErr, Effect=Allele2, Baseline=Allele1)

combined2 <- merge(combined1, METAr_2_19_fp1, all=TRUE)
ordervec <- c("Lawrenson (2019)", "Phelan (2017)", "Meta")
combined2$study <- factor(combined2$study, levels=ordervec)
combined2 <- with(combined2, combined2[order(study),])
combined2$ncc <- paste(combined2$N.cases, combined2$N.controls, sep="/")
study.table <- combined2[,c(8,3,4,7,12,20)]
summary.table <- combined2[, c("log_OR", "stderr")]

viz_forest(summary.table, type="study_only",
      variant="rain",
      x_trans_function = exp,
      xlab="OR SNP 19:17374507", annotate_CI = TRUE,
      study_table = study.table,
      table_headers=c("Study", "Effect", "Baseline", "P", "MAF", "NCase/NControl"),
      table_layout = matrix(c(1,1,2,3), 2, 2, byrow=TRUE))
#Make significant results match up with input alleles to determine allele swaps:
alleleframe <- METAr_2_sig
phelan_alleles <- rbind(phelan_17,phelan_19,phelan_2,phelan_3,phelan_8,phelan_9)
alleleframe <- merge(alleleframe, phelan_alleles, by.x="MarkerName", by.y="chrpos", all.x=TRUE)
alleleframe <- alleleframe[,c(1:13,16:20,22,23,25)]
alleleframe2 <- alleleframe[,-c(14:15,18:21)]
alleleframe2 <- alleleframe2 %>% filter(is.na(Effect.y) == FALSE) %>%
  filter(Effect.y != "GAG") %>% filter(Effect.y != "GTT")
alleleframe2$Effect.y <- gsub("<INS:ME:ALU>", "G", alleleframe2$Effect.y)
alleleframe2 <- alleleframe2[,c(1:5,15:14,6:13)]


alleleframe2 <- alleleframe2 %>% rename(SNP=MarkerName, `Input Effect Allele` = Effect.y,
                          `Input Base Allele` = Baseline, `Output Effect Allele` = Allele1,
                          `Output Base Allele` = Allele2, Beta = Effect.x, SE=StdErr,
                          `P-Value` = P.value, `Het. I-Squared` = HetISq, `Het. Chi-Squared` = HetChiSq,
                          `Het. DF` = HetDf, `Het. P-Value` = HetPVal, Chromosome = chromosome,
                          Position = Position.x)

alleleframe2 <- arrange(alleleframe2, alleleframe2$Chromosome, alleleframe2$Position)
alleleframe2$`Output Effect Allele` <- toupper(alleleframe2$`Output Effect Allele`)
alleleframe2$`Output Base Allele` <- toupper(alleleframe2$`Output Base Allele`)

alleleframe2_res <- alleleframe2[,c(1:3,6:7,4:5,8:11)]
alleleframe2_het <- alleleframe2[,c(1:3,11:15)]


library(stargazer)
stargazer(alleleframe2_res, type="html", summary=FALSE, rownames=FALSE, out="ovarian_meta_results_table.doc",
      title="Ovarian-Only Meta-Analysis Significant Results")
stargazer(alleleframe2_het, type="html", summary=FALSE, rownames=FALSE,
out="ovarian_meta_results_table_2.doc",
      title="Ovarian-Only Meta-Analysis Significant Results - Heterogeneity Test")
#Make table for variants lost across flowchart steps:
#precalculated values

varflow <- data.frame(Study = c("O'Mara (2018)", "Phelan (2017)", "Lawrenson (2019)", "UK Biobank (Cerv., 2018)",
"UK Biobank (Ovar., 2018)"),
                `Number of Variants at Start` = c(9518973,20594648,20246914,13791467,13791467),
                `Number After Filtering Nonsense Entries` = c(9518964,20000699,20030609,13788358,13788363),
                `Number After Filtering Low Allele Frequency` = c(9518964,10236573,9156543,9664368,9664368))
varflow <- varflow %>% rename(`Number of Variants at Start`=Number.of.Variants.at.Start,
                    `Number After Filtering Nonsense Entries`=Number.After.Filtering.Nonsense.Entries,
