main

April 20, 2022

1 SMR manhattan plot

[1]:
```
suppressMessages({
    library(dplyr)
    library(ggtext)
    library(ggplot2)
})
```

1.1 Functions

[2]:
```
save_plots <- function(p, fn){
    for(ext in c('.png', '.pdf', '.svg')){
        ggsave(paste0(fn,ext), p, width=30, height=15, units="cm")
    }
}
```

1.2 Prepare data

[3]:
```
annot_file = "/ceph/projects/v4_phase3_paper/inputs/counts/gene_annotation/_m/
↪gene_annotation.tsv"
gene_annot <- data.table::fread(annot_file) %>% janitor::clean_names()
gene_annot %>% head(2)
```

A data.table: 2 × 11

feature_id <chr>	length <int>	gencode_id <chr>	ensembl_id <chr>	gene_type <chr>
ENSG00000223972.5	1735	ENSG00000223972.5	ENSG00000223972	transcribed_u
ENSG00000227232.5	1351	ENSG00000227232.5	ENSG00000227232	unprocessed_p

1.2.1 Preparing SMR data

[4]:
```
smr_file = "../../_m/eqtl_genes.eqtl_p1e-04.gwas_p5e-08.csv"
qsmr_data_load <- data.table::fread(smr_file) %>% janitor::clean_names() %>%
    filter(fdr < 0.05, p_heidi > 0.01) %>% select(probe_id, probe_bp, fdr) %>%␣
↪distinct

smr_data_load <- data.table::fread(smr_file) %>% janitor::clean_names() %>%
    inner_join(gene_annot, by=c("probe_id"="feature_id")) %>%
```

```
    select(probe_id, probe_bp, probe_chr, p_smr, p_heidi, top_snp, symbol) %>%␣
↪distinct %>%
    rename(chr=probe_chr)

data_cum <- smr_data_load %>% group_by(chr) %>% summarise(max_bp =␣
↪max(probe_bp)) %>%
    mutate(bp_add = lag(cumsum(as.numeric(max_bp)), default=0)) %>% select(chr,␣
↪bp_add)
smr_data <- smr_data_load %>% inner_join(data_cum, by="chr") %>%
    mutate(bp_cum = probe_bp + bp_add) %>% left_join(qsmr_data_load,␣
↪by=c("probe_id", "probe_bp")) %>%
    mutate(fdr = tidyr::replace_na(fdr, 1))
smr_data %>% head(2)
```

A data.table: 2 × 10

probe_id <chr>	probe_bp <int>	chr <int>	p_smr <dbl>	p_heidi <dbl>	top_snp <chr>	symbo <chr>
ENSG00000118292.8	150268200	1	0.38392490	0.3198645	rs35413284	C1orf5
ENSG00000118298.10	150257158	1	0.02700349	0.4391561	rs12401456	CA14

