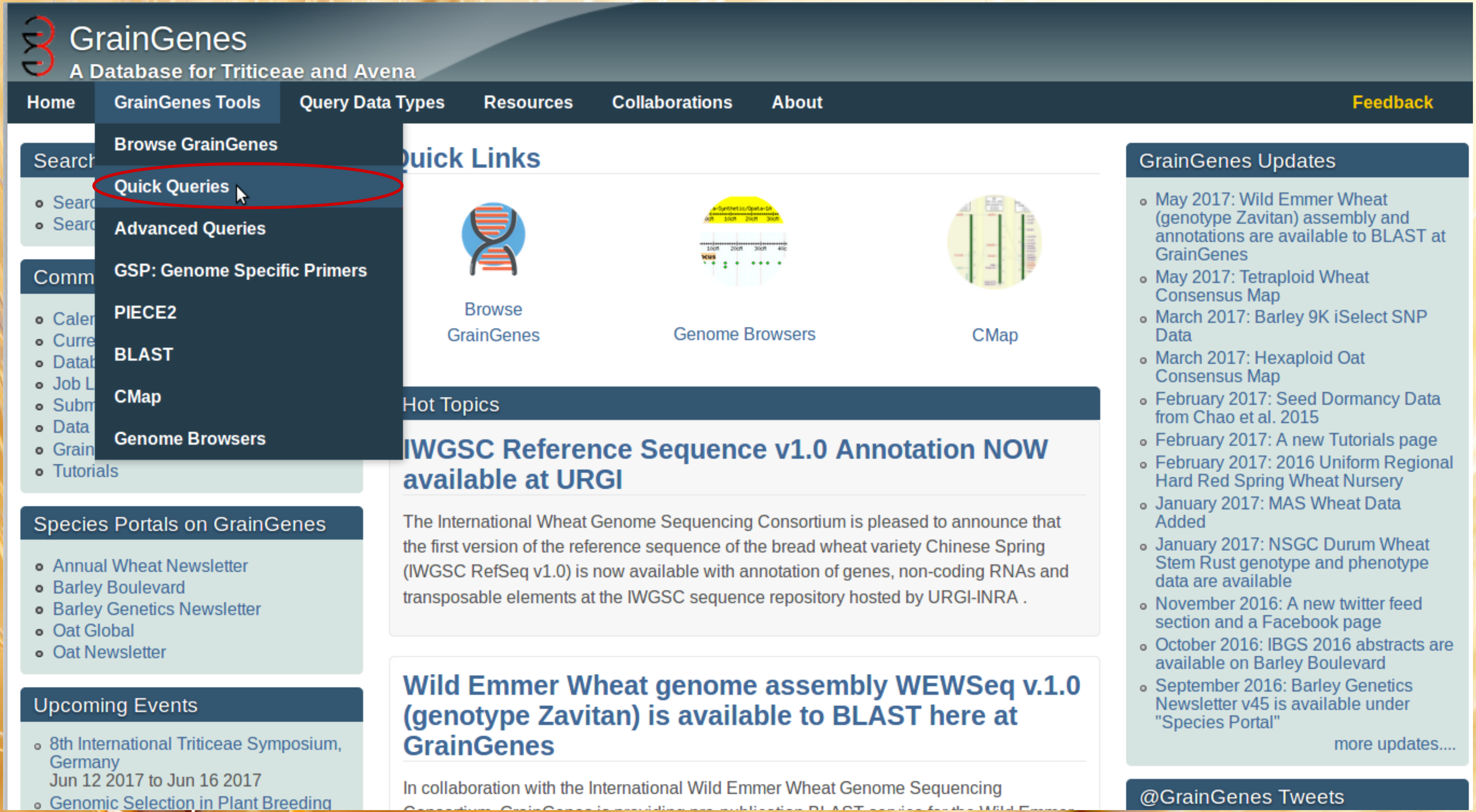
GrainGenes Tutorial: Finding Markers with Quick Queries

https://wheat.pw.usda.gov

Quick Queries can be accessed from the GrainGenes Tools dropdown menu



Species Portals on GrainGenes

- Annual Wheat Newsletter
- Barley Boulevard
- Barley Genetics Newsletter
- Oat Global
- Oat Newsletter

Upcoming Events

- 8th International Triticeae Symposium, Germany
  Jun 12 2017 to Jun 16 2017
- Genomic Selection in Plant Breeding

IWGSC Reference Sequence v1.0 Annotation NOW available at URGI

The International Wheat Genome Sequencing Consortium is pleased to announce that the first version of the reference sequence of the bread wheat variety Chinese Spring (IWGSC RefSeq v1.0) is now available with annotation of genes, non-coding RNAs and transposable elements at the IWGSC sequence repository hosted by URGI-INRA .

Wild Emmer Wheat genome assembly WEWSeq v.1.0 (genotype Zavitan) is available to BLAST here at GrainGenes