```

```
                `Number After Filtering Low Allele Frequency`=Number.After.Filtering.Low.Allele.Frequency)
stargazer(varflow, type="html", summary=FALSE, rownames=FALSE, out="var_flow.doc",
          title="Analysis Variants by Process Flow Step")
```

Bibliography

Abbott, L., Bryant, S., Churchhouse, C., Ganna, A., Howrigan, D., Palmer, D., Neale, B., Walters, R., Carey, C., and The Hail Team. (2018). UK Biobank GWAS Round 2 Results. http://www.nealelab.is/uk-biobank; data retrieved from https://docs.google.com/spreadsheets/d/1kvPoupSzsSFBNSztMzl04xMoSC3Kcx3CrjVf4yBmESU/edit?ts=5b5f17db#gid=227859291

Bacalbasa, N., Balescu, I. C., Diaconu, C., Dima, S., Iliescu, L., Vilcu, M., Filipescu, A., Halmaciu, I., Cretoiu, D., & Brezean, I. (2020). Synchronous Cervical Adenocarcinoma and Ovarian Serous Adenocarcinoma-A Case Report and Literature Review. Medicina (Kaunas, Lithuania), 56(4), 152.

Baxter, J. S., Leavy, O.C., Dryden, N.H. et al. (2018). Capture Hi-C identifies putative target genes at 33 breast cancer risk loci. Nat Commun, 9(1028).

Berger, A. C., Korkut, A., Kanchi, R. S., Hegde, A. M., Lenoir, W., Liu, W., Liu, Y., Fan, H., Shen, H., Ravikumar, V., Rao, A., Schultz, A., Li, X., Sumazin, P., Williams, C., Mestdagh, P., Gunaratne, P. H., Yau, C., Bowlby, R., Robertson, A. G., … Akbani, R. (2018). A comprehensive Pan-Cancer molecular study of gynecologic and breast cancers. Cancer Cell, 33(4), 690-705.

Bhutia, Y. D., Babu, E., Ramachandran, S., Yang, S., Thangaraju, M., & Ganapathy, V. (2016). SLC transporters as a novel class of tumour suppressors: identity, function and molecular mechanisms. The Biochemical journal, 473(9), 1113-1124.

Borenstein, M, Hedges, L. V., Higgins, J.P., and Rothstein, H. R. (2009). "Identifying and Quantifying Heterogeneity." Introduction to Meta-Analysis. Chichester, UK: John Wiley & Sons, Ltd, 2009. 107-125. Web.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. Research synthesis methods, 1(2), 97-111.

Buniello, A., MacArthur, JAL., Cerezo, M., Harris, LW., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, PL., Amode, R., Guillen, JA., Riat, HS., Trevanion, SJ., Hall, P., Junkins, H., Flicek, P., Burdett, T., Hindorff, LA., Cunningham, F,. and Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Research, 47, D1005-D1012.

Bycroft, C., Freeman, C., Petkova, D. et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature, 562, 203-209.

Carvajal-Carmona, L. G., O'Mara, T.A., Painter, J.N, et al. (2015). Candidate locus analysis of the TERT–CLPTM1L cancer risk region on chromosome 5p15 identifies multiple independent variants associated with endometrial cancer risk. Hum Genet, 134(2), 231-245.

Cesaratto, L., Grisard, E., Coan, M., Zandonà, L., De Mattia, E., Poletto, E., Cecchin, E., Puglisi, F., Canzonieri, V., Mucignat, M. T., Zucchetto, A., Stocco, G., Colombatti, A., Nicoloso, M. S., & Spizzo, R. (2016). BNC2 is a putative tumor suppressor gene in high-grade serous ovarian carcinoma and impacts cell survival after oxidative stress. Cell Death and Disease, 7(9), e2374.

Cheng, L., Li, Z., Huang, Y. Z., Zhang, X., Dai, X. Y., Shi, L., Xi, P. W., Wei, J. F., & Ding, Q. (2019). TCDD-Inducible Poly-ADP-Ribose Polymerase (TIPARP), A Novel Therapeutic Target Of Breast Cancer. Cancer management and research, 11, 8991-9004.

