

Plant Genomes Central: Integrated Resources for Plant Genomics

Plant Genomes Central (PGC) is an integrated, Web-based portal for plant genomics data and tools at the National Center for Biotechnology Information (NCBI). This article explains the goals of Plant Genomes Central, and provides some examples of how to use databases and tools at PGC to access and analyze plant genomic data.

NCBI provides a large suite of integrated, general-purpose tools for accessing, analyzing, and visualizing biological data. These tools include:

- *Basic Local Alignment and Search Tool (BLAST)*—a suite of tools for sequence alignment
- *MapView*—a system to search and display genomic information by chromosomal position
- *UniGene*—an automated system for creating clusters of GenBank sequences related to individual genes
- *Entrez*—a text-based search and retrieval system for all NCBI biological data and publications databases

Many of these tools are designed to work with all sorts of biological data, from all types of species. Because the tools are so general, they can be challenging to use in specific research contexts. Furthermore, most researchers are only interested in the subset of data that is relevant to their research. NCBI serves research communities with specific interests by creating information “portals” that provide access to general-purpose NCBI tools and databases, customized to the needs of the specific community.

PGC has three purposes. First, it integrates all of the genetic maps for each plant species genome with observed sequence information for that species. Second, PGC provides interfaces to NCBI databases and search, analysis, and visualization tools that are customized to support the needs of plant biology researchers. Third, PGC provides extensive hyperlinking to databases outside of NCBI for additional information about particular species.

PGC offers three types of resources that are specific to plant biology:

- Prepared Entrez queries (in the form of Web hyperlinks) provide single-click access to data generated by various EST sequencing projects
- MapViewer can be used to display and browse high-resolution genetic maps for species in PGC that have such maps
- PGC can also display sequence maps for public genomes

Most of the genetic maps and sequence data in NCBI's databases come originally from collaborators outside of NCBI. A list of the plant genomics collaborators whose data appear in PGC appears in the appendix at the end of this article. PGC adds value to these existing databases by integrating them into a single system, providing a standardized set of analysis tools, and providing tool interfaces customized to the needs of plant biologists.

Species for which more than 70,000 EST sequences exist are listed in UniGene, an experimental system for automatically partitioning GenBank sequences into non-redundant sets of gene-oriented clusters. Each UniGene cluster contains sequences that represent a specific gene, plus related information such as tissue-specific expression data and map location.

Plant Genomes Central Main Page

The PGC main page can be accessed using a standard Web browser at the Web address <http://ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>.

Down the left side of the main page, in the blue navigation bar, is a large list of related resources. A group of links to *NCBI FTP SITES* links to sites for downloading the data for entire genomes. Two more groups of *Related Resources* link to other plant genome projects, organized by function and by organism. Finally, a section called *Plant UniGene* links to the UniGene database main page for each species.

The body of the main page consists of four sections:

- *Plant-oriented resources at NCBI* contains links to an multiple-genome search page and a special BLAST page for plant genomes
- *Large-Scale Sequencing Projects* contains links to data from whole-genome sequencing projects.
- *Genetic Maps* contains links to genetic maps in PGC. It includes chromosomal maps, and in some cases also maps for plastids, mitochondria, and mitochondrial plasmids.
- *Large-scale EST Sequencing Projects* contains links to EST data and summary pages for each organism.

The next few sections of this article will explain each of these in detail.

Plant-oriented Resources at NCBI

The top section of the main page appears in Figure 1 below.

NEW Plant-oriented Resources at NCBI

- Multi-species plant genome map [search](#)
- Customized plant genome [BLAST](#)
This BLAST database includes three datasets:
 - GenBank sequences used to identify mapped genetic loci in *Avena sativa* (oat), *Hordeum vulgare* (barley), *Oryza sativa* (rice), and *Triticum aestivum* (wheat)
 - The collection of contigs developed by the Chinese WGS endeavor
 - The *Arabidopsis thaliana* genome

<http://ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>

Figure 1. Plant resources section of PGC main page

The *Multiple genome search* page performs searches across mapped plant genomes. The special *BLAST* page is customized for searching plant genomes. Other links are occasionally added to this section as new resources are added.

The Multiple Genome Search Page

The top of the multiple genome search page appears in Figure 2 below.

Plant genomes query

The screenshot shows a web interface for searching plant genomes. At the top, there is a search bar with the text "Search for [] in all selected organisms" and buttons for "Find" and "Reset". Below this is a section for selecting chromosomes/maps, with a "View all chromosomes/maps selected below" link and a "Display" button. There are also "Select all", "Clear", and "Query Help" options. The interface lists three species: **Arabidopsis thaliana** (Chromosome: 1), **Avena sativa (oat)** (Chromosome: A), and **Hordeum vulgare (barley)** (Chromosome: 1H). Each species has a "Maps:" section with various map types and their corresponding checkboxes. For Arabidopsis thaliana, the maps include Marker, Clone_gb, Clone_tigr, Genes_seq, AtUniG, At ESTs, HvUniG, Hv ESTs, OsUniG, Os ESTs, TaUniG, Ta ESTs, ZmUniG, ZmESTs, Contig, and RNA. For Avena sativa, the maps are AxH_92 and SxW_00. For Hordeum vulgare, the maps are Gene, Gene2, NARGMP (SxM), and NARGMP (HxT). The URL at the bottom is http://ncbi.nlm.nih.gov/mapview/map_search.cgi?chr=plants.inf.

Figure 2. Multi-genome search page

This page can be used in comparative genomics to search the multiple plant species maps in a single operation. Each species name is linked to a species “home page”, where NCBI maintains a centralized set of resources for the species.