In collaboration with the International Wild Emmer Wheat Genome Sequencing

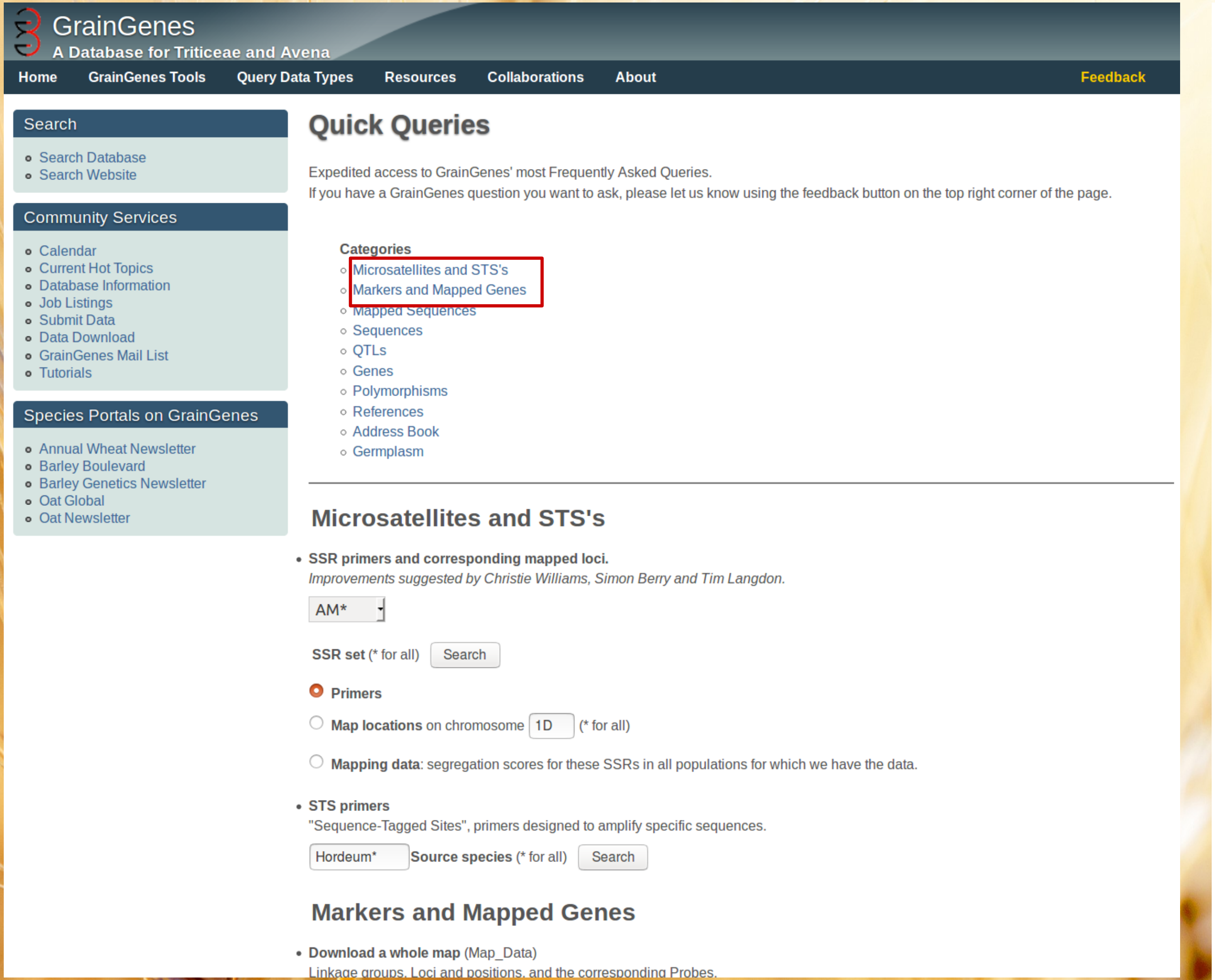
GrainGenes Updates

- May 2017: Wild Emmer Wheat (genotype Zavitan) assembly and annotations are available to BLAST at GrainGenes
- May 2017: Tetraploid Wheat Consensus Map
- March 2017: Barley 9K iSelect SNP Data
- March 2017: Hexaploid Oat Consensus Map
- February 2017: Seed Dormancy Data from Chao et al. 2015
- February 2017: A new Tutorials page
- February 2017: 2016 Uniform Regional Hard Red Spring Wheat Nursery
- January 2017: MAS Wheat Data Added
- January 2017: NSGC Durum Wheat Stem Rust genotype and phenotype data are available
- November 2016: A new twitter feed section and a Facebook page
- October 2016: IBGS 2016 abstracts are available on Barley Boulevard
- September 2016: Barley Genetics Newsletter v45 is available under "Species Portal"

more updates....

@GrainGenes Tweets

Quick Queries



GrainGenes
A Database for Triticeae and Avena

Home | GrainGenes Tools | Query Data Types | Resources | Collaborations | About | Feedback

Search
- Search Database
- Search Website

Community Services
- Calendar
- Current Hot Topics
- Database Information
- Job Listings
- Submit Data
- Data Download
- GrainGenes Mail List
- Tutorials

Species Portals on GrainGenes
- Annual Wheat Newsletter
- Barley Boulevard
- Barley Genetics Newsletter
- Oat Global
- Oat Newsletter

Quick Queries

Expedited access to GrainGenes' most Frequently Asked Queries.
If you have a GrainGenes question you want to ask, please let us know using the feedback button on the top right corner of the page.

Categories
- Microsatellites and STS's
- Markers and Mapped Genes
- Mapped Sequences
- Sequences
- QTLs
- Genes
- Polymorphisms
- References
- Address Book
- Germplasm

Microsatellites and STS's

- **SSR primers and corresponding mapped loci.**
  Improvements suggested by Christie Williams, Simon Berry and Tim Langdon.
  AM*
  SSR set (* for all) Search
  ● **Primers**
  ○ **Map locations** on chromosome 1D (* for all)
  ○ **Mapping data**: segregation scores for these SSRs in all populations for which we have the data.
- **STS primers**
  "Sequence-Tagged Sites", primers designed to amplify specific sequences.
  Hordeum* **Source species** (* for all) Search

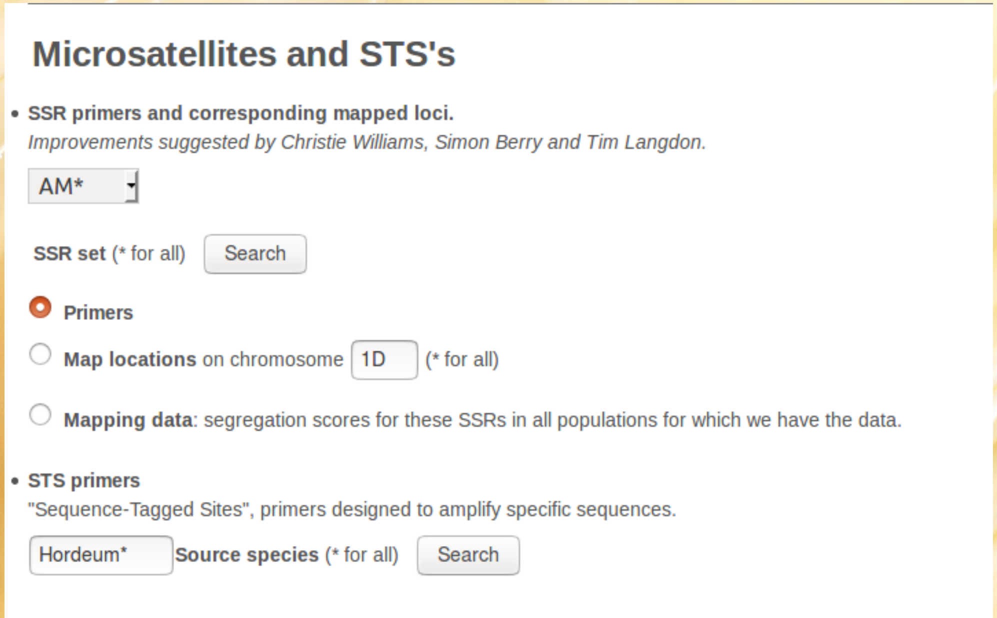
Markers and Mapped Genes

- **Download a whole map** (Map_Data)
  Linkage groups, Loci and positions, and the corresponding Probes.

- The Quick Queries page contains pre-made queries for common database searches
- At the top of the page is a list of categories for searches
- Clicking on one will navigate you to the corresponding part of the page
- For this tutorial, we will use the Microsetallites and STS's QuickQuery
- We'll also look at the Markers and Mapped Genes query

Quick Queries - Microsatellites and STS’s

- The Microsatellites and STS’s Quick Query allows you to get a list of SSR or STS probes that are available in GrainGenes
- One can also search for primer sequences for the probes, get map locations of mapped SSR probes, and segregation scores for SSRs, if available
- STS probes can be searched by search species
- SSR probes can be searched by prefix (AM, BCD, etc.)



Microsatellites and STS's

- **SSR primers and corresponding mapped loci.**
  Improvements suggested by Christie Williams, Simon Berry and Tim Langdon.

