# MANUAL FOR QUERY

## Step 1: Find genes

### 1.1 Find genes from the database

- You can search for genes associated with a clinical trait by EWAS, GWAS or both. If you select both **“Implication by EWAS”** and “**Implication by GWAS**”, you must select **Union** or **Intersection** in the drop-down menu. **Union** will give you both EWAS and GWAS associations for a trait, **Intersection** will give you trait associations identified with both EWAS and GWAS in the same locus.

---

- **Implication by EWAS**: by checking this box you indicate to look for associations by EWAS. There are three different types of associations, (1) gene expression, (2) protein levels and (3) clinical and metabolite traits. If you select more than one type of association, you must specify if you wish you see the **Union** or **Intersection** of these associations using the drop-down menu.
- **Association with gene expression**. By checking this box you look for EWAS associations for gene expression traits. Indicate the *p*-value threshold for the association in the box. **Distance between gene and methylation site ≤** allows you to select the distance threshold between the gene and the associated CpG using the drop-down menu, for example <= 1Mb. The option **Number of Citeline clinical trials on gene ≥**, allows you to filter genes targeted by drugs in the Citeline database. The default is 0, meaning no minimum number of drugs targeting the gene.
- **Association with protein levels**. By checking this box you look for EWAS associations for protein levels. Indicate the *p*-value threshold for the association in the box. **Distance between gene and methylation site ≤** allows you to select the distance threshold between the gene and the associated CpG using drop-down menu, for example <= 1Mb. The option **Number of Citeline clinical trials on gene ≥**, allows you to filter genes targeted by drugs in the Citeline database. The default is 0, meaning no minimum number of drugs targeting the gene.
- **Association with clinical and metabolite traits**. By checking this box you look for EWAS associations for clinical and/or metabolite traits. Select the trait(s) in the box below, you can select one or more, traits are grouped by *Clinical traits* or *Metabolite traits*. Indicate the *p*-

value threshold for the association in the box. To generate a list of candidate genes for an association, you have two choices **(1)** **Link trait and gene through association with gene/protein expression** or **(2)** **Link trait and gene through proximity**.

  - You can select **(1)** **Link trait and gene through association with gene/protein expression**, to look for candidate genes that also have *cis*-expression or *cis*-protein associations in the locus (aka *cis*-eQTL and *cis*-pQTL). You can specify the *p*-value threshold for the gene expression and/or protein associations in the corresponding gene expression or protein level menus above, as well as a distance threshold to consider an association as *cis*, for example <= 1Mb.
  - *OR*, you can select **(2)** **Link trait and gene through proximity**. For a given clinical trait association, this will list all candidate genes in the associated locus within a certain distance of the associated CpG, for example <= 1Mb. Select the distance threshold from the drop-down menu labeled **Distance between gene and methylation site ≤**.
  - The option **Number of Citeline clinical trials on gene ≥**, allows you to filter genes targeted by drugs in the Citeline database. The default is 0, meaning no minimum number of drugs targeting the gene.

---

- **Implication by GWAS**: by checking this box you indicate to look for associations by GWAS. There are three different types of associations, (1) gene expression, (2) protein levels and (3) clinical and metabolite traits. If you select more than one type of association, you must specify if you wish you see the **Union** or **Intersection** of these associations using the drop-down menu.
- **Association with gene expression**. By checking this box you look for GWAS associations with gene expression traits. Indicate the *p*-value threshold for the association in the box. **Distance between gene and SNP site ≤** allows you to select the distance threshold between the gene and the associated SNP using drop-down menu, for example <= 1Mb. The option **Number of Citeline clinical trials on gene ≥**, allows you to filter genes targeted by drugs in the Citeline database. The default is 0, meaning no minimum number of drugs targeting the gene.
- **Association with protein levels**. By checking this box you look for GWAS associations with protein levels. Indicate the *p*-value threshold for the association in the box. **Distance between gene and SNP site ≤** allows you to select the distance threshold between the

gene and the associated SNP using drop-down menu, for example <= 1Mb. The option **Number of Citeline clinical trials on gene ≥**, allows you to filter genes targeted by drugs in the Citeline database. The default is 0, meaning no minimum number of drugs targeting the gene.

- **Association with clinical and metabolite traits**. By checking this box you look for GWAS associations for clinical and/or metabolite traits. Select the trait(s) in the box below, you can select one or more, traits are grouped by *Clinical traits* or *Metabolite traits*. Indicate the *p*-value threshold for the association in the box. To generate a list of candidate genes for an association, you have two choices **(1)** **Link trait and gene through association with gene/protein expression** or **(2)** **Link trait and gene through proximity**.
    - You can select **(1)** **Link trait and gene through association with gene/protein expression**, to look for candidate genes that also have *cis*-expression or *cis*-protein associations in the locus (aka *cis*-eQTL and *cis*-pQTL). You can specify the *p*-value threshold for the gene expression and/or protein associations in the corresponding gene expression or protein level menus above, as well as a distance threshold to consider an association as *cis*, for example <= 1Mb.
    - *OR*, you can select **(2)** **Link trait and gene through proximity**. For a given clinical trait association, this will list all candidate genes in the associated locus within a certain distance of the associated CpG, for example <= 1Mb. Select the distance threshold from the drop-down menu labeled **Distance between gene and methylation site ≤**.
    - The option **Number of Citeline clinical trials on gene ≥**, allows you to filter genes targeted by drugs in the Citeline database. The default is 0, meaning no minimum number of drugs targeting the gene.

----------------------------------------------------------------------------------------------------------------

### Step 2: Choose terms

- In this step you can select a term or list of terms related to candidate genes in the literature. For example, if you select the term “diabetes”, the output result will list the number of times a candidate gene was found to co-occur with the term “diabetes” in the literature. The number will indicate the total number of publications. You have two options to select terms, these are **2.1 Choose curated terms**, or **2.2 Append your own terms**:

### 2.1 Choose curated terms

- You can select from a list of curated terms and/or genes related to atherosclerosis, diabetes and obesity.

### 2.2 Append your own terms

- OR, you can insert your own list of terms.

### Step 3: Choose search scope

- The search scope indicates whether the term co-occurs with a gene in the literature only in the **Title and abstract** of the publication, or the **Full text** i.e. the entire publication. We suggest selecting **Title and abstract**. Selecting the entire publication may lead to a very broad relationship between the gene and the term.