The user interaction for this page is still being designed, and so it will change over time. The form allows the user to search the maps for these species using Entrez syntax. After searching, the user may select genetic maps to display in the MapViewer. An example of how to use the MapViewer appears later in this article.

Special BLAST Interface for Plant Genomics

The plant genome BLAST databases are a special collection of mapped and unmapped plant genomic sequences. The user may search mapped genetic loci in wheat, barley, oats, or rice; may search the unmapped collection of contigs developed by the Chinese rice WGS project; or may search the *A. thaliana* genome. Each BLAST result includes a link to a MapViewer display that graphically displays BLAST hits in the context of other chromosomal landmarks.

The BLAST link on the PGC main page leads to a BLAST form for plant genomes, as shown in Figure 3 below.

BLAST against DNA/protein collections mapped in, or associated with, plant genomes

Proteins and mRNA data for BLAST search is available for Arabidopsis thaliana and Oryza sativa only.

Source: Oryza sativa

Database: proteins (A.thaliana or Oryza sativa)

Program: blastx MegaBlast

Enter an accession, GI, or a sequence in FASTA format:

Optional parameters:

[Expect](#) [Filter](#) [Descriptions](#) [Alignments](#)

0.01 default 100 100

[Advanced options:](#)

<http://www.ncbi.nlm.nih.gov/BLAST/Genome/PlantBlast.shtml>

Figure 3. Plant genome BLAST form

This BLAST form provides a choice of the plant source species available from PGC. All source species have DNA databases. Two species, *A. thaliana* and *O. sativa*, have public, large-scale sequencing projects, so these two also have mRNA and protein databases. Choosing *Oryza sativa*, as shown Figure 3, results in three choices of database: DNA sequence, proteins, and mRNA sequences. Choosing a database automatically limits the Program selection immediately following to those BLAST variants that apply to the selected type of data.

To perform a BLAST search, choose a source species, database, and BLAST program. Either enter an accession or gi number, or paste FASTA-formatted sequence data into the box. You may use the controls directly under the text box to tune BLAST parameters and add options. Click **Begin Search** to start the BLAST.

Large-Scale Sequencing Projects

Returning to the PGC main page, the second section provides access to data from large-scale genome sequencing projects. This section of the page appears in Figure 4 below.

Large-Scale Sequencing Projects

- [5] [Arabidopsis thaliana](#) (thale cress)
chromosomes: [1](#), [2](#), [3](#), [4](#), [5](#), [plastid](#), [mitochondrion](#)
- [12] [Lycopersicon esculentum](#) (tomato)
chromosomes: [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#), [9](#), [10](#), [11](#), [12](#), [plastid](#),
[mitochondrion](#)
- [8] [Medicago truncatula](#) (barrel medic)
chromosomes: 1, 2, 3, 4, 5, 6, 7, 8 [plastid](#), [mitochondrion](#)
- [12] [Oryza sativa](#) (rice)
chromosomes: [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#), [9](#), [10](#), [11](#), [12](#), [plastid](#) ,
[mitochondrion](#), [mitochondrial plasmid B1](#), [mitochondrial plasmid B2](#),

<http://ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>

Figure 4. Large-scale genome sequencing resources

In this section, the number of chromosomes for the species precedes the species name. The species name is linked to the species home page. Chromosome numbers are linked to the MapViewer for the chromosome, and organelles are linked to the Entrez Genome page for the organelle's genome. Absence of a link ("1" for *M. truncatula*, for example) indicates that no data are as yet available for that molecule. This list will grow as data are submitted to NCBI.

High-Resolution Genetic Maps

The next section of the main page links to PGC's high-resolution genetic maps, as shown in Figure 5 below.

Genetic Maps

- [7] [Avena sativa](#) (oat)
chromosomes: [A](#), [B](#), [C](#), [D](#), [E](#), [F](#), [G](#), plastid, mitochondrion
- [20] [Glycine max](#) (soybean)
chromosomes: [A1](#), [A2](#), [B1](#), [B2](#), [C1](#), [C2](#), [D1a](#), [D1b](#), [D2](#), [E](#), [F](#), [G](#), [H](#), [I](#), [J](#), [K](#), [L](#), [M](#), [N](#), [O](#), plastid, mitochondrion
- [7] [Hordeum vulgare](#) (barley)
chromosomes: [1H](#), [2H](#), [3H](#), [4H](#), [5H](#), [6H](#), [7H](#), plastid, mitochondrion
- [12] [Lycopersicon esculentum](#) (tomato)
chromosomes: [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#), [9](#), [10](#), [11](#), [12](#), plastid, mitochondrion
- [12] [Oryza sativa](#) (rice)
chromosomes: [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#), [9](#), [10](#), [11](#), [12](#), [plastid](#), mitochondrion, [mitochondrial plasmid B1](#), [mitochondrial plasmid B2](#),
- [10] [Sorghum](#)
chromosomes: [A](#), [B](#), [C](#), [D](#), [E](#), [F](#), [G](#), [H](#), [I](#), [J](#), plastid, mitochondrion
- [7] [Triticum aestivum](#) (bread wheat)
chromosomes: [1A](#), [1B](#), [1D](#), [2A](#), [2B](#), [2D](#), [3A](#), [3B](#), [3D](#), [4A](#), [4B](#), [4D](#), [5A](#), [5B](#), [5D](#), [6A](#), [6B](#), [6D](#), [7A](#), [7B](#), [7D](#), [plastid](#), mitochondrion
- [10] [Zea mays](#) (corn)
chromosomes: [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#), [9](#), [10](#), [plastid](#), mitochondrion, [mitochondrial 1.9 kb plasmid](#)

<http://ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>

Figure 5. Links to high-resolution genetic maps

As with the large-scale sequencing projects, the species names lead to the species home pages, while links for chromosomes and organelles lead to MapViewer and Entrez Genome, respectively.