AM*

SSR set (* for all)

● **Primers**

○ **Map locations** on chromosome 1D (* for all)

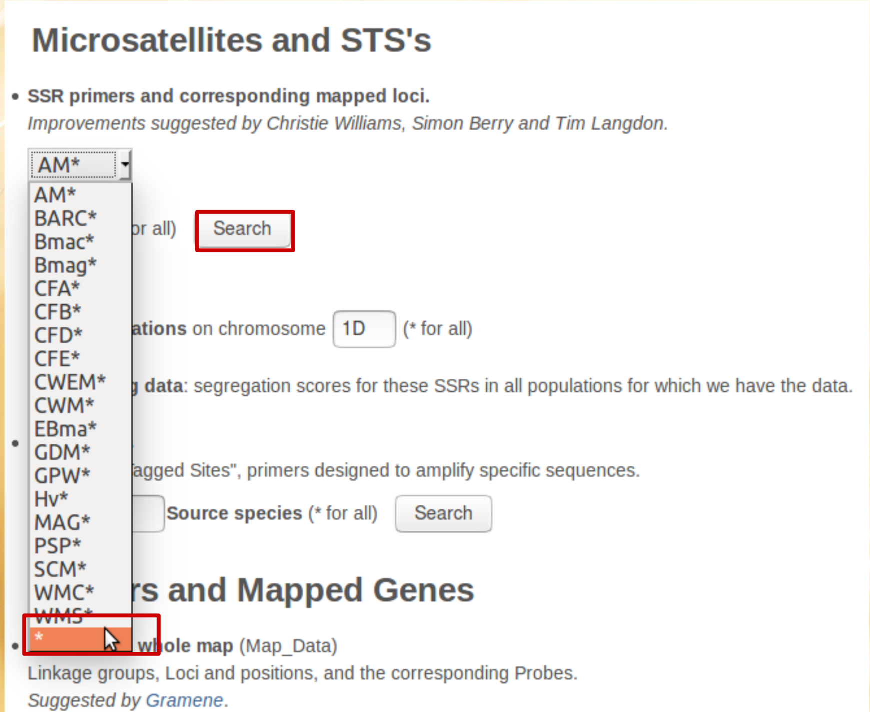
○ **Mapping data**: segregation scores for these SSRs in all populations for which we have the data.

- **STS primers**
  "Sequence-Tagged Sites", primers designed to amplify specific sequences.

Hordeum* **Source species** (* for all)

Quick Queries -
Microsatellites and STS’s

- As with most search tools in GrainGenes, the wildcard (*) can be used to ‘search all’
- For example, if we wanted to find all SSR probes and primer sequences in GrainGenes
- To do this, we would choose the ‘*’ option from the dropdown menu to select all prefix types
- To obtain the results of this query, click the Search button



Microsatellites and STS's

- **SSR primers and corresponding mapped loci.**
  Improvements suggested by Christie Williams, Simon Berry and Tim Langdon.

or all)

ations on chromosome 1D (* for all)

g data: segregation scores for these SSRs in all populations for which we have the data.

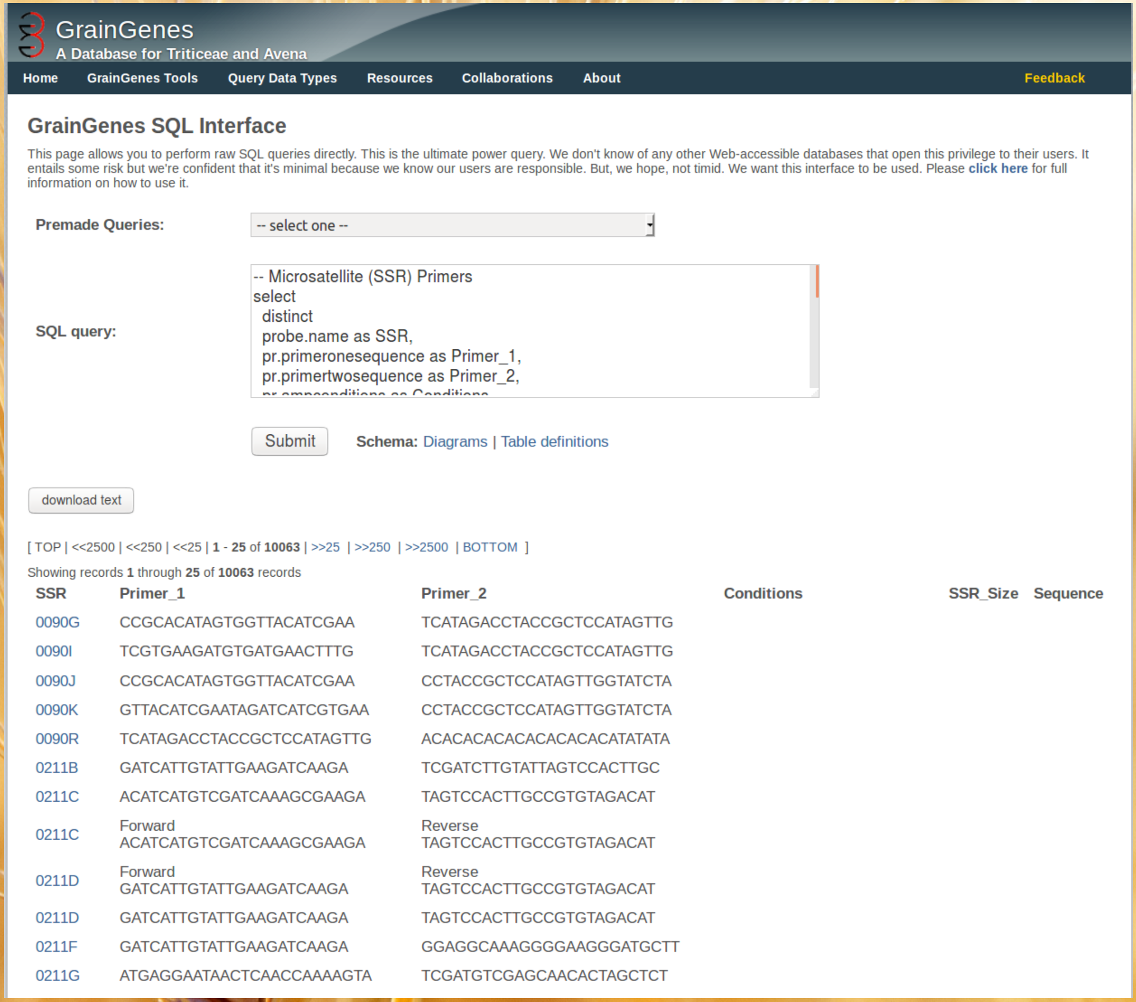
agged Sites", primers designed to amplify specific sequences.

Source species (* for all)

rs and Mapped Genes

- **whole map** (Map_Data)
  Linkage groups, Loci and positions, and the corresponding Probes.
  Suggested by Gramene.

Quick Queries - Query Results



GrainGenes SQL Interface

This page allows you to perform raw SQL queries directly. This is the ultimate power query. We don't know of any other Web-accessible databases that open this privilege to their users. It entails some risk but we're confident that it's minimal because we know our users are responsible. But, we hope, not timid. We want this interface to be used. Please **click here** for full information on how to use it.

Premade Queries: -- select one --

SQL query:

```
-- Microsatellite (SSR) Primers
select
  distinct
  probe.name as SSR,
  pr.primeronesequence as Primer_1,
  pr.primertwosequence as Primer_2,
```

Schema: Diagrams | Table definitions

[TOP | <<2500 | <<250 | <<25 | **1** - **25** of **10063** | >>25 | >>250 | >>2500 | BOTTOM]

Showing records **1** through **25** of **10063** records

SSR	Primer_1	Primer_2	Conditions	SSR_Size	Sequence
0090G	CCGCACATAGTGGTTACATCGAA	TCATAGACCTACCGCTCCATAGTTG			
0090I	TCGTGAAGATGTGATGAACTTTG	TCATAGACCTACCGCTCCATAGTTG			
0090J	CCGCACATAGTGGTTACATCGAA	CCTACCGCTCCATAGTTGGTATCTA			
0090K	GTTACATCGAATAGATCATCGTGAA	CCTACCGCTCCATAGTTGGTATCTA			
0090R	TCATAGACCTACCGCTCCATAGTTG	ACACACACACACACACACATATATA			
0211B	GATCATTGTATTGAAGATCAAGA	TCGATCTTGTATTAGTCCACTTGC			
0211C	ACATCATGTCGATCAAAGCGAAGA	TAGTCCACTTGCCGTGTAGACAT			
0211C	Forward ACATCATGTCGATCAAAGCGAAGA	Reverse TAGTCCACTTGCCGTGTAGACAT			
0211D	Forward GATCATTGTATTGAAGATCAAGA	Reverse TAGTCCACTTGCCGTGTAGACAT			
0211D	GATCATTGTATTGAAGATCAAGA	TAGTCCACTTGCCGTGTAGACAT			
0211F	GATCATTGTATTGAAGATCAAGA	GGAGGCAAAGGGGAAGGGATGCTT			
0211G	ATGAGGAATAACTCAACCAAAAGTA	TCGATGTCGAGCAACACTAGCTCT			

