



RTWG Workshop  
2/25/06

# Using Gramene in a Breeding Program



# What is Gramene

- A resource for Comparative Grass Genomics
- Adds value to public info
  - Automated and curated relationships
  - Web displays

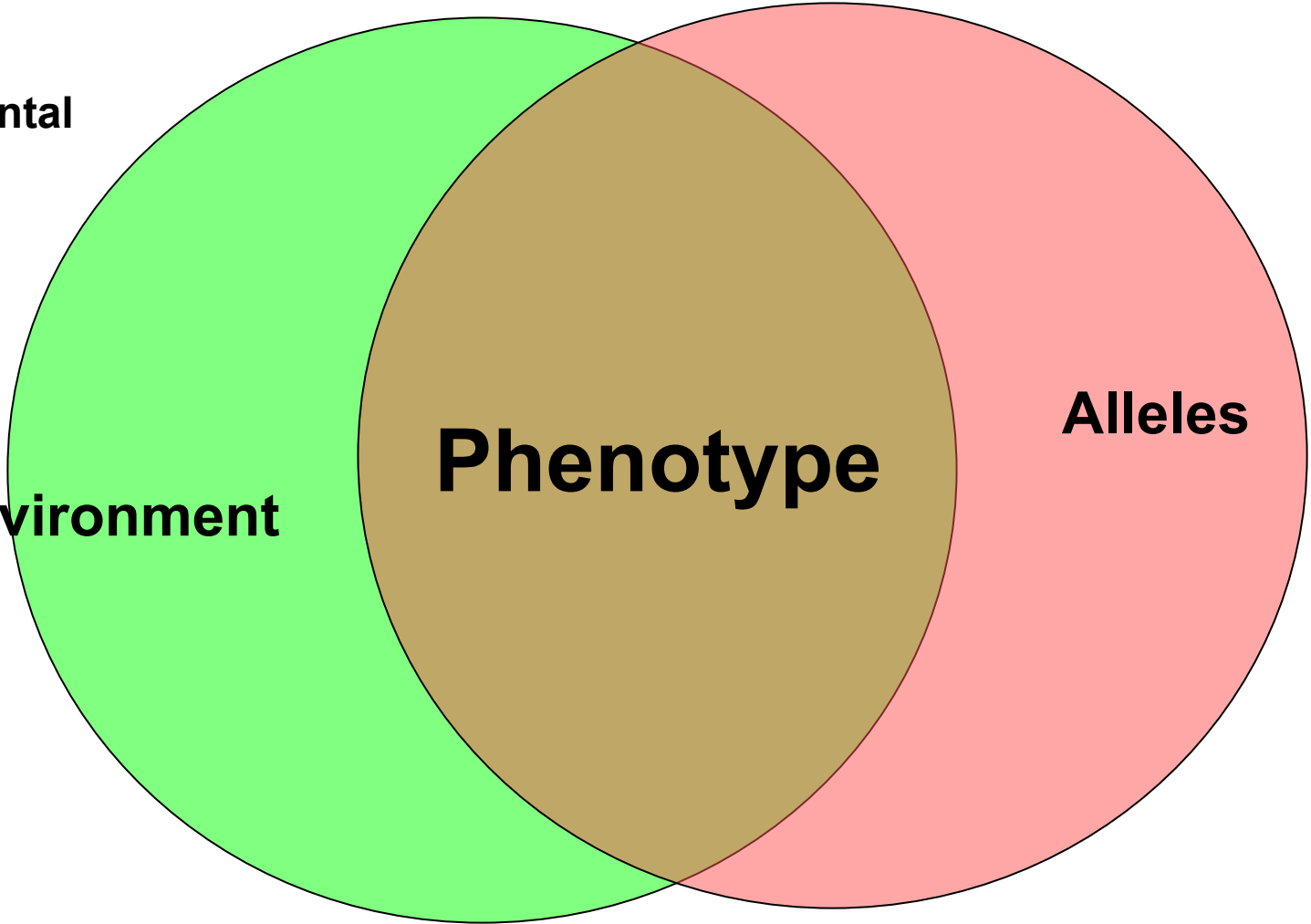


# Phenotype = Genotype X Environment

Developmental stage

Length of exposure

Intensity



# Breeders are concerned with the selection of favorable phenotypes



They have always endeavored to 'read' genome sequence by peering through the lens of the phenotype.