Multiple maps exist for many of the species shown here. Pulling all of the public maps into one resource provides the ability to compare the multiple maps that exist for a single species.

Large-scale EST Sequencing Projects

Finally, the PGC main page ends with a list of large-scale EST sequencing projects, as shown in Figure 6 below.

<http://ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>

Large-Scale EST Sequencing Projects

Avena sativa (oat)	EST sequences
Beta vulgaris subsp. vulgaris (beet)	EST sequences
Glycine max (soybean)	EST sequences
Gossypium arboreum (tree cotton)	EST sequences
Gossypium hirsutum (upland cotton)	EST sequences
Hordeum vulgare (barley)	EST sequences
Lotus japonicus	EST sequences
Lycopersicon esculentum (tomato)	EST sequences
Lycopersicon hirsutum	EST sequences
Lycopersicon pennellii	EST sequences
Marchantia polymorpha (liverwort)	EST sequences
Medicago sativa (alfalfa)	EST sequences

Figure 6. EST project data

In this section, the species name links to the NCBI Taxonomy Browser page for that species. The taxonomy browser includes such information as various formal and informal species name, lineage, genetic code, and a summary of how many of each type of record exists in Entrez. It also links to a viewer for each chromosome or organelle, and provides a comprehensive list of “link outs”—links to resources external to NCBI.

The [EST sequences](#) link for the species leads to an Entrez Nucleotide search that yields all ESTs in GenBank for that species. When Entrez appears, its search box is already initialized with a query that can be modified, if desired, to make the query more specific.

All grass ESTs (barley, wheat, corn, and rice) have precomputed, stored alignments to the *Arabidopsis* genome. EST data are also used to build UniGene clusters. Both UniGene clusters and ESTs aligned to the *A. thaliana* genome can be viewed in the MapViewer.

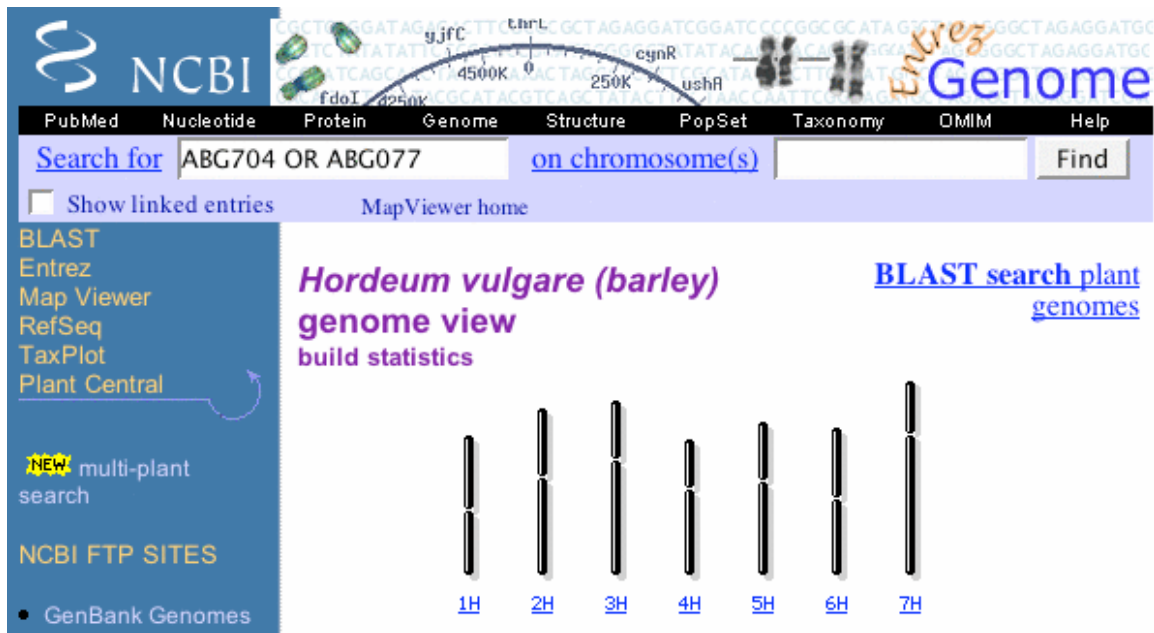
Having completed a tour of the resources available on the PGC main page, we move on to some examples of how to use PGC in research.

Example: Identifying Candidate Gene Regions With MapViewer

One way to identify a candidate gene for a biological phenomenon is to identify two genetic markers that co-segregate with the trait. The example shows how to use MapViewer to identify a candidate gene in *H. vulgare* (barley) associated with resistance or susceptibility to barley stem rust, a plant infection of the fungus *Puccinia graminis* f. sp. *Tritici*.

Given two trait-linked genetic markers identified in barley, MapViewer can be used to identify their physical location in the genome, determine the relative distance between them, and find other genes linked to resistance or susceptibility. Laboratory research has shown that resistance and susceptibility are linked to inheritance of a region near barley markers ABG704 and ABG077.

Start at the barley species home page, accessible from the PGC main page in the *Genetic Maps* section. Click on the species name. The barley genome view appears, with a search form at the top, followed by a schematic diagram of the set of barley chromosomes, taxonomic classification of the species, and other explanatory text about the barley genome. To search for the two markers, enter the expression “ABG704 OR ABG077” in the search box as shown in Figure 7. Since we don’t yet know on which chromosome these markers lie, leave the on chromosome(s) search box empty.