- We will then be brought to the GrainGenes SQL Interface
- Below the box containing our SQL query is a list of results
- The query found 10,063 SSR probes in GrainGenes, although there are some probes listed more than once because there was more than one primer pair listed with them
- For this tutorial, we will try to narrow down this list so that it contains only SSR markers that are mapped for hexaploid wheat (Triticum aestivum)
- In order to do this, we will need to make some edits to the SQL query

Quick Queries -
Editing the Query

Original Query:

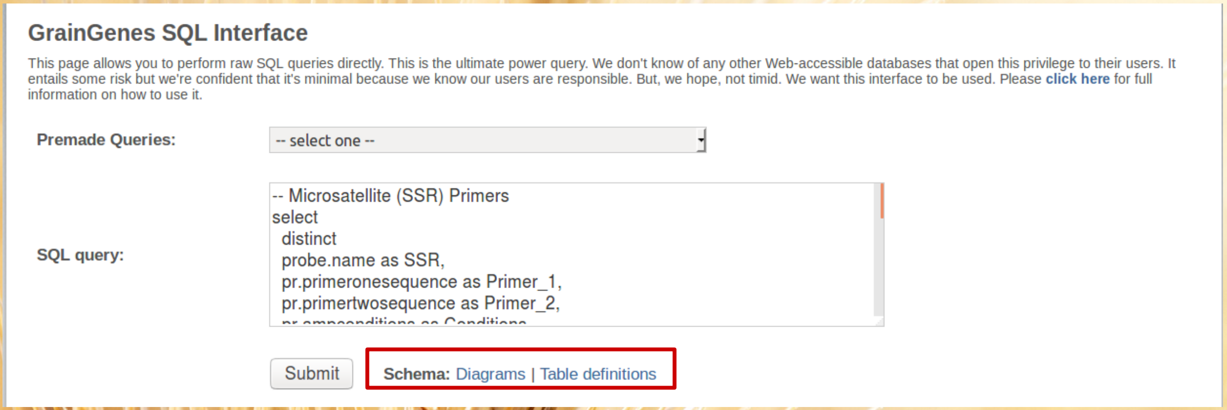
```
-- Microsatellite (SSR) Primers
select
  distinct
  probe.name as SSR,
  pr.primeronesequence as Primer_1,
  pr.primertwosequence as Primer_2,
  pr.ampconditions as Conditions,
  prSSR.size as SSR_Size,
  sequence.name as Sequence
from probe
  inner join probetype on probe.id = probetype.probeid
  -- get primer sequences and conditions:
  inner join probeprimer pr on probe.id = pr.probeid
    and pr.primeronesequence is not null
  -- get any size data:
  left join probeprimer prSSR on prSSR.probeid = probe.id
    and prSSR.sizetype = 'SSR_size'
  left join sequenceprobe on probe.id = sequenceprobe.probeid
  left join sequence on sequenceprobe.sequenceid = sequence.id
where probetype.type = 'SSR'
  and probe.name like '%'
```

Edited Query:

```
-- Microsatellite (SSR) Primers
select
  distinct
  probe.name as SSR,
  pr.primeronesequence as Primer_1,
  pr.primertwosequence as Primer_2,
  pr.ampconditions as Conditions,
  prSSR.size as SSR_Size
    -- sequence.name as Sequence
from probe
   inner join probesourcespecies on probe.id =
   probesourcespecies.probeid
   inner join species on probesourcespecies.speciesid = species.id
  inner join probetype on probe.id = probetype.probeid
  -- get primer sequences and conditions:
  inner join probeprimer pr on probe.id = pr.probeid
    and pr.primeronesequence is not null
  -- get any size data:
  left join probeprimer prSSR on prSSR.probeid = probe.id
    and prSSR.sizetype = 'SSR_size'
--left join sequenceprobe on probe.id = sequenceprobe.probeid
--left join sequence on sequenceprobe.sequenceid = sequence.id
where probetype.type = 'SSR'
  and probe.name like '%'
    and species.name='Triticum aestivum';
```

- With the edits shown above, we are selecting only the SSR markers for wheat. To simplify our results, we are also removing Sequence from our ‘select’ statement

Quick Queries - Table Diagrams



GrainGenes SQL Interface

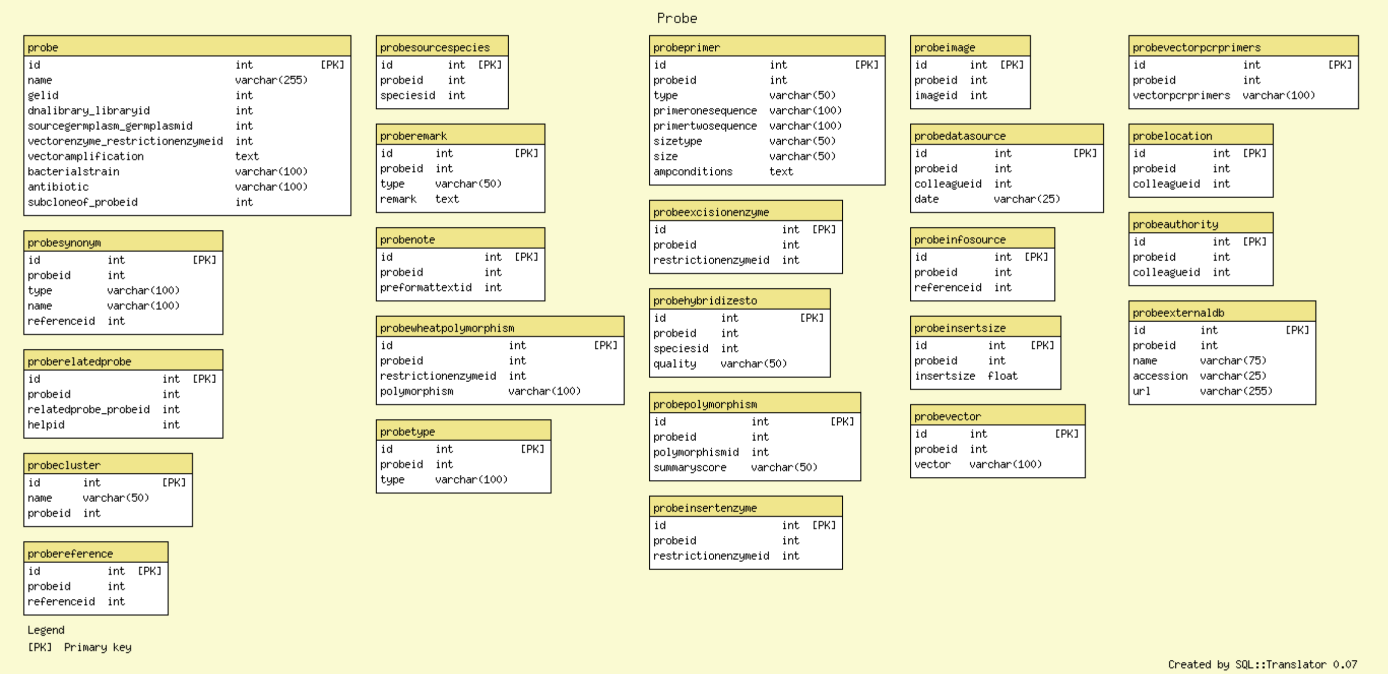
This page allows you to perform raw SQL queries directly. This is the ultimate power query. We don't know of any other Web-accessible databases that open this privilege to their users. It entails some risk but we're confident that it's minimal because we know our users are responsible. But, we hope, not timid. We want this interface to be used. Please **click here** for full information on how to use it.

