README.txt in the RWAS_weights folder.

This folder contains weight files that can be used together with GWAS summary statistics to conduct Regulome Wide Association-Studies (RWAS) using the FUSION software package.

Each weight file corresponds to a single accessible element in the FUSION format. Instructions on how to use FUSION for RWAS can be found at http://gusevlab.org/projects/fusion/. RWAS is conducted in the same manner as TWAS using these RWAS weights.

The weight files were generated using a pan-cancer peak set. A BED file containing pan-cancer peaks can be found in the peaks_hg19 folder along with other BED files containing cancer-type specific peaks. Cancer-type specific peak files can be used to restrict the RWAS analysis to regulatory elements active in specific cancer types.

The data generation and analyses are described in detail in the methods section of Grishin D. and Gusev A. Allelic imbalance of chromatin accessibility in cancer identifies likely causal risk variants and their mechanisms. (2022).

—----------

README.txt in the stratAS_results folder.

This folder contains the results of analyses of allele-specific chromatin accessibility for 23 cancer types + pan-cancer using the stratAS software. The analyses identified allelically imbalanced genomic regions from 406 cancer ATAC-Seq samples. The files are named after each cancer type (or “pancancer” for the pan-cancer analysis) and contain the following columns:

CHR - Chromosome
POS - Position of test SNP
RSID - ID of test SNP
P0 - Start of gene/peak
P1 - End of gene/peak
NAME - Name of gene/peak
CENTER - Center position of peak (or TSS for gene)
N.HET - Number of heterozygous individuals tested
N.READS - Number of reads tested in total
ALL.AF - Allelic fraction estimate from beta binomial test across both conditions
ALL.BBINOM.P - Beta-binomial test for imbalance across both conditions
C0.AF - Allelic fraction estimate from condition 0
C0.BBINOM.P - Beta-binomial test for imbalance in condition 0
C1.AF - Allelic fraction estimate from condition 1
C1.BBINOM.P - Beta-binomial test for imbalance in condition 1
DIFF.BBINOM.P - Beta-binomial test for difference between conditions

IND.C0 - Number of each condition 0 individual included in this test (comma separated)
IND.C0.COUNT.REF - condition 0 REF allele counts of each individual included in this test (comma separated)
IND.C0.COUNT.ALT - condition 0 ALT allele counts of each individual included in this test (comma separated)
IND.C1 - Number of each condition 1 individual included in this test (comma separated)
IND.C1.COUNT.REF - condition 1 REF allele counts of each individual included in this test (comma separated)
IND.C1.COUNT.ALT - condition 1 ALT allele counts of each individual included in this test (comma separated)

In our analysis, samples that belong to a cancer type that is analyzed for allelic imbalance are assigned to Condition 1 (C1) while samples belonging to the remaining 22 cancer types are assigned to Condition 0 (C0). For the pan-cancer analysis, samples from all 23 cancer types are assigned to C1. *P*-values denoting the significance of allelic imbalance in a given cancer type (or pan-cancer analysis) are listed in the column C1.BBINOM.P. *P*-values denoting the significance of differential allelic imbalance between a given cancer type and the remaining 22 cancer types are listed in column DIFF.BBINOM.P (set to NA for pan-cancer analysis since all samples are assigned to C1).

More information about stratAS and its output format can be found at https://github.com/gusevlab/stratAS.

The conducted analyses are described in detail in the methods section of Grishin D. and Gusev A. Allelic imbalance of chromatin accessibility in cancer identifies likely causal risk variants and their mechanisms. (2022).