The screenshot shows the NCBI MapViewer interface for *Hordeum vulgare* (barley). At the top, there is a search bar with the text "ABG704 OR ABG077" and a dropdown menu set to "on chromosome(s)". Below the search bar, there is a "Find" button and a "Show linked entries" checkbox. The main content area displays the "Hordeum vulgare (barley) genome view" with a schematic of seven chromosomes labeled 1H through 7H. A red highlight is visible on chromosome 7H, indicating the location of the search results. The interface also includes a navigation menu on the left with options like BLAST, Entrez, Map Viewer, RefSeq, TaxPlot, and Plant Central, and a "BLAST search plant genomes" link on the right.

Figure 7. Searching for two loci at once

Clicking “Find” produces two hits, as shown in Figure 8. The results indicate that the two markers map near or within the telomeric regions of the short arm of chromosome 7H.

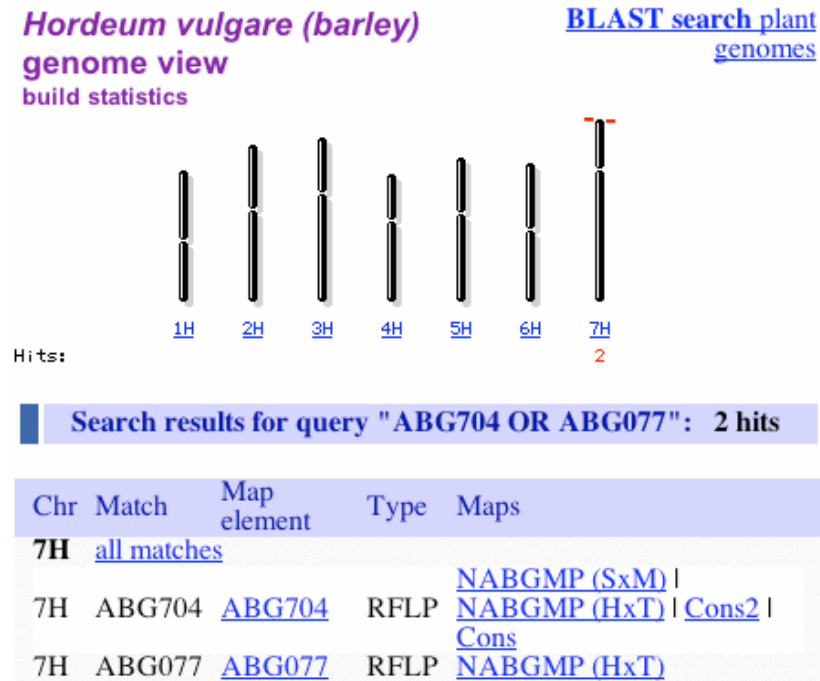


Figure 8. Result of genome search

The table below the chromosomes in Figure 8 shows that four maps contain at least one of the markers: NABGMP (SxM), NABGMP (HxT), Cons2, and Cons. NABGMP (HxT) contains both.

Clicking on the link 7H at the bottom of the corresponding chromosome displays a MapViewer for the four maps that matched. (Often, a MapViewer comes up as a “compressed” view, where locus names are replaced by small diamonds. Uncheck the “Compress View” box to see the locus names.) This MapViewer appears in

Figure 9 below.

Note that two loci, Plc and dRpg1, lie very close to marker ABG704. Clicking on the names in the Symbol and Probe columns of this page will lead to the GrainGenes database, which provides more comprehensive information about these two loci. The Links column has an [acc](#) link for Plc that leads to the GenBank accession for the locus.

Further research of these two loci reveals some information about their function. Plc encodes a precursor of plastocyanin, a Photosystem I electron donor in the thylakoid lumen in chloroplasts. dRpg1 is a receptor kinase-like gene that has been demonstrated to be defective in stem rust-susceptible barley strains.

The **Maps & Options** links on the MapViewer page provide a great deal of flexibility and customization. Click the Help link on the MapViewer page to learn more about how to customize MapViewer for your needs.