Premade Queries: -- select one --

SQL query:

```
-- Microsatellite (SSR) Primers
select
  distinct
  probe.name as SSR,
  pr.primeronesequence as Primer_1,
  pr.primertwosequence as Primer_2,
```

Schema: Diagrams | Table definitions



Probe

probe		
id	int	[PK]
name	varchar(255)	
gelid	int	
dnalibrary_libraryid	int	
sourcegermplasm_germplasmid	int	
vectorenzyme_restrictionenzymeid	int	
vectoramplification	text	
bacterialstrain	varchar(100)	
antibiotic	varchar(100)	
subcloneof_probeid	int	

probesynonym		
id	int	[PK]
probeid	int	
type	varchar(100)	
name	varchar(100)	
referenceid	int	

proberelatedprobe		
id	int	[PK]
probeid	int	
relatedprobe_probeid	int	
helpid	int	

probecluster		
id	int	[PK]
name	varchar(50)	
probeid	int	

probereference		
id	int	[PK]
probeid	int	
referenceid	int	

Legend
[PK] Primary key

probesourcespecies		
id	int	[PK]
probeid	int	
speciesid	int	

proberemark		
id	int	[PK]
probeid	int	
type	varchar(50)	
remark	text	

probenote		
id	int	[PK]
probeid	int	
preformattextid	int	

probewheatpolymorphism		
id	int	[PK]
probeid	int	
restrictionenzymeid	int	
polymorphism	varchar(100)	

probetype		
id	int	[PK]
probeid	int	
type	varchar(100)	

probeprimer		
id	int	[PK]
probeid	int	
type	varchar(50)	
primeronesequence	varchar(100)	
primertwosequence	varchar(100)	
sizetype	varchar(50)	
size	varchar(50)	
ampconditions	text	

probeexcisionenzyme		
id	int	[PK]
probeid	int	
restrictionenzymeid	int	

probehybridizesto		
id	int	[PK]
probeid	int	
speciesid	int	
quality	varchar(50)	

probepolymorphism		
id	int	[PK]
probeid	int	
polymorphismid	int	
summaryscore	varchar(50)	

probeinsertenzyme		
id	int	[PK]
probeid	int	
restrictionenzymeid	int	

probeimage		
id	int	[PK]
probeid	int	
imageid	int	

probedatasource		
id	int	[PK]
probeid	int	
colleagueid	int	
date	varchar(25)	

probeinfosource		
id	int	[PK]
probeid	int	
referenceid	int	

probeinsertsize		
id	int	[PK]
probeid	int	
insertsize	float	

probevector		
id	int	[PK]
probeid	int	
vector	varchar(100)	

probevectorpcrprimers		
id	int	[PK]
probeid	int	
vectorpcrprimers	varchar(100)	

probelocation		
id	int	[PK]
probeid	int	
colleagueid	int	

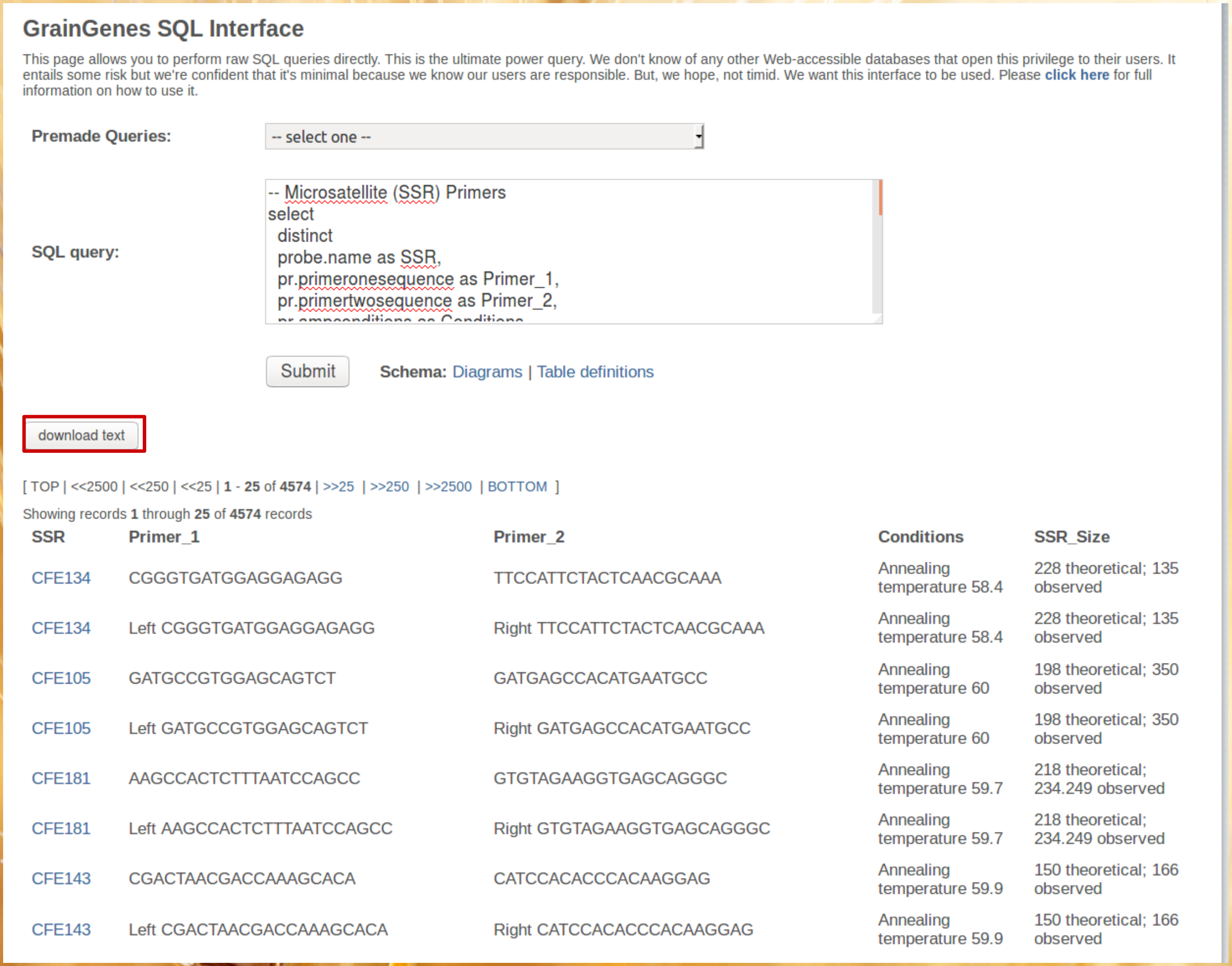
probeauthority		
id	int	[PK]
probeid	int	
colleagueid	int	

probeexternaldb		
id	int	[PK]
probeid	int	
name	varchar(75)	
accession	varchar(25)	
url	varchar(255)	

Created by SQL::Translator 0.07

- To better understand the edits made in the previous slide, information on the GrainGenes database schema is available underneath the SQL query box
- By clicking the Diagrams link and then selecting the data type that you would like to query (i.e probe), you can view an image of the database tables and how they are interconnected (this will require some previous understanding of relational databases)

Quick Queries -
Edited Query Results



GrainGenes SQL Interface

This page allows you to perform raw SQL queries directly. This is the ultimate power query. We don't know of any other Web-accessible databases that open this privilege to their users. It entails some risk but we're confident that it's minimal because we know our users are responsible. But, we hope, not timid. We want this interface to be used. Please **click here** for full information on how to use it.