Cheng, T. H., Thompson, D. J., O'Mara, T. A., Painter, J. N., Glubb, D. M., Flach, S., Lewis, A., French, J. D., Freeman-Mills, L., Church, D., Gorman, M., Martin, L., National Study of Endometrial Cancer Genetics Group (NSECG), Hodgson, S., Webb, P. M., Australian National Endometrial Cancer Study Group (ANECS), Attia, J., Holliday, E. G., McEvoy, M., Scott, R. J., … Spurdle, A. B. (2016). Five endometrial cancer risk loci identified through genome-wide association analysis. . Nature genetics, 48(6), 667-674.

Constantinou, P., Tischkowitz, M. (2017). Genetics of gynaecological cancers. Best Practice & Research Clinical Obstetrics and Gynaecology, 42, 114-124.

Das, H., Koizumi, T., Sugimoto, T., Chakraborty, S., Ichimura, T., Hasegawa, K., Nishimura, R. (2000). Quantitation of Fas and Fas ligand gene expression in human ovarian, cervical and endometrial carcinomas using real-time quantitative RT-PCR. British Journal of Cancer, 82(10), 1682-1688.

Diaz-Padilla, I., Amir, E., Marsh, S., Liu, G., Mackay, H. (2012). Genetic polymorphisms as predictive and prognostic biomarkers in gynecological cancers: a systematic review. Gynecol Oncol, 124, 354-365.

Dratwa, M., Wysoczańska, B., Łacina, P., Kubik, T., Bogunia-Kubik, K. (2020). TERT-Regulation and Roles in Cancer Formation. Front Immunol, 19(11), 589929.

Evangelou, E., Ioannidis, J. (2013). Meta-analysis methods for genome-wide association studies and beyond. Nat Rev Genet, 14, 379-389.

Fejzo, M. S., Chen, H., Anderson, L., McDermott, M.S.J., Karlan, B., Konecny, G.E., Slamon, D.J. (2021). Analysis in epithelial ovarian cancer identifies KANSL1 as a biomarker and target gene for immune response and HDAC inhibition. Gynecol Oncol, 160(2), 539-546.

Gates, M. A., Wolpin, B. M., Cramer, D. W., Hankinson, S. E., & Tworoger, S. S. (2011). ABO blood group and incidence of epithelial ovarian cancer. . International journal of cancer, 128(2), 482-486.

Goode, E., Chenevix-Trench, G., Song, H., Ramus, S., et al. (2010). A Genome-Wide Association Study Identifies Susceptibility Loci for Ovarian Cancer at 2q31 and 8q24. Nat Genet, 42(10), 874-879.

Guo, Y., Liu, J., Luo, J., You, X., Weng, H., Wang, M., Ouyang, T., Li, X., Liao, X., Wang, M., Lan, Z., Shi, Y., & Chen, S. (2020). Molecular Profiling Reveals Common and Specific Development Processes in Different Types of Gynecologic Cancers. Front Oncol, 10, 584793.

Guolo, A., Varin, C. (2017). Random-effects meta-analysis: the number of studies matters. Statistical Methods in Medical Research, 26(3), 1500-1518.

Herrinton, L. J., Voigt, L.F., Weiss, N.S., Beresford, S.A.A., & Wingo, P.A. (2001). Risk Factors for Synchronous Primary Endometrial and Ovarian Cancers. Annals of Epidemiology, 11(8), 529-533.

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, MJ., Welch, VA. (2021). Cochrane Handbook for Systematic Reviews of Interventions. In. Cochrane. https://training.cochrane.org/handbook/current/chapter-10

Howrigan, D., Abbott, L., Churchhouse, C., Palmer, D. (2017). Details and Considerations of the UK Biobank GWAS. Blog Post. http://www.nealelab.is/blog/2017/9/11/details-and-considerations-of-the-uk-biobank-gwas

Ioannidis, J. P., Patsopoulos, N.A., Evangelou, E. (2007). Heterogeneity in Meta-Analyses of Genome-Wide Association Investigations. PLoS ONE, 2(9), e841.