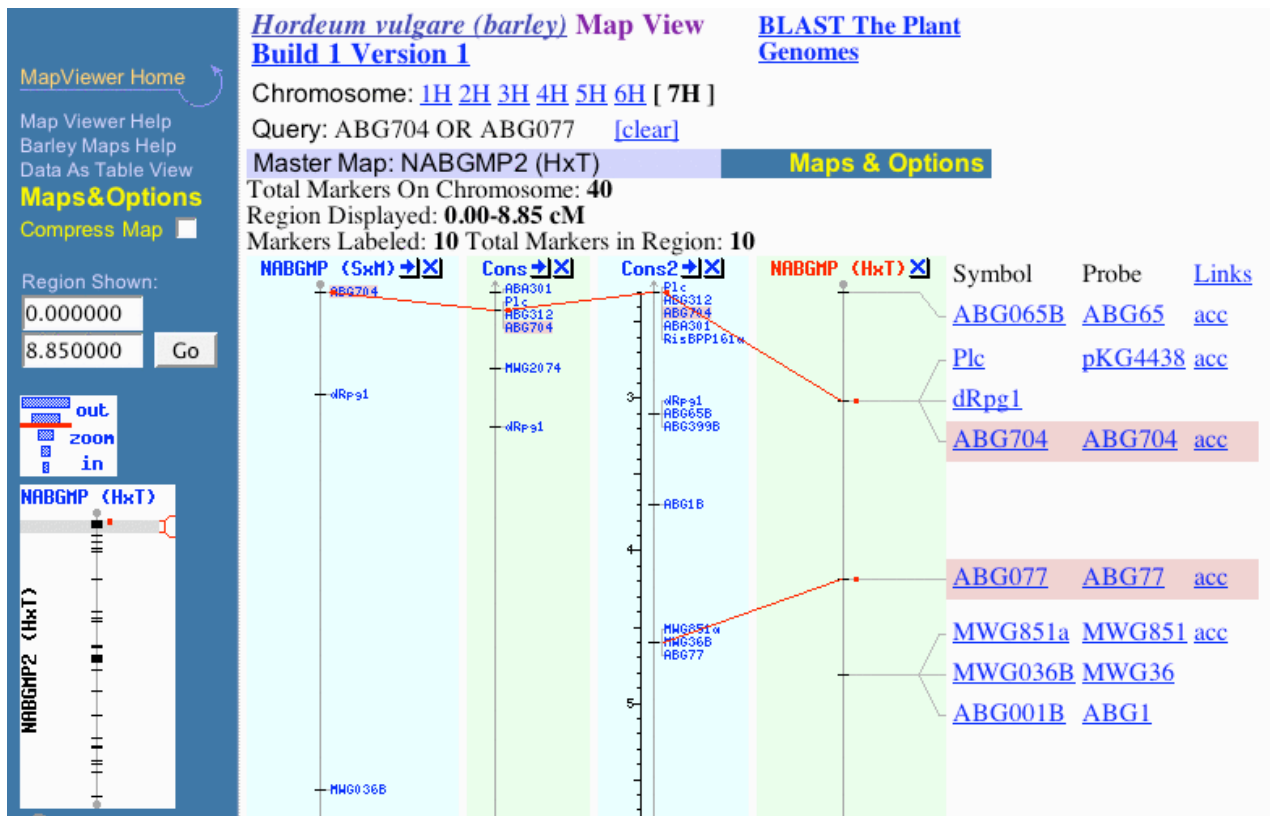


Figure 9. Map view of two markers in the *H. vulgare* genome on four maps

Example: EST Identification Using Cross-Species Plant Genome BLAST

This example demonstrates how to use the Plant Genome BLAST utility to identify an uncharacterized EST sequence from a species that lacks a genetic map. The identification approach is to search for homologs to the EST sequence in a mapped species, and then use resources for that species to infer structure and function of the uncharacterized gene.

In this example we use blastx to search the *A. thaliana* protein database with an uncharacterized EST from Poplar (GenBank accession #BU898322), looking for homologs. This example also demonstrates how integrating search results with other resources makes interpreting those results easier.

This example uses an uncharacterized EST sequence from a Poplar wood cDNA library. Because we have no genetic map for Poplar, we instead use blastx against the *A. thaliana* protein database, looking for homologs.

Begin at the PGC main page. Click on the [BLAST](#) link in *Plant-oriented resources at NCBI*. For the Source species, select *Arabidopsis thaliana*; for the Database, select *proteins (A. thaliana or Oryza sativa)*; and for Program, select *blastx*. In the text box, enter the accession number of the uncharacterized Poplar EST, #BU898322. Click *Begin Search*. You will be redirected to the “formatting BLAST” page; click the *Format!* button. The result will be a page containing links to several types of information.

First on the BLAST results page is a button marked *Genome View*. Clicking this button shows a MapViewer that displays the location of the BLAST hits on the *A. thaliana* genome. The BLAST hits are color-coded by score. Beneath the graphic is a long list of hits, sorted by chromosome. Click [Sort results by score](#), at the top right of the hit table. The graphic and part of the table *after* sorting appears in Figure 10. (This graphic does not display an associated URL, because the data are volatile.)

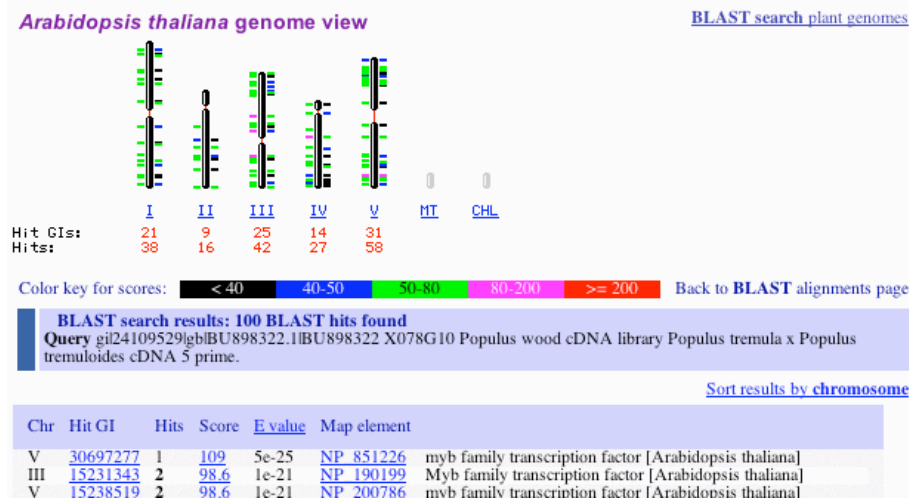


Figure 10. BLAST hits of EST against *A. thaliana* genome

The most-significant hit, with a score of 109 and an E-value of 5e-25, is on chromosome V. The element this hit references is a transcription factor of the myb family. In plants, myb transcription factors play important roles in cell cycle regulation, secondary metabolism, and cell fate. Clicking on the accession number (NP_851226) in the *Map element* column of the result table shows a MapViewer for that locus. Part of the graphical representation of the gene appears in Figure 11.



