[RT #200617] Effect / ALT alleles for results returned via Gtex webporal

Ayellet Segre via RT <gtex-help@broadinstitute.org>
Wed 29/03/2017 16:19

To: Philip Haycock <philip.haycock@bristol.ac.uk>

Hi Philip,

I checked the SNPs you said were missing from the reference table, and I DO see them in the table. It might be that for some reason you did not download the table in its entirety?
I would recommend re-downloading it. The file should have 11959406 lines including the header (files size: 143079527 bytes).

GTEx_Analysis_2015-01-12_OMNI_2.5M_5M_450Indiv_chr1-22+X_genot_imput_info04_maf01_HWEp1E6_variant_id_lookup.txt.gz

Best,
Ayellet

On Wed Mar 08 10:30:03 2017, philip.haycock@bristol.ac.uk wrote:
> Dear Tim
>
> I am trying to work out the effect allele for the results returned
> using
> the Gtex webporal, ie which allele corresponds to the report effect
> size.
> for example, for the gene BIRC5, the first row of the results is.
> Which
> allele corresponds to the reported effect size?
>
>
> Gencode Id
> Gene Symbol
> SNP
> P-Value
> Effect Size
> Tissue
> Actions
> ENSG00000089685.10 BIRC5 rs2661686dbSNP
> <http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=2661686>
> 5.7e-14
> -0.37 Testis eQTL box plot, IGV eQTL Browser, Multi-tissue eQTL Plot
>
> many thanks
>
> Philip
>
> On 19 May 2015 at 15:39, PC Haycock <philip.haycock@bristol.ac.uk>
> wrote:
>
> > Hi Tim
> >
> > Thanks for the additional information.
> >
> > best
> > Philip
> >
> > On 19 May 2015 at 15:29, Tim Sullivan via RT <gtex-
> > help@broadinstitute.org
> > > wrote:

> >
> >> Hi Philip,
> >>
> >> This paragraph from that paper best describes the strand orientation
> >> used
> >> for our SNPs, the "plus genomic strand":
> >> "In all human reference chromosomes, as for other eukaryotes [3],
> >> the plus
> >> (+) strand is defined as the strand with its 5′ end at the tip of
> >> the
> >> short
> >> arm 4 and 5 (Genome Reference Consortium, personal communication,
> >> March
> >> 27,
> >> 2012). SNP alleles reported on the same strand as the (+) strand are
> >> called
> >> ‘plus’ alleles and those on the (−) strand are called ‘minus’
> >> alleles.
> >> Providing SNP alleles on the plus genomic strand is the convention
> >> in
> >> publicly available SNP datasets such as the HapMap (www.hapmap.org)
> >> and
> >> 1000 Genomes Projects ([http://www.1000genomes.org).]www.1000genomes.org)."
> >>
> >> Tim Sullivan
> >> Associate Computational Biologist
> >>
> >> On Tue, May 19, 2015 at 4:21 AM, philip.haycock@bristol.ac.uk via RT
> >> <
> >> gtex-help@broadinstitute.org> wrote:
> >>
> >> >
> >> > <URL: https://rt.broadinstitute.org/Ticket/Display.html?id=175850
> >> > >
> >> >
> >> > Dear Tim
> >> >
> >> > Could you clarify what you mean by "The reference strand is always
> >> > the
> >> > forward or positive strand". As far as I know, forward and
> >> > positive
> >> strands
> >> > are not necessarily equivalent.
> >> > http://www.ncbi.nlm.nih.gov/pubmed/22658725
> >> >
> >> > Many thanks
> >> >
> >> > Philip
> >> >
> >> >
> >> > On 18 May 2015 at 21:02, PC Haycock <philip.haycock@bristol.ac.uk>
> >> wrote:
> >> >
> >> > > That sounds great. Thanks for the quick reply.
> >> > >
> >> > > best
> >> > > Philip
> >> > >
> >> > > On 18 May 2015 at 21:00, Tim Sullivan via RT <
> >> > gtex-help@broadinstitute.org
> >> > > > wrote:

> >> > >
> >> > >> Hi Philip,
> >> > >>
> >> > >> We do not currently have a file available which annotates the
> >> > >> effect
> >> > >> allele
> >> > >> frequency for all of our SNPs. In June, we are planning on
> >> releasing a
> >> > >> batch search/calculation service on our portal, which will
> >> > >> return the
> >> > >> effect allele frequencies for the SNPs searched, and in
> >> > >> September we
> >> are
> >> > >> planning on releasing an updated eQTL dataset, which will have
> >> allele
> >> > >> frequencies annotated for all SNPs.
> >> > >>
> >> > >> The reference strand is always the forward or positive strand
> >> > >> -- the
> >> > > same
> >> > >> strand used by the hg19 genome reference.
> >> > >>
> >> > >> Tim Sullivan
> >> > >> Associate Computational Biologist
> >> > >>
> >> > >> On Mon, May 18, 2015 at 3:49 PM, philip.haycock@bristol.ac.uk
> >> > >> via
> >> RT <
> >> > >> gtex-help@broadinstitute.org> wrote:
> >> > >>
> >> > >>>
> >> > >>> <URL:
> >> > >>> https://rt.broadinstitute.org/Ticket/Display.html?id=175850
> >> > >
> >> > >>>
> >> > >>> Hi Tim
> >> > >>>
> >> > >>> Thanks for that explanation. It was very clear.
> >> > >>>
> >> > >>> I noticed that your summary eQTL data does not include the
> >> > >>> effect
> >> > > allele
> >> > >>> frequency. Would it be possible to obtain a file indicating
> >> > >>> the
> >> effect
> >> > >>> allele frequency for each SNP? Could you also confirm the
> >> > >>> reference
> >> > >>> strand
> >> > >>> used in your study?
> >> > >>>
> >> > >>> Many thanks for your assistance.
> >> > >>>
> >> > >>> bw
> >> > >>> Philip
> >> > >>>
> >> > >>>
> >> > >>> On 18 May 2015 at 19:20, Tim Sullivan via RT <
> >> > >> gtex-help@broadinstitute.org
> >> > >>>>
> >> > >> wrote:
> >> > >>>

> >> > >> > > Hi Phillip,
> >> > >> > >
> >> > >> > > 1. The effect/modeled allele does not necessarily
> >> > >> > > correspond to
> >> the
> >> > >> > minor
> >> > >> > > allele in GTEx.
> >> > >> > >
> >> > >> > > 2. Our definitions of REF and ALT are always with respect
> >> > >> > > to the
> >> > >> > > hg19/GRCh37 genome reference. The effect size we report is
> >> > >> > > with
> >> > >> respect
> >> > >> > to
> >> > >> > > an increase or decrease in expression among individuals
> >> > >> > > with ALT
> >> > >> > genotypes.
> >> > >> > >
> >> > >> > > An example from our portal might be the best way to
> >> > >> > > demonstrate
> >> > >> this:
> >> > >> > >
> >> > >> > >
> >> > >> >
> >> > >> >
> >> > > http://www.gtexportal.org/home/eqtls/calc?tissueName=
> >> Artery_Tibial&geneId=ENSG00000005075.11&snpId=rs4403344
> >> > >> > >
> >> > >> > > This SNP, rs4403344, in hg19 and in the file
> >> > >> GTEx_genot_imputed_variants_
> >> > >> > > info4_maf05_CR95_CHR_POSb37_ID_REF_ALT.txt has a REF allele
> >> > >> > > of
> >> C and
> >> > >> an
> >> > >> > ALT
> >> > >> > > allele of T.
> >> > >> > >
> >> > >> > > This means for the regression, "Homo Ref" refers to C/C
> >> individuals,
> >> > >> > "Het"
> >> > >> > > refers to C/T, and "Homo Alt" refers to T/T individuals.
> >> > >> > > This
> >> > >> regression
> >> > >> > > has a positive effect size because the expression increases
> >> > >> > > from
> >> C/C
> >> > >> to
> >> > >> > C/T
> >> > >> > > to T/T individuals.
> >> > >> > >
> >> > >> > > Tim Sullivan
> >> > >> > > Associate Computational Biologist
> >> > >> > >
> >> > >> > >
> >> > >> > >
> >> > >> > >
> >> > >> > > On Mon, May 18, 2015 at 2:01 PM,
> >> > >> > > philip.haycock@bristol.ac.uk
> >> via
> >> > >> RT
> >> > >> > <