Jacob, F., Hitchins, M. P., Fedier, A., Brennan, K., Nixdorf, S., Hacker, N. F., Ward, R., & Heinzelmann-Schwarz, V. A. (2014). Expression of GBGT1 is epigenetically regulated by DNA methylation in ovarian cancer cells. BMC Molecular Biology, 15(24).

Kent, W., Sugnet, CW., Furey, TS., Roskin, KM., Pringle, TH., Zahler, AM., Haussler, D. (2002). The human genome browser at UCSC. Genome Res, 12(6), 996-1006. https://genome.ucsc.edu/cgi-bin/hgLiftOver

Kim, J., Hong, C.M., Park, S.M., Shin, D.H., Kim, J.Y., Kwon, S.M., Kim, J.H., Kim, C.D., Lim, D.S., Lee, D. (2018). SURF4 has oncogenic potential in NIH3T3 cells. . Biochem Biophys Res Commun, 502(1), 43-47.

Kostovska, I. M., Wang, J., Bogdanova, N., Schürmann, P., Bhuju, S., Geffers, R., Dürst, M., Liebrich, C., Klapdor, R., Christiansen, H., Park-Simon, T., Hillemanns, P., Plaseska-Karanfilska, D., Dörk, T. (2016). Rare ATAD5 missense variants in breast and ovarian cancer patients. Cancer Letters, 376(1), 173-177.

Krishnamurthy, N., Ngam, C., Berdis, A. et al. (2011). The exonuclease activity of hPMC2 is required for transcriptional regulation of the QR gene and repair of estrogen-induced abasic sites. Oncogene, 30, 4731–4739.

Kuchenbaecker, K. B., Ramus, S. J., Tyrer, J., Lee, A., Shen, H. C., Beesley, J., Lawrenson, K., McGuffog, L., Healey, S., Lee, J. M., Spindler, T. J., Lin, Y. G., Pejovic, T., Bean, Y., Li, Q., Coetzee, S., Hazelett, D., Miron, A., Southey, M., Terry, M. B., … Consortium of Investigators of Modifiers of BRCA1 and BRCA2. (2015). Identification of six new susceptibility loci for invasive epithelial ovarian cancer. Nature genetics, 47(2), 164-171.

Lan, L., Cao, H., Chi, W., Meng, W., Zhao, L., Cui, W., Wang, B. (2020). Aberrant DNA hypermethylation-silenced LINC00886 gene accelerates malignant progression of laryngeal carcinoma. Pathology Research and Practice, 216(4), 152877.

Lawrenson, K., Kar, S., McCue, K. et al. (2016). Functional mechanisms underlying pleiotropic risk alleles at the 19p13.1 breast–ovarian cancer susceptibility locus. Nat Commun, 7(12675).

Lawrenson, K., Song, F., Hazelett, D. J., Kar, S. P., Tyrer, J., Phelan, C. M., Corona, R. I., Rodríguez-Malavé, N. I., Seo, J. H., Adler, E., Coetzee, S. G., Segato, F., Fonseca, M., Amos, C. I., Carney, M. E., Chenevix-Trench, G., Choi, J., Doherty, J. A., Jia, W., Jin, G. J., … Goodman, M. T. (2019). Genome-wide association studies identify susceptibility loci for epithelial ovarian cancer in east Asian women. Gynecol Oncol, 153(2), 343-355.

Ledford, L., Lockwood, S. (2019). Scope and epidemiology of gynecologic cancers: an overview. Seminars in Oncology Nursing, 35(2), 147-150.

Leo, P. J., Madeleine, M. M., Wang, S., Schwartz, S. M., Newell, F., Pettersson-Kymmer, U., Hemminki, K., Hallmans, G., Tiews, S., Steinberg, W., Rader, J. S., Castro, F., Safaeian, M., Franco, E. L., Coutlée, F., Ohlsson, C., Cortes, A., Marshall, M., Mukhopadhyay, P., Cremin, K., … Brown, M. A. (2017). Defining the genetic susceptibility to cervical neoplasia—A genome-wide association study. PLoS Genet, 13(8), e1006866.