Figure 11. Map view of most-significant BLAST hit

The small arrow pointing up indicates that the gene is encoded on the (-) strand. At this point, the integrated NCBI tool set offers a wealth of information about this gene:

- Clicking on the gene name [At5g59780](#) leads to Entrez Gene, a gene-centric view of data. This view appears in Figure 13 on the next page. Among other things, the gene view shows that this gene has three splice variants.
- Clicking [sv](#) leads to the Sequence Viewer for this gene, shown in Figure 13.
- Clicking [pr](#) leads to Entrez Protein, for access to the GenBank protein record for this gene; see Figure 14.
- The protein link page also provides access to *BLink*, which is a link to precomputed BLAST results that find proteins similar to this one, as shown in Figure 15.
- From the BLink page, clicking the *Taxonomy Report* button organizes the BLAST results taxonomically, as shown in Figure 16.

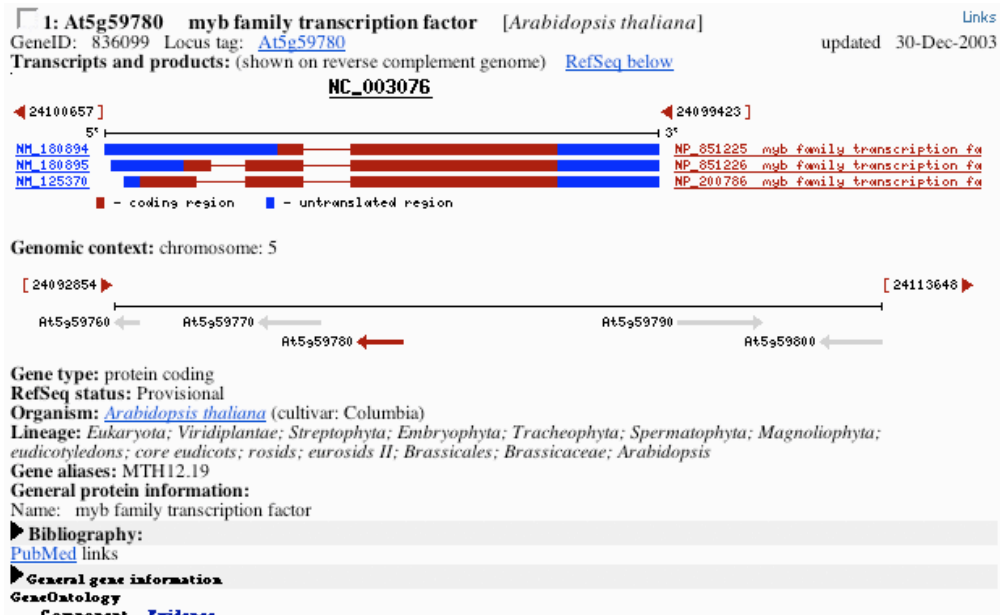


Figure 12. Gene view of myb locus

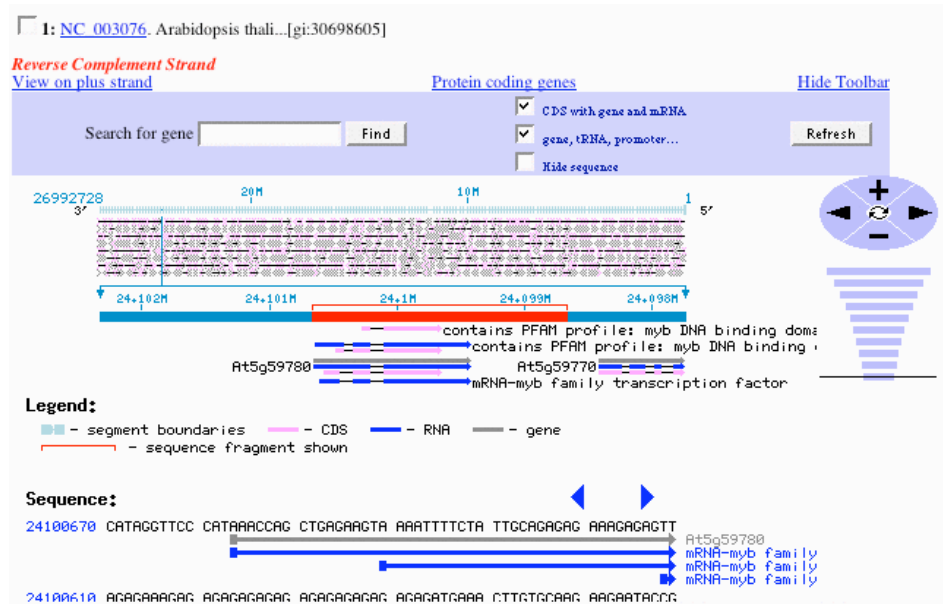


Figure 13. Sequence view of myb locus



Figure 14. Link to GenBank accession for protein record for myb locus

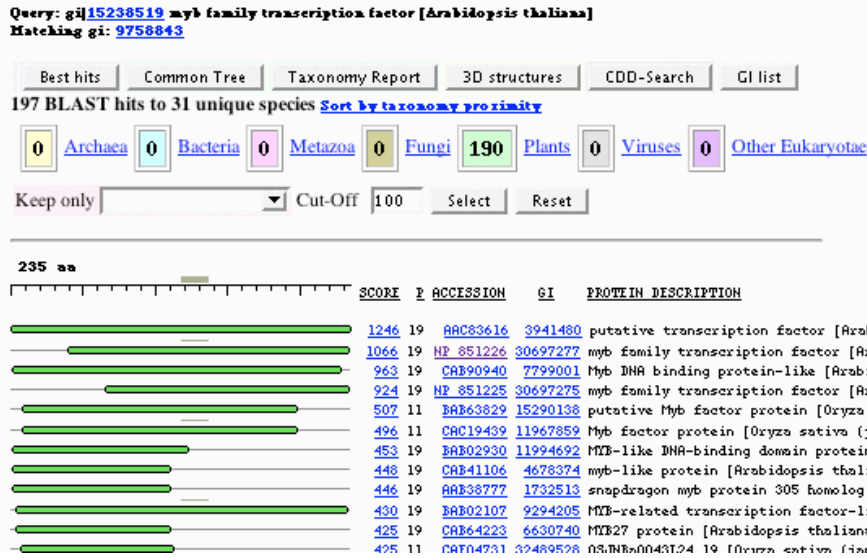


Figure 15. Precomputed BLAST results for myb

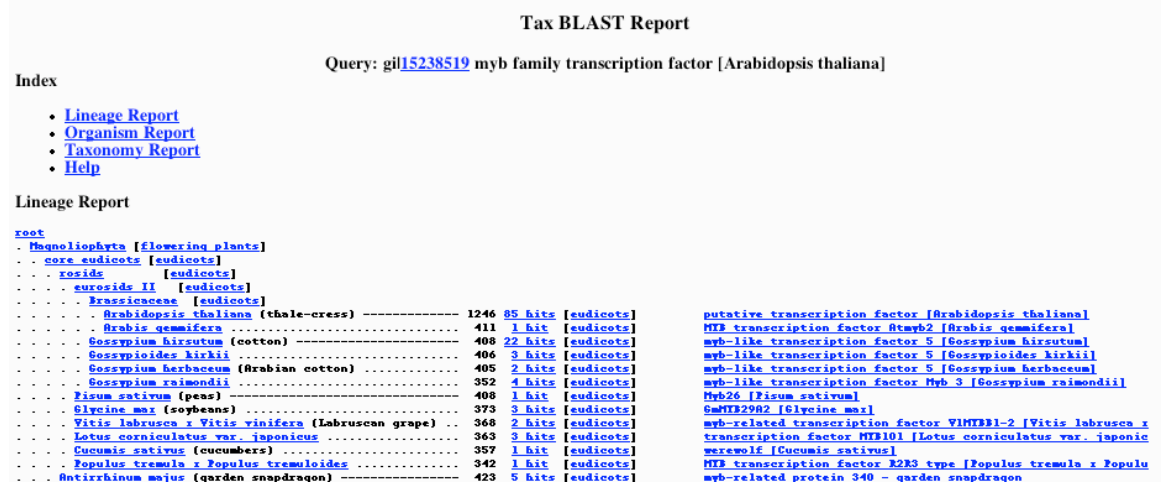


Figure 16. Taxonomic categorization of BLAST results

Much more information is available. On the Protein page (Figure 14), clicking the [Domains](#) link displays a graphical representation of protein domains, while the [Links](#) link connects to other Entrez databases, such as Entrez Gene, Related Sequences, Domain Relatives, Genome, MapViewer, Nucleotide Taxonomy, and Link-Out.

The BLAST results clearly indicate that the myb family of proteins is evolutionarily conserved across many plant species.

Example: Finding Conserved and Syntenic Regions of Plant Genomes

Restriction fragment length polymorphisms (RFLPs) are commonly used as molecular probes to identify conserved and non-conserved regions of plant genomes. NCBI's multiple species query page searches against all of the available high-resolution genetic maps in PGC. At this writing, eight species have such maps, and so are included on the multiple genome search page. In this example, we will search the plant genomes for RFLP marker cdo718.

Begin at the main page of Plant Genomes Central, and click *Multi-species plant genome search* in the first section of the page. The result is the multiple genome search form. Notice that, by default, all species on this page are already selected.