Premade Queries: -- select one --

SQL query:

```
-- Microsatellite (SSR) Primers
select
  distinct
  probe.name as SSR,
  pr.primeronesequence as Primer_1,
  pr.primertwosequence as Primer_2,
```

Schema: Diagrams | Table definitions

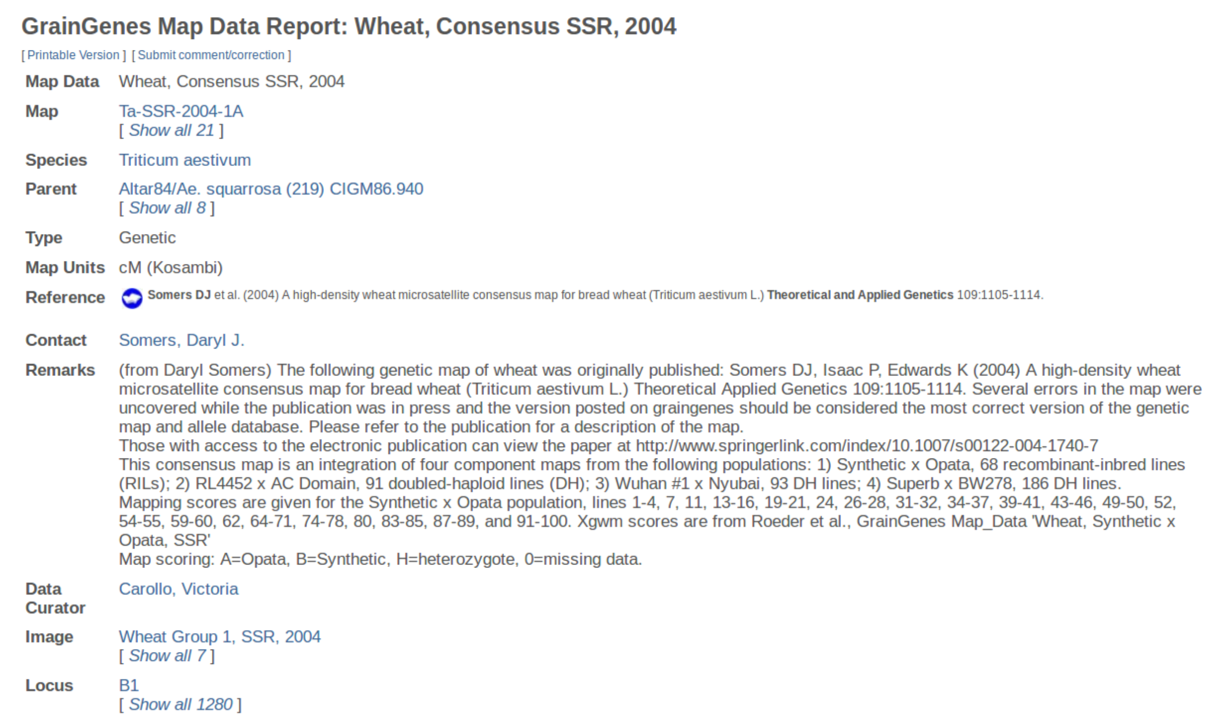
[TOP | <<2500 | <<250 | <<25 | **1** - **25** of **4574** | >>25 | >>250 | >>2500 | BOTTOM]

Showing records **1** through **25** of **4574** records

SSR	Primer_1	Primer_2	Conditions	SSR_Size
CFE134	CGGGTGATGGAGGAGAGG	TTCCATTCTACTCAACGCAAA	Annealing temperature 58.4	228 theoretical; 135 observed
CFE134	Left CGGGTGATGGAGGAGAGG	Right TTCCATTCTACTCAACGCAAA	Annealing temperature 58.4	228 theoretical; 135 observed
CFE105	GATGCCGTGGAGCAGTCT	GATGAGCCACATGAATGCC	Annealing temperature 60	198 theoretical; 350 observed
CFE105	Left GATGCCGTGGAGCAGTCT	Right GATGAGCCACATGAATGCC	Annealing temperature 60	198 theoretical; 350 observed
CFE181	AAGCCACTCTTTAATCCAGCC	GTGTAGAAGGTGAGCAGGGC	Annealing temperature 59.7	218 theoretical; 234.249 observed
CFE181	Left AAGCCACTCTTTAATCCAGCC	Right GTGTAGAAGGTGAGCAGGGC	Annealing temperature 59.7	218 theoretical; 234.249 observed
CFE143	CGACTAACGACCAAAGCACA	CATCCACACCCACAAGGAG	Annealing temperature 59.9	150 theoretical; 166 observed
CFE143	Left CGACTAACGACCAAAGCACA	Right CATCCACACCCACAAGGAG	Annealing temperature 59.9	150 theoretical; 166 observed

- After the proper edits have been made, our query returns 4,574 results of SSR markers for hexaploid wheat!
- The results can be downloaded into a text file by clicking the ‘download text’ button

Quick Queries - Markers and Mapped Genes



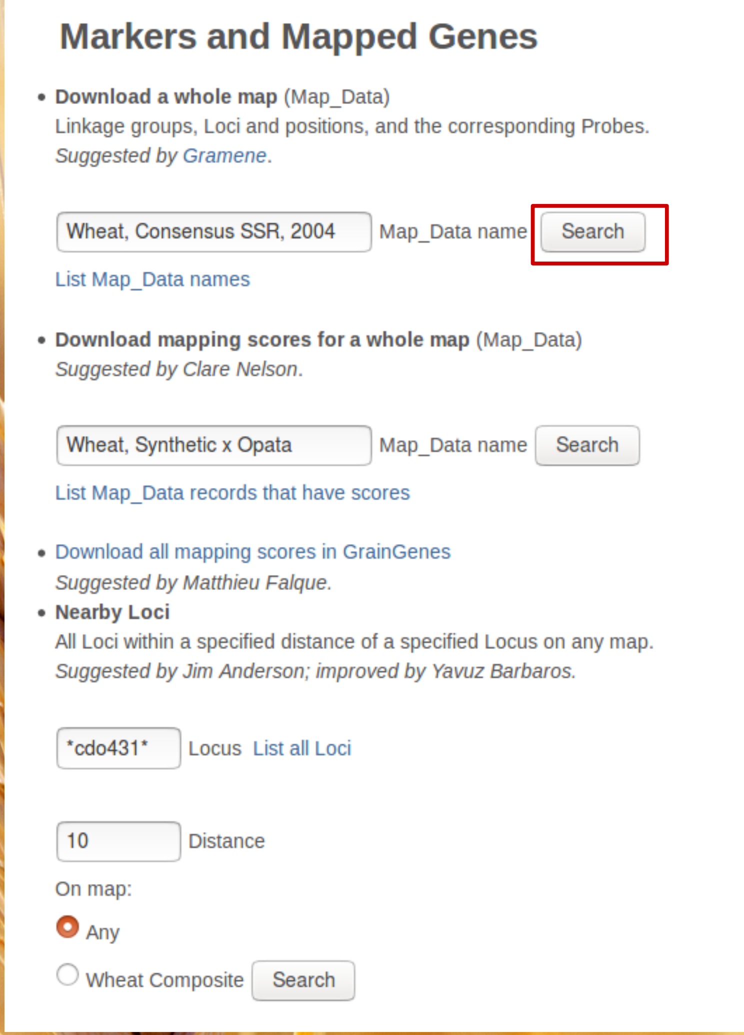
GrainGenes Map Data Report: Wheat, Consensus SSR, 2004