> >> > >> > > gtex-help@broadinstitute.org> wrote:
> >> > >> > >
> >> > >> > > >
> >> > >> > > > Mon May 18 14:01:09 2015: Request 175850 was acted upon.
> >> > >> > > > Transaction: Ticket created by
> >> > >> > > > philip.haycock@bristol.ac.uk
> >> > >> > > > Queue: gtex-help
> >> > >> > > > Subject: Definition of REF and ALT alleles
> >> > >> > > > Owner: Nobody
> >> > >> > > > Requestors: philip.haycock@bristol.ac.uk
> >> > >> > > > Status: new
> >> > >> > > > Ticket <URL:
> >> > >> > > > https://rt.broadinstitute.org/Ticket/Display.html?id=175850
> >> > >> > > > >
> >> > >> > > >
> >> > >> > > >
> >> > >> > > > Hi there
> >> > >> > > >
> >> > >> > > > I am using summary eQTL data from GTEx. Could you confirm
> >> > >> > > > that
> >> the
> >> > >> > > > effect/modelled allele corresponds to the minor allele in
> >> > >> > > > GTEx?
> >> > >> Also,
> >> > >> > > > could
> >> > >> > > > you confirm that REF and ALT refer to the effect alleles
> >> > >> > > > and
> >> > >> non-effect
> >> > >> > > > alleles respectively, in the file
> >> > >> > > >
> >> > >> GTEx_genot_imputed_variants_info4_maf05_CR95_CHR_POSb37_
> >> ID_REF_ALT.txt
> >> > >> > > ?
> >> > >> > > >
> >> > >> > > > Many thanks
> >> > >> > > > Philip
> >> > >> > > >
> >> > >> > > >
> >> > >> > > >
> >> > >> > >
> >> > >> > >
> >> > >> > >
> >> > >> --
> >> > >> Philip Haycock, PhD
> >> > >> Research Associate
> >> > >> MRC Integrative Epidemiology Unit
> >> > >>
> >> > >> University of Bristol
> >> > >>
> >> > >> Oakfield House, Oakfield Grove
> >> > >>
> >> > >> Bristol, BS8 2BN, UK
> >> > >>
> >> > >> T +44 (0)117 3310089
> >> > >>
> >> > >>
> >> > >>
> >> > >>

> >> > > --
> >> > > Philip Haycock, PhD
> >> > > Research Associate
> >> > > MRC Integrative Epidemiology Unit
> >> > >
> >> > > University of Bristol
> >> > >
> >> > > Oakfield House, Oakfield Grove
> >> > >
> >> > > Bristol, BS8 2BN, UK
> >> > >
> >> > > T +44 (0)117 3310089 <+44%20117%20331%200089>
> >> > >
> >> >
> >> >
> >> >
> >> > --
> >> > Philip Haycock, PhD
> >> > Research Associate
> >> > MRC Integrative Epidemiology Unit
> >> >
> >> > University of Bristol
> >> >
> >> > Oakfield House, Oakfield Grove
> >> >
> >> > Bristol, BS8 2BN, UK
> >> >
> >> > T +44 (0)117 3310089 <+44%20117%20331%200089>
> >> >
> >> >
> >>
> >>
> >
> >
> > --
> > Philip Haycock, PhD
> > Research Associate
> > MRC Integrative Epidemiology Unit
> >
> > University of Bristol
> >
> > Oakfield House, Oakfield Grove
> >
> > Bristol, BS8 2BN, UK
> >
> > T +44 (0)117 3310089 <+44%20117%20331%200089>
> >