Enter the RFLP name cdo718 in the *Search for* box and click the *Find* button. The result indicates how many times the marker was found in each genome: once each in oats, wheat, barley, and corn, and five times in rice. Checkboxes indicate which maps contain hits, and chromosome drop-down boxes report how many hits occurred on each chromosome.

<input checked="" type="checkbox"/> Avena sativa (oat)	1 hit	A (1 hit)
Maps: <input checked="" type="checkbox"/> AxH_92 <input type="checkbox"/> SxW_00		
<input checked="" type="checkbox"/> Hordeum vulgare (barley)	1 hit	3H (1 hit)
Maps: <input checked="" type="checkbox"/> Cons <input type="checkbox"/> Cons2 <input type="checkbox"/> NABGMP (SxM) <input type="checkbox"/> NABGMP (HxT)		
<input checked="" type="checkbox"/> Oryza sativa (rice)	5 hits	2 (5 hits)
Maps: <input type="checkbox"/> HD00 <input type="checkbox"/> HD98 <input checked="" type="checkbox"/> R <input checked="" type="checkbox"/> RC01		
<input checked="" type="checkbox"/> RC94 <input type="checkbox"/> RD00 <input type="checkbox"/> RD01 <input type="checkbox"/> RI96		
<input type="checkbox"/> RJ94 <input type="checkbox"/> RK98 <input type="checkbox"/> RM00 <input type="checkbox"/> RS98		
<input checked="" type="checkbox"/> RW99 <input type="checkbox"/> RX97 <input checked="" type="checkbox"/> RC00 <input type="checkbox"/> fpc_contig		
<input type="checkbox"/> fpc_clone <input type="checkbox"/> genes <input type="checkbox"/> RNA <input type="checkbox"/> contig		
<input type="checkbox"/> clone		
<input checked="" type="checkbox"/> Triticum aestivum (bread wheat)	1 hit	3B (1 hit)
Maps: <input type="checkbox"/> Ta-Gene <input checked="" type="checkbox"/> S-O <input type="checkbox"/> Bin		
<input checked="" type="checkbox"/> Zea mays (corn)	1 hit	5 (1 hit)
Maps: <input type="checkbox"/> QTL <input type="checkbox"/> IBM00 <input checked="" type="checkbox"/> MW99 <input type="checkbox"/> Em35		
<input type="checkbox"/> Bin <input type="checkbox"/> IBM02 <input type="checkbox"/> UMC98		

Figure 17. Result of RFLP marker search

Clicking the name of a particular species on this page shows just the markers for that species. Clicking the Display button (further up the page) shows all of the maps together in the MapViewer. The red line connecting the loci on the different maps shows the placement of cd0718 RFL on each of the maps. Uncheck the "Compress Maps" box to reveal all of the locus names.