Molecular markers offer a new lens for identifying genes associated with phenotypes of interest.



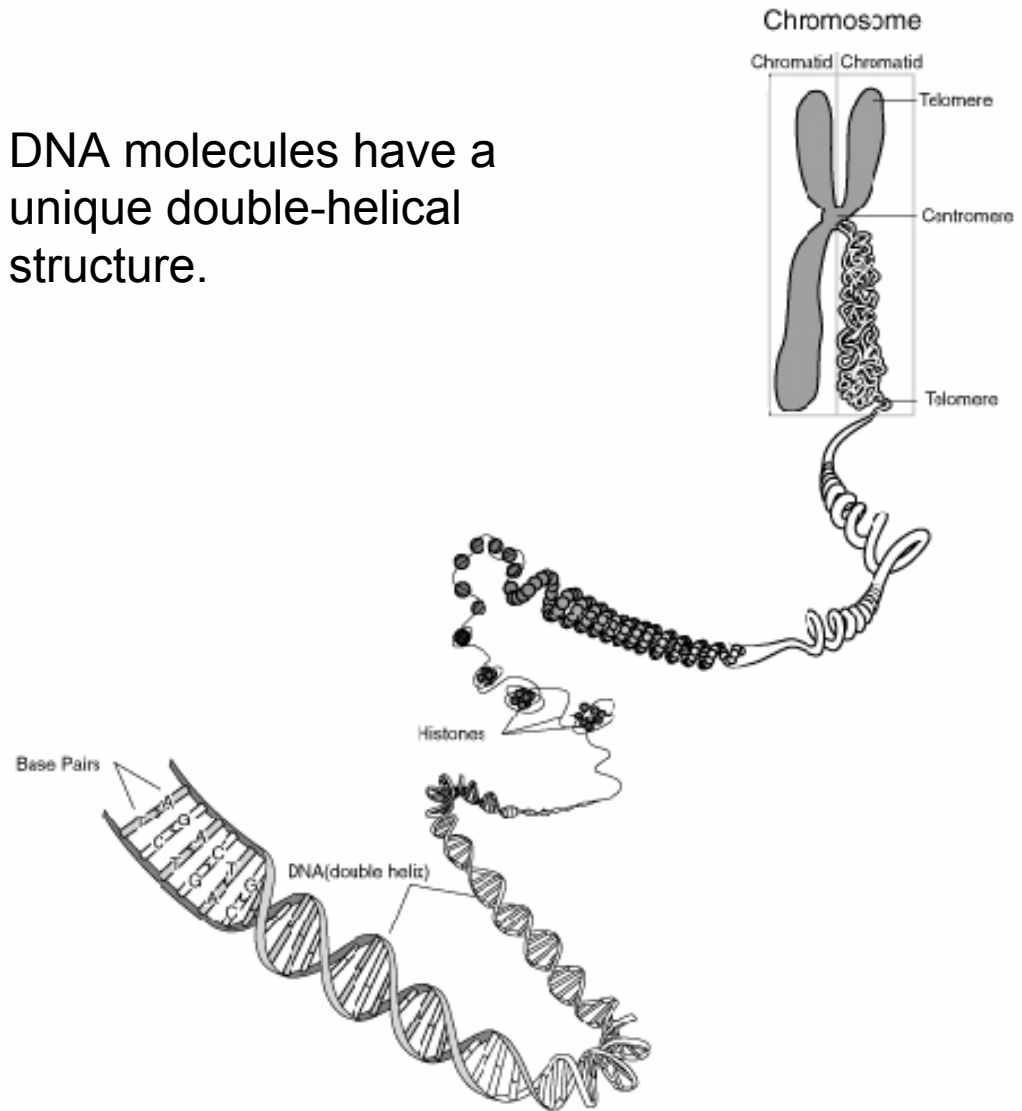
# Example MAS

