

# Gene Search and Results

- [Using the Gene Search](#)
- [Search by Name](#)
  - [Search by Name, Description, UniProt ID](#)
  - [Name types](#)
- [Search by Gene List and Bulk Downloads](#)
- [Search by Keyword](#)
  - [Keyword Term](#)
  - [Keyword type](#)
  - [Evidence](#)
- [Restrict by Features](#)
  - [Gene Model Type](#)
  - [has associated literature](#)
  - [genetic \(unsequenced\) loci only](#)
- [Restrict by Map Location](#)
  - [Chromosome](#)
  - [Map Type](#)
  - [Range](#)

## Using the Gene Search

---

The TAIR Gene Search window provides three ways of searching for a gene: simple search by name only, feature search using more limits, and search by position. These are offered in three distinct sections of the window.

### Search by Name

At the top of the TAIR Gene Search window are the Search by Name options.

This simple search is case insensitive and allows wildcard searching (see "[Using Wildcards](#)").

#### Search by Name, Description, UniProt ID

Use this drop down menu to search either by gene name, description, phenotype or UniProt ID or TAIR locus ID and TAIR accession. You can choose exact, contains or start with search for names, descriptions and phenotypes. Identifier searches are exact matches.

#### Name types

In TAIR, there are four types of names associated with a Gene. Searching by name will search all gene names and gene aliases. For more information about the name types see below:

- **Symbol**  
This is the mnemonic naming used for gene names by researchers. Examples include AG (Agamous), and QRT1 (Quartet1). A symbol for a gene is designated when a gene has been published or if the name and symbol have been registered at Oklahoma State, currently maintained by David Meinke's group or in GenBank. (<http://mutant.lse.okstate.edu/genepage/genepage.html>).
- **ORF name**  
An open reading frame (ORF) name from the Arabidopsis Genome Initiative (AGI) groups' annotations. Usually, the convention for naming an ORF in Arabidopsis has been using the clone name followed by a number suffix (e.g., F23H14.13). For chromosome arms that have been completely sequenced, a standard ORF name designation is used:
  - AT (Arabidopsis thaliana)
  - 2 (chromosome number)
  - G (for Gene)
  - 01130 (Number)Examples: At2g01130, AT4g00010
- **Full name**  
The full descriptive name of a gene. Examples: Agamous, Aspartate aminotransferase deficient 3
- **Gene product name**  
Name of a gene product. Examples: ASPARTATE AMINOTRANSFERASE, CYTOPLASMIC ISOZYME 1, CHALCONE SYNTHASE. For the genes that do not have a full name or symbol (largely the predicted genes (ORFs) from AGI sequencing and annotation), the following product names have been used:
  - Hypothetical protein  
Gene models without any database matches.
  - Unknown protein  
Gene models with only EST matches.

### Search by Gene List and Bulk Downloads

Users can now search in bulk and upload a list of AGI Locus identifiers. This feature can also be used as a starting point for bulk downloading of data (see [Bulk Downloading Gene Data help](#) or our [YouTube Tutorial](#)).

1. Paste or upload a list of AGI identifiers (one ID per row)
2. Apply any additional search parameters to the gene list or click submit
3. On the results screen at the top will be a number of options for retrieving data for all or a subset of the results
  - a. Get GO annotations- retrieve Gene Ontology annotations
  - b. Get PO annotations-retrieve Plant Ontology annotations
  - c. Get Sequences-retrieve FASTA formatted sequences
  - d. Get Gene Descriptions- retrieves gene summaries and aliases
  - e. Get Locus History- retrieves history of loci such as create date, merges , deletes etc.
  - f. Get Microarray Elements- retrieves corresponding microarray element IDs

## Search by Keyword

This option allows you to search for genes by keywords. Genes are annotated with keywords describing molecular function, subcellular localization and biological process with controlled vocabularies from the [Gene Ontology Consortium](#). Genes are also annotated with controlled vocabulary terms for plant structures and developmental stages for describing phenotypes and where and when the genes are expressed. Plant structure and development terms are from the [Plant Ontology Consortium](#). Each annotation is associated to an evidence code and one or more reference (the source of the evidence for the annotation).

### Keyword Term

This option allows you to input a given term and choose the type of search. Exact searches are most restrictive, contains searches are least restrictive but slower. If you are not sure of the format of the term, choose the 'contains' option. For example exact search for protein kinase will return results for only those terms that exactly match protein kinase, whereas a search for contains protein kinase will include matches to terms such as serine-threonine protein kinase AND histidine protein kinase. You can also limit the search to keywords that start with or end with a given term.

### Keyword type

The keyword types represent the categories of keywords used for annotation. You can restrict your search to include only one type of keyword or choose multiple types. To select more than one type of keyword hold down the Apple (MAC) or CTRL (PC) while clicking with the mouse. If you are not sure of which type the keyword belongs to, use the ANY option which searches all types of keywords.

### Evidence

Each annotation in TAIR is associated to some form of evidence that supports the annotation. The evidence can be used to filter annotations, for example, to exclude annotations based on computational methods (inferred by electronic annotation). Evidence is also used to quickly assess the quality of a given annotation. The evidence is also linked to the source of the evidence (i.e. the experimental data or analysis method).

## Restrict by Features

The Restrict by Features options are below the Search by Name options on the TAIR Gene Search window.

These options let you restrict your search by selecting one or more features. Selecting multiple attributes indicates an "AND" relationship.

### Gene Model Type

This feature allows you to restrict your search to specific types of gene models. The default search returns any gene model type. To select multiple gene model types after your first selection click on additional ones while holding down either the CTRL key (PCs) or the Apple key (Mac).

### has associated literature

Checking this box will limit your search to include only genes with associated publications.

### genetic (unsequenced) loci only

This option allows to choose only genes which are represented only as genetic loci (e.g. have not been cloned or associated with a sequence yet).

## Restrict by Map Location

The bottom section of the window lets you restrict your search by location.

The options in this section let you use three parameters to restrict your search: Chromosome, Map Type, and Range.

### Chromosome

Lets you limit your search to a single chromosome. There are five nuclear chromosomes in *Arabidopsis*: 1, 2, 3, 4, and 5 and also the mitochondrial and chloroplast genomes.

### Map Type

Lets you search entities by their position on a particular map and is to be used with the Range parameters. Currently, you can search on only one map type at a time. The options are Reference Assembly (currently Araport11 which corresponds to the TAIR9 genome assembly), physical (BACs) and genetic maps.

## Range

Lets you specify a range search by the upper and lower bounds. The value is interpreted based on the selected range units. You can specify the range by physical distance (kb).