Plant Genomes Comparative Map View [BLAST The Plant Genomes](#)

Query: cdo718 [\[clear\]](#)

Master Map: MW99 Maps & Options

Total Markers On Chromosome: 62

Region Displayed: 0.00-300.40 cM

Markers Labeled: 20 Total Markers in Region: 62

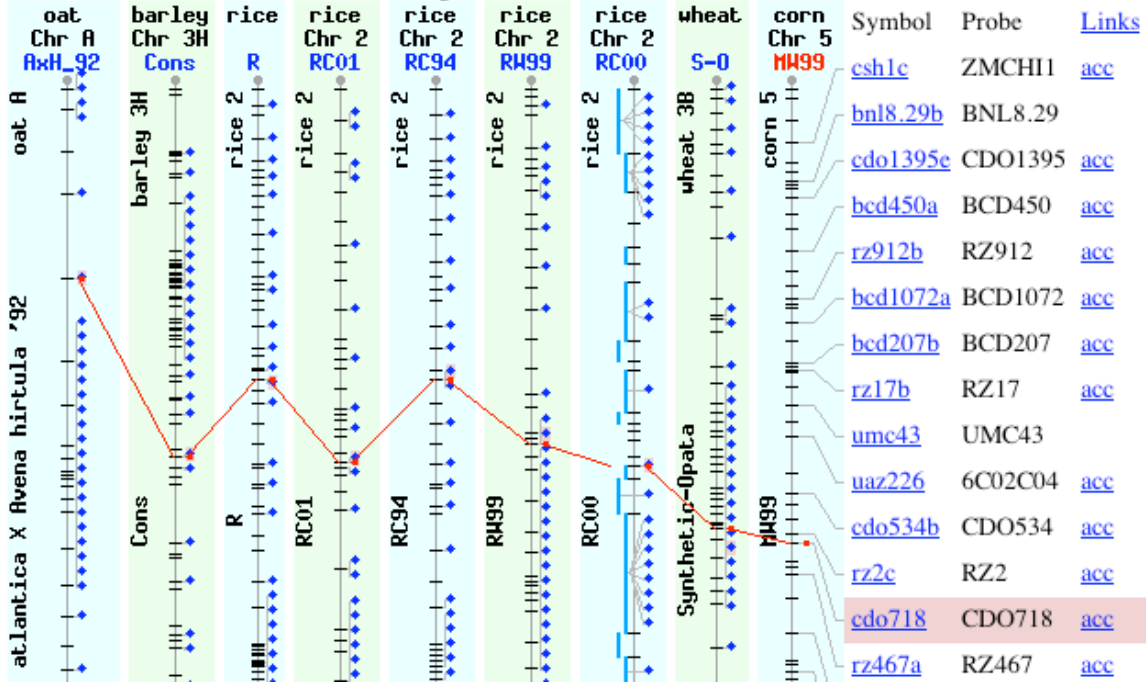


Figure 18. Multi genome search results for RFLP cd0718

Resources

Taxonomy

All species in GenBank are cross-listed in the Entrez Taxonomy database. The Taxonomy database is a curated resource that provides taxonomical classification of each species, its scientific and common name, a tabulated display of the number of organism-specific hits in all relevant Entrez databases, and a Link-Out feature to selected external databases. Find the taxonomy database at

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Taxonomy>

Sequence downloads

Direct access to genomic, cDNA, EST, STS, genome survey sequences, and high-throughput genomic sequences can be found in Entrez Nucleotides at

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=nucleotide>

The plant databases for Standalone BLAST are available at

<ftp://ftp.ncbi.nih.gov/genomes/PLANTS/BLASTDB>

<ftp://ftp.ncbi.nih.gov/refseq/release/plant> (protein and RNA files)

Whole chromosome records are available for *A. thaliana* at

ftp://ftp.ncbi.nih.gov/genbank/genomes/A_thaliana/OLD/and

Reference Sequence records for all plants are at

<ftp://ftp.ncbi.nih.gov/refseq/release/plant/>

Genome records for plant organelles are located at

<ftp://ftp.ncbi.nih.gov/refseq/release/mitochondrion/>

<ftp://ftp.ncbi.nih.gov/refseq/release/plastid/>

Database Collaborators for Plant Genomes Central

ACWW	http://www.genome.arizona.edu/shotgun/rice/
GrainGenes	http://wheat.pw.usda.gov/index.shtml
Gramene	http://www.gramene.org/
MaizeDB/MaizeGDB	http://www.maizegdb.org/
Oryzabase	http://www.shigen.nig.ac.jp/rice/oryzabase/top/top.jsp
RGP	http://rgp.dna.affrc.go.jp/
RiceGenes	http://ukcrop.net/cgi-bin/WebAce/webace?db=RiceGenes
Solanaceae Genomics Network	http://www.sgn.cornell.edu/
Soybase	http://soybase.agron.iastate.edu/
TIGR	http://www.tigr.org/tdb/e2k1/ath1/
TAIR	http://arabidopsis.org/index.jsp
MIPS	http://mips.gsf.de/proj/thal/db/index.html