[Printable Version] [Submit comment/correction]

Map Data	Wheat, Consensus SSR, 2004
Map	Ta-SSR-2004-1A [*Show all 21*]
Species	Triticum aestivum
Parent	Altar84/Ae. squarrosa (219) CIGM86.940 [*Show all 8*]
Type	Genetic
Map Units	cM (Kosambi)
Reference	**Somers DJ** et al. (2004) A high-density wheat microsatellite consensus map for bread wheat (Triticum aestivum L.) **Theoretical and Applied Genetics** 109:1105-1114.
Contact	Somers, Daryl J.
Remarks	(from Daryl Somers) The following genetic map of wheat was originally published: Somers DJ, Isaac P, Edwards K (2004) A high-density wheat microsatellite consensus map for bread wheat (Triticum aestivum L.) Theoretical Applied Genetics 109:1105-1114. Several errors in the map were uncovered while the publication was in press and the version posted on graingenes should be considered the most correct version of the genetic map and allele database. Please refer to the publication for a description of the map. Those with access to the electronic publication can view the paper at http://www.springerlink.com/index/10.1007/s00122-004-1740-7 This consensus map is an integration of four component maps from the following populations: 1) Synthetic x Opata, 68 recombinant-inbred lines (RILs); 2) RL4452 x AC Domain, 91 doubled-haploid lines (DH); 3) Wuhan #1 x Nyubai, 93 DH lines; 4) Superb x BW278, 186 DH lines. Mapping scores are given for the Synthetic x Opata population, lines 1-4, 7, 11, 13-16, 19-21, 24, 26-28, 31-32, 34-37, 39-41, 43-46, 49-50, 52, 54-55, 59-60, 62, 64-71, 74-78, 80, 83-85, 87-89, and 91-100. Xgwm scores are from Roeder et al., GrainGenes Map_Data 'Wheat, Synthetic x Opata, SSR' Map scoring: A=Opata, B=Synthetic, H=heterozygote, 0=missing data.
Data Curator	Carollo, Victoria
Image	Wheat Group 1, SSR, 2004 [*Show all 7*]
Locus	B1 [*Show all 1280*]

- There are multiple ways that Quick Queries can be used, depending on the information that you are looking for
- For example, say that we want information on probes that were used to construct a particular map in GrainGenes, such as the 2004 Wheat SSR Consensus map (Somers et al. 2004).
- First, find the name of the map data set that you wish to search as it is recorded in GrainGenes using the Database Browser. Our example is listed as ‘Wheat, Consensus SSR, 2004’

Quick Queries -
Markers and Mapped Genes



Markers and Mapped Genes

- **Download a whole map** (Map_Data)
  Linkage groups, Loci and positions, and the corresponding Probes.
  Suggested by Gramene.

Wheat, Consensus SSR, 2004 Map_Data name

List Map_Data names

- **Download mapping scores for a whole map** (Map_Data)
  Suggested by Clare Nelson.

Wheat, Synthetic x Opata Map_Data name

List Map_Data records that have scores

- Download all mapping scores in GrainGenes
  Suggested by Matthieu Falque.
- **Nearby Loci**
  All Loci within a specified distance of a specified Locus on any map.
  Suggested by Jim Anderson; improved by Yavuz Barbaros.

cdo431 Locus List all Loci

10 Distance

On map:

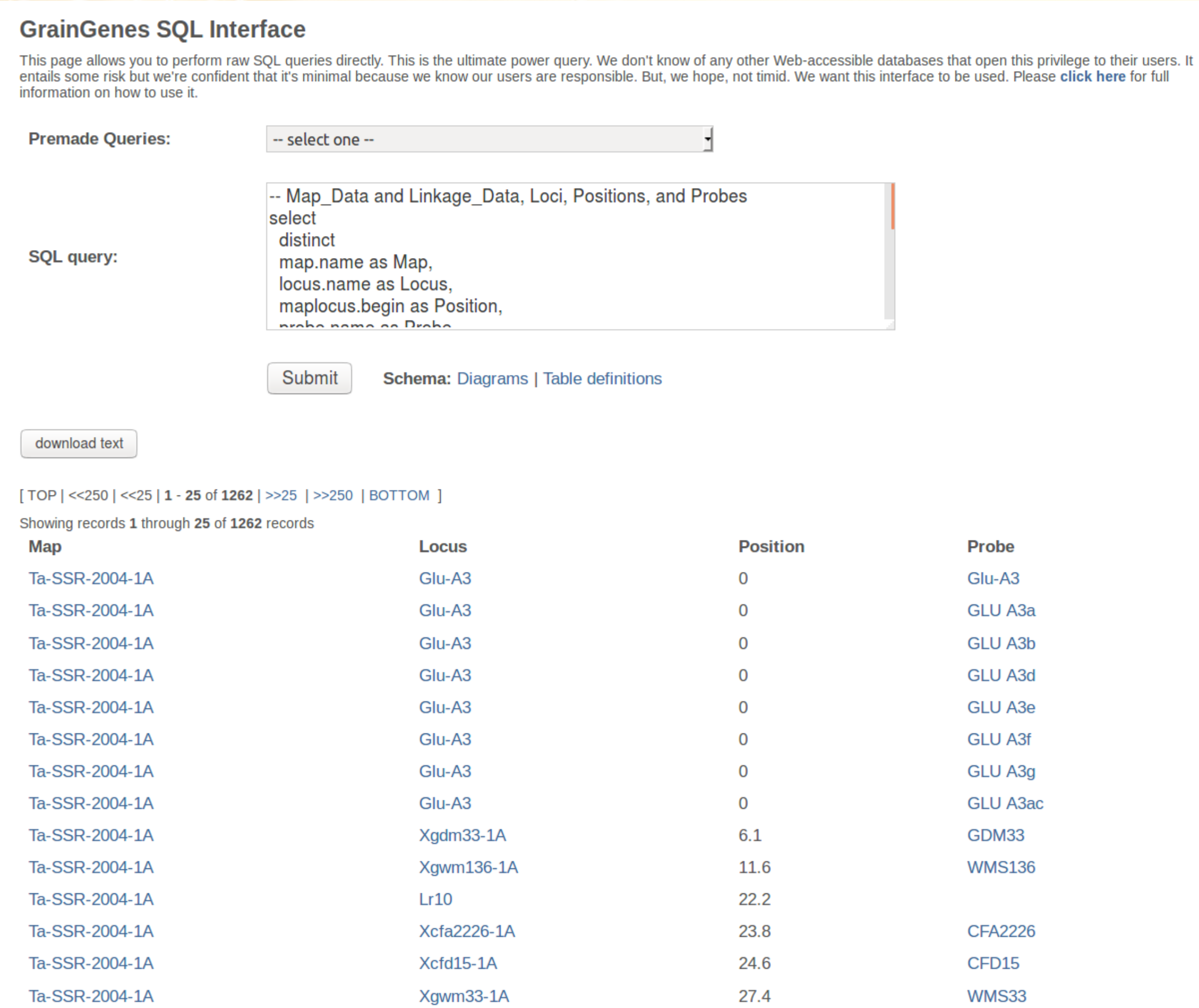
● Any

○ Wheat Composite

- Back on the Quick Queries page, click the link for Markers and Mapped Genes
- In the ‘Download a whole map’ search box, type the name of the map set of interest. In this example, that would be ‘Wheat, Consensus SSR, 2004’
- This query allows you to find linkage groups, loci, map positions, and probes for a particular map set
- You can also obtain mapping scores for a whole map set, or identify loci that are within a certain mapping distance to a locus of interest for maps within GrainGenes
- Click the ‘Search’ button to obtain the results