Li, H., Che, J., Jiang, M., et al. (2020). CLPTM1L induces estrogen receptor β signaling-mediated radioresistance in non-small cell lung cancer cells. Cell Commun Signal, 18(152).

Li, X., Zhu, X. (2017). Cross-Phenotype Association Analysis Using Summary Statistics from GWAS. Methods in molecular biology, 1666, 455-467.

Lichou, F., Trynka, G. (2020). Functional studies of GWAS variants are gaining momentum. Nat Commun, 11.

Lin, E., Tong, T., Chen, Y., Wang, Y. (2020). Fixed-effects model: the most convincing moel for meta-analysis with few studies. arXiv:2002.04211 [Preprint] 11 Feb 2020.

Liu, Y., Wang, F., Liu, Y., Yao, Y., Lv, X., Dong, B., Li, J., Ren, S., Yao, Y., & Xu, Y. (2016). RNF135, RING finger protein, promotes the proliferation of human glioblastoma cells in vivo and in vitro via the ERK pathway. Scientific reports, 6, 20642.

Madan, E., Pelham, C.J., Nagane, M. et al. (2019). Flower isoforms promote competitive growth in cancer. Nature, 572, 260-264.

Marcotte, R., Brown, K. R., Suarez, F., Sayad, A., Karamboulas, K., Krzyzanowski, P. M., Sircoulomb, F., Medrano, M., Fedyshyn, Y., Koh, J., van Dyk, D., Fedyshyn, B., Luhova, M., Brito, G. C., Vizeacoumar, F. J., Vizeacoumar, F. S., Datti, A., Kasimer, D., Buzina, A., Mero, P., … Moffat, J. (2012). Essential gene profiles in breast, pancreatic, and ovarian cancer cells. Cancer Discovery, 2(2), 172-189.

Masuda, T., Low, SK., Akiyama, M., et al. (2020). GWAS of five gynecologic diseases and cross-trait analysis in Japanese. Eur J Hum Genet, 28, 95-107.

Nakamura, K., Reid, B.M., Sellers, T.A., Karreth, F.A. (July 2020). Abstract B34: LINC00886, a risk locus-associated long noncoding RNA, promotes ovarian cancer progression. Clinical Cancer Research,

O'Mara, T. A., Glubb, D.M., Amant, F. et al. (2018). Identification of nine new susceptibility loci for endometrial cancer. Nat Commun, 9(3166), 1-12.

Pan, Y., Pan, Y., Cheng, Y., Yang, F., Yao, Z., & Wang, O. Cell & bioscience, 8, 7. https://doi.org/10.1186/s13578-018-0207-5. (2018). Knockdown of LncRNA MAPT-AS1 inhibites proliferation and migration and sensitizes cancer cells to paclitaxel by regulating MAPT expression in ER-negative breast cancers. Cell & Bioscience, 8(7).

Pereira, T. V., Patsopoulos, N. A., Salanti, G., & Ioannidis, J. P. (2009). Discovery properties of genome-wide association signals from cumulatively combined data sets. American journal of epidemiology, 170(10), 1197-1206.

Permuth, J., Pirie, A., Chen, Y.A., Lin, H. et al. (2016). Exome genotyping arrays to identify rare and low frequency variants associated with epithelial ovarian cancer risk., 25(16), 3600-3612.

Phelan, C. M., Kuchenbaecker, K. B., Tyrer, J. P., Kar, S. P., Lawrenson, K., Winham, S. J., Dennis, J., Pirie, A., Riggan, M. J., Chornokur, G., Earp, M. A., Lyra, P. C., Jr, Lee, J. M., Coetzee, S., Beesley, J., McGuffog, L., Soucy, P., Dicks, E., Lee, A., Barrowdale, D., … Pharoah, P. (2017). Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. Nat Jenet, 49(5), 680-691.

