**Genetic regulation of RNA splicing in human pancreatic islets**

Goutham Atla, Silvia Bonas-Guarch, Mirabai Cuenca-Ardura, Anthony Beucher, Daniel J.M. Crouch, Javier Garcia-Hurtado, Ignasi Moran, the T2DSystems consortium, Manuel Irimia, Rashmi B Prasad, Anna L. Gloyn, Lorella Marselli, Mara Suleiman, Thierry Berney, Eelco J P de Koning, Julie Kerr-Conte, Francois Pattou, John A. Todd, Lorenzo Piemonti, Jorge Ferrer

Contact:

Sílvia Bonàs-Guarch ([silvia.bonas@crg.eu](mailto:silvia.bonas@crg.eu))

Goutham Atla ([goutham.atla@gmail.com](mailto:goutham.atla@gmail.com))

Here we describe all the summary statistics generated in Atla, A. et al. 2021 (currently in Biorxiv, <https://www.biorxiv.org/content/10.1101/2021.11.11.468254v1>), mapping gene expression and alternative splicing QTLs (eQTLs and sQTLs, respectively) in 399 human pancreatic islet samples.

**Files:**

1. Lead variant-to-phenotype associations identified using permutations. Contains data for all phenotypes (eGene/Junction) with an eQTL or sQTL effect at FDR 1%.

Significant lead sQTLs: *Significant\_sQTLs.txt.gz*

Significant lead eQTLs: *Significant\_sQTLs.txt.gz*

Headers:

**Gene:** gene symbol ID.

**Junction (only applicable to sQTLs)**: chromosome location (hg19 genome build) of the splicing junction, the intron cluster that belongs to, and the ENSEMBL ID of the corresponding gene (chr:st:en:clu\_XX\_NA\_ENSGXXXXXXX).

**Lead\_SNP**: lead QTL variant association for a given gene/splicing junction.

\*Variant id is given as chr:position (hg19 genome build):Ref(allele):Alt(allele).

**Slope:** beta for the association between expression/splicing junction usage and SNP dosage. The direction of the effect needs to be interpreted for the alternate allele.

**Nominal\_pvalue**:nominal p-value for the association.

**Permutation\_pvalue**: QTLtools adjusted p-value for the number of variants tested in cis given by the fitted beta distribution.

**q\_value**

**NumbersofGTExTissuesQTLshared:** Number of GTEx (V8) tissues (excluding Pancreas and Testis) an eGene or Junction is shared with (according to overlapping QTL signals).

**GTExTissuesQTLshared:** GTEx (V8) tissues (excluding Pancreas and Testis) an eQTL or sQTL an eGene or Junction is shared with (according to overlapping QTL signals).

1. All significant variant-to-gene associations for phenotypes with a QTL effect at FDR 1%.

Nominally significant sQTLs: *NominallySignificant\_sQTLs.txt.gz*

Nominally significant eQTLs: *NominallySignificant\_eQTLs.txt.gz*

Headers:

**Gene:** gene symbol ID.

**Junction**: junction id based on the chromosome location (hg19 genome build), the intron cluster that belongs to, and the ENSEMBL ID of the corresponding gene. (chr:st:en:clu\_XX\_NA\_ENSGXXXXXXX). **Only applicable to sQTLs.**

**Variant\_id**: significant QTL variants associated with a given gene/splicing junction. \*Variant id is given as chr:position (hg19 genome build):Ref(allele):Alt(allele).

**Slope:** beta for the association between expression/junction usage and SNP dosage. The direction of the effect needs to be interpreted for the alternate allele.

**Lead\_(e/s)SNP**:lead eQTL or sQTL for a given eGene or junction, respectively.

1. Fine-mapped eQTL and sQTLsl

Credible sets for sQTLs: *sQTL\_CredibleSets.txt.zip* contains *SupplementaryData3\_sQTL\_CrediblSets.txt*

Credible sets for eQTLs: *eQTLs\_CredibleSets.txt.zip* contains *SupplementaryData4\_eQTL\_CredibleSets.txt*

Associated headers:

**sQTL\_junction (in sQTLs) or eQTL\_gene (in eQTLs)**: sQTL junctions are provided as chromosome location (hg19 genome build) of the junction, the intron cluster that belongs to, and the ENSEMBL ID of the corresponding gene. e.g. chr:st:en:clu\_XX\_NA\_ENSGXXXXXXX. For eQTLs we provide the gene symbol ID.

**variant**: fine-mapped eQTL or sQTL variants associated with the expression of a given gene or the relative usage of a given splicing junction.

\*\*Variant id is given as chr:position (hg19 genome build):Ref(allele):Alt(allele).

**cpp**: posterior probability of each variant to be causal calculated by CAVIAR.

**genic\_annotation (for sQTLs) or regulome\_annotation (for eQTLs)**: for sQTLs, whether fine-mapped variants overlap exonic, intronic, non-genic, 5’ or 3’ splice-sites according to GENCODE annotations. For eQTls, whether fine-mapped variants overlap class I, II or II active enhancers (Active\_enhancers\_I, Active\_enhancers\_II, Active\_enhancers\_III), inactive enhancers (Inactive\_enhancers), gene promoters (Active\_promoters), CTCF sites (CTCF), inactive open chromatin regions (inactive\_regions) or closed chromatin in human pancreatic islets. More information about islet regulome annotations is provided in Miguel-Escalada, I. (2019).