Quick Queries -
Markers and Mapped Genes

- Listed in the results are all of the loci in the selected map set, as well as their linkage group, map position, and corresponding probe
- This information is also downloadable with the ‘download text’ button



GrainGenes SQL Interface

This page allows you to perform raw SQL queries directly. This is the ultimate power query. We don't know of any other Web-accessible databases that open this privilege to their users. It entails some risk but we're confident that it's minimal because we know our users are responsible. But, we hope, not timid. We want this interface to be used. Please **click here** for full information on how to use it.

Premade Queries: -- select one --

SQL query:

```
-- Map_Data and Linkage_Data, Loci, Positions, and Probes
select
 distinct
 map.name as Map,
 locus.name as Locus,
 maplocus.begin as Position,
```

Schema: Diagrams | Table definitions

[TOP | <<250 | <<25 | **1** - **25** of **1262** | >>25 | >>250 | BOTTOM]

Showing records **1** through **25** of **1262** records

Map	Locus	Position	Probe
Ta-SSR-2004-1A	Glu-A3	0	Glu-A3
Ta-SSR-2004-1A	Glu-A3	0	GLU A3a
Ta-SSR-2004-1A	Glu-A3	0	GLU A3b
Ta-SSR-2004-1A	Glu-A3	0	GLU A3d
Ta-SSR-2004-1A	Glu-A3	0	GLU A3e
Ta-SSR-2004-1A	Glu-A3	0	GLU A3f
Ta-SSR-2004-1A	Glu-A3	0	GLU A3g
Ta-SSR-2004-1A	Glu-A3	0	GLU A3ac
Ta-SSR-2004-1A	Xgdm33-1A	6.1	GDM33
Ta-SSR-2004-1A	Xgwm136-1A	11.6	WMS136
Ta-SSR-2004-1A	Lr10	22.2	
Ta-SSR-2004-1A	Xcfa2226-1A	23.8	CFA2226
Ta-SSR-2004-1A	Xcfd15-1A	24.6	CFD15
Ta-SSR-2004-1A	Xgwm33-1A	27.4	WMS33

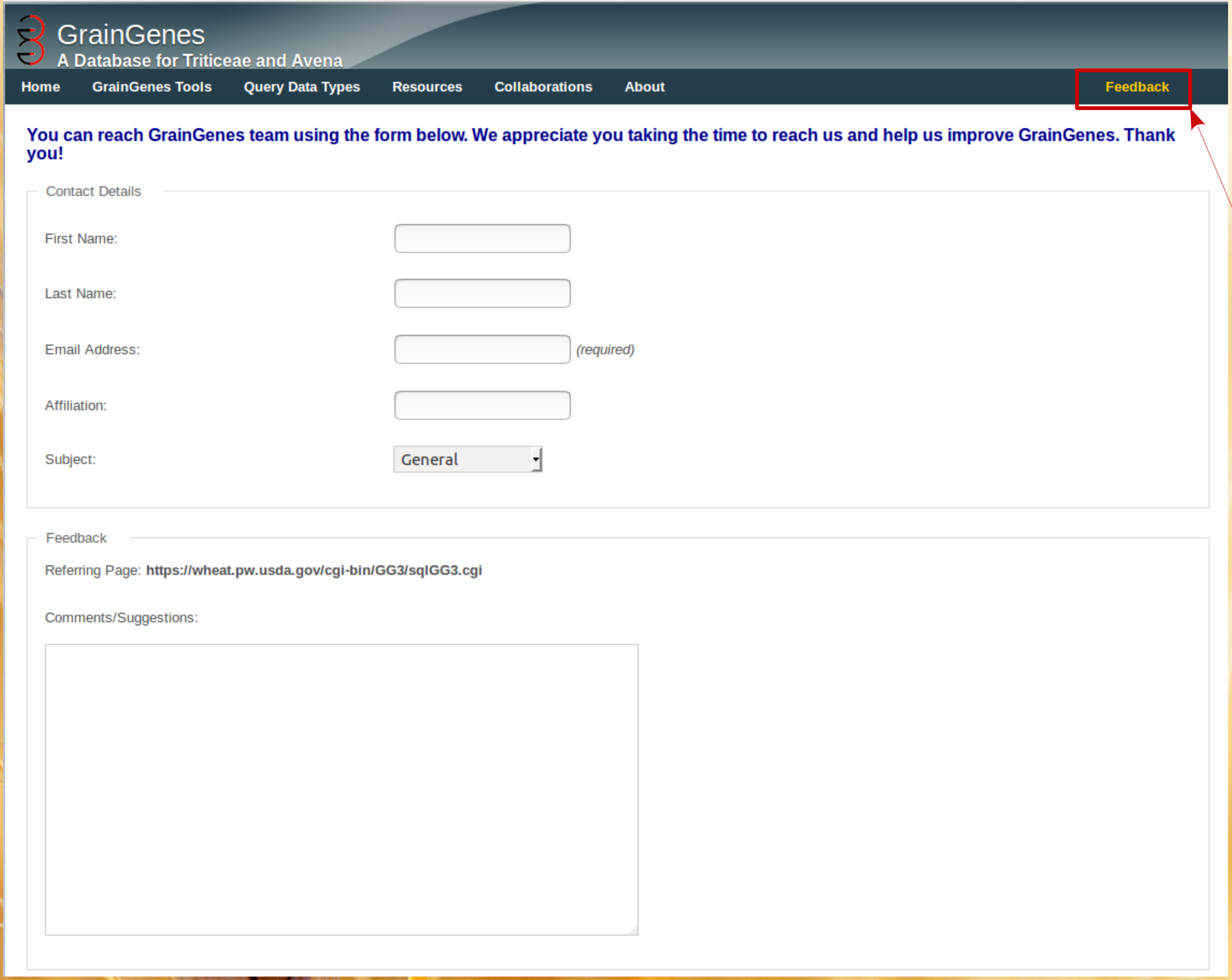
Advanced Queries-
Probes to Sequences

GrainGenes Tutorial: Obtaining FASTA Sequences from Markers

- To learn how to obtain FASTA sequences from a list of probe names, check out our video Tutorial here:

  https://www.youtube.com/watch?v=EC4NOo03_5g&t=1s

Feedback



You can reach GrainGenes team using the form below. We appreciate you taking the time to reach us and help us improve GrainGenes. Thank you!

Contact Details

First Name:	
Last Name:	
Email Address:	*(required)*
Affiliation:	
Subject:	General

Feedback

Referring Page: **https://wheat.pw.usda.gov/cgi-bin/GG3/sqlGG3.cgi**

Comments/Suggestions:

- If you have any questions, have suggestions for new Quick Queries, want to submit your data to GrainGenes, or anything else, please contact us!
- Our Feedback button is at the top right corner of every page, and we'd love the hear from you!