Ramakrishna, M., Williams, L. H., Boyle, S. E., Bearfoot, J. L., Sridhar, A., Speed, T. P., Gorringe, K. L., & Campbell, I. G. (2010). Identification of candidate growth promoting genes in ovarian cancer through integrated copy number and expression analysis. PLoS ONE, 5(4), e9983.

Rashkin, S. R., Graff, R.E., Kachuri, L. et al. (2020). Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. Nat Commun, 11(4423), 1-14.

Sankaranarayanan, R., Ferlay, J. (2006). Worldwide burden of gynaecological cancer: The size of the problem. Best Practice & Research Clinical Obstetrics and Gynaecology, 20(2), 207-225.

Schmandt, R. E., Iglesias, D. A., Co, N. N., & Lu, K. H. (2011). Understanding obesity and endometrial cancer risk: opportunities for prevention. American journal of obstetrics and gynecology, 205(6), 518-525.

Shukla, V., Mahadev, R., Zhang, H., Beers, J., Wangsa, Dar., Wangsa, Dan., Buishand, F.O., Wang, Y., Yu, Z., Stevenson, H.S.,... Schrump, D. (2017). ASXL3 is a novel pluripotency factor in human respiratory epithelial cells and a potential therapeutic target in small cell lung cancer. Cancer Res, 77(22), 6267-6281.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., et al. (2015). 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(3), e1001779.

Takeda, T., Banno, K., Yanokura, M., Anko, M., Kobayashi, A., Sera, A., Takahashi, T., Adachi, M., Kobayashi, Y., Hayashi, S., Nomura, H., Hirasawa, A., Tominaga, E., & Aoki, D. (2018). Synchronous endometrial and ovarian cancer in Lynch syndrome with a MSH2 germline mutation: A case report. Molecular and clinical oncology, 9(5), 479-484.

Team., R. C. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. American journal of human genetics, 101(1), 5-22.

Wan, L., Wang., Y., Zhang, Z., et al. (2021). Elevated TEFM expression promotes growth and metastasis through activation of ROS/ERK signaling in hepatocellular carcinoma. Cell Death Dis, 12(325).

Wang, M., Li, L., Liu, J., & Wang, J. (2018). A gene interaction network-based method to measure the common and heterogeneous mechanisms of gynecological cancer. Molecular medicine reports, 18(1), 230-242.

Wang-Hong, X., Zheng, W., Xiang, Y., Shu, X. (2011). ABO blood type is associated with endometrial cancer risk in Chinese women. Clin J Cancer, 30(11), 766-771.

Willer, C. J., Li, Y., & Abecasis, G. R. . (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics, 26(17), 2190-2191.

Yu, Y., Xin, F., Lu, D., Ge, L., Zhai, C., Shen, X. (2020). Weighted gene coexpression network analysis identifies critical genes in different subtypes of acute myeloid leukaemia. Biotechnology & Biotechnology Equipment, 34(1).

Zhang, J., Zhang, Y., Zhang, Z. (2018). Association of rs2279744 and rs117039649 promoter polymorphism with the risk of gynecological cancer. Medicine, 97(2), e9554.

Zhou, J. X., Yang, X., Ning, S., Wang, L., Wang, K., Zhang, Y., Yuan, F., Li, F., Zhuo, D. D., Tang, L., & Zhuo, D. (2017). Identification of KANSARL as the first cancer predisposition

fusion gene specific to the population of European ancestry origin. Oncotarget, 8(31), 50594-50607.

Zhu, Z., Anttila, V., Smoller, JW., and Lee, PH. (2018). Statistical power and utility of meta-analysis methods for cross-phenotype genome-wide association studies. PLoS ONE, 13(3), e0193256.

Zorn, K., Bonome, T., Gangi, L., Chandramouli, G., Awtrey, C., Gardner, G., Barrett, J., Boyd, J., Birrer, M. (2005). Gene Expression Profiles of Serous, Endometrioid, and Clear Cell Subtypes of Ovarian and Endometrial Cancer. Clin Cancer Res, 11(18), 6422-6430.