1.2.2 SMR caudate, SZ PGC3

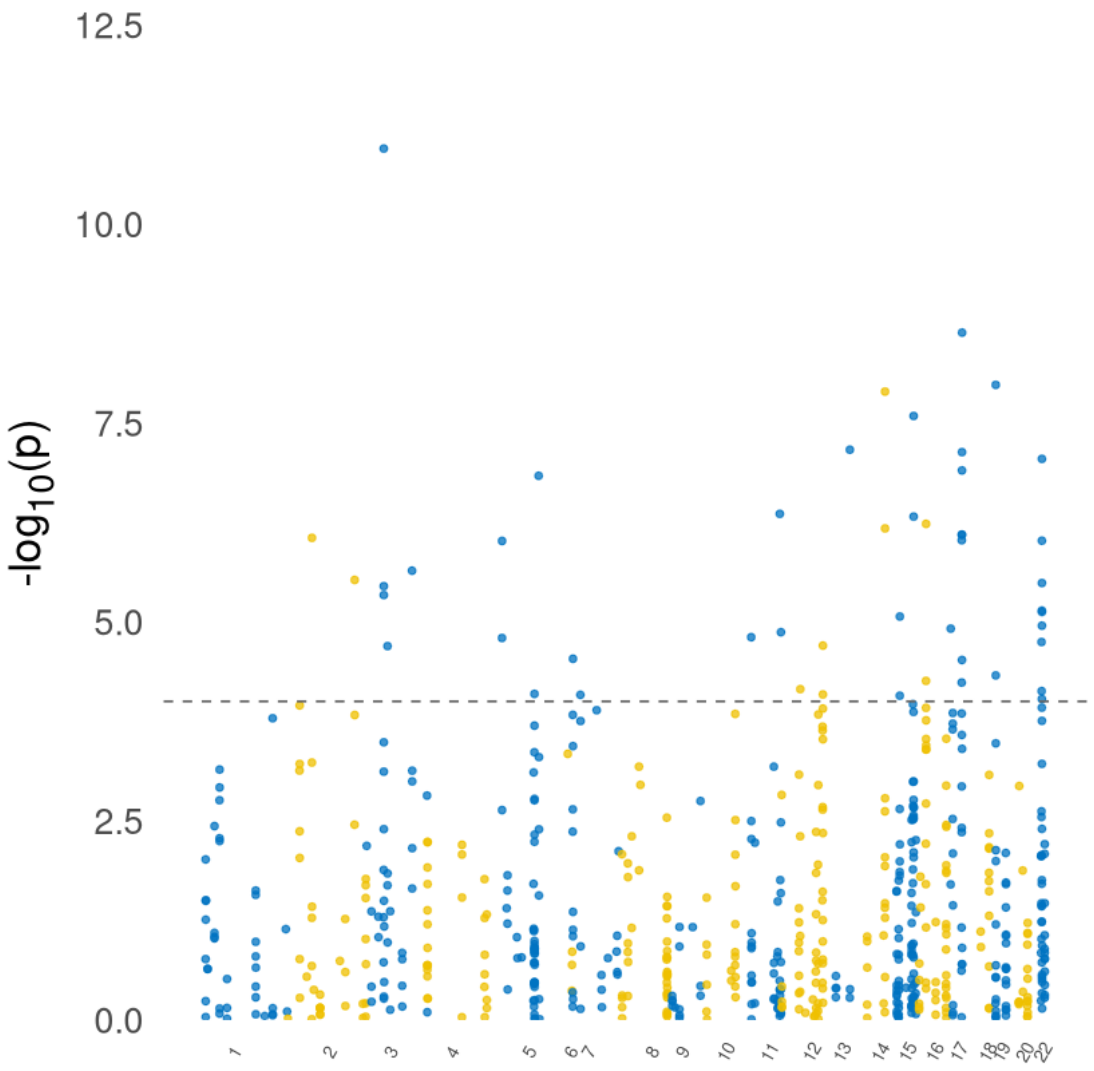
1.2.3 Plotting settings

[5]:
```
axis_set <- smr_data %>% group_by(chr) %>% summarize(center = mean(bp_cum))
ylim <- smr_data %>% filter(p_smr == min(p_smr)) %>%
    mutate(ylim  = abs(floor(log10(p_smr))) + 2) %>% pull(ylim)
ylim = ylim[1]
sig <- 1e-4
```

1.2.4 Plot

[6]:
```
manhplot1 <- ggplot(smr_data, aes(x=bp_cum, y=-log10(p_smr),␣
↪size=-log10(p_smr))) +
    geom_point(alpha=0.75, size=1.25, aes(color=forcats::as_factor(chr))) +
    geom_hline(yintercept = -log10(sig), color = "grey40", linetype = "dashed")␣
↪+
    scale_x_continuous(label = axis_set$chr, breaks = axis_set$center) +
    scale_y_continuous(expand = c(0,0), limits = c(0, ylim)) +
    scale_color_manual(values = rep(ggpubr::get_palette("jco", 2),␣
↪unique(length(axis_set$chr)))) +
    labs(x = NULL, y = "-log<sub>10</sub>(p)") +
    theme_minimal(base_size=20) +
    theme(
        legend.position="none", panel.border=element_blank(),
        panel.grid.major=element_blank(), panel.grid.minor=element_blank(),
        axis.title.y=element_markdown(), axis.text.x=element_text(angle=60,␣
↪size=8, vjust=0.5)
```

```
    )
manhplot1
```



[7]:
```
save_plots(manhplot1, 'caudate_smr_manhattanplot')
```

[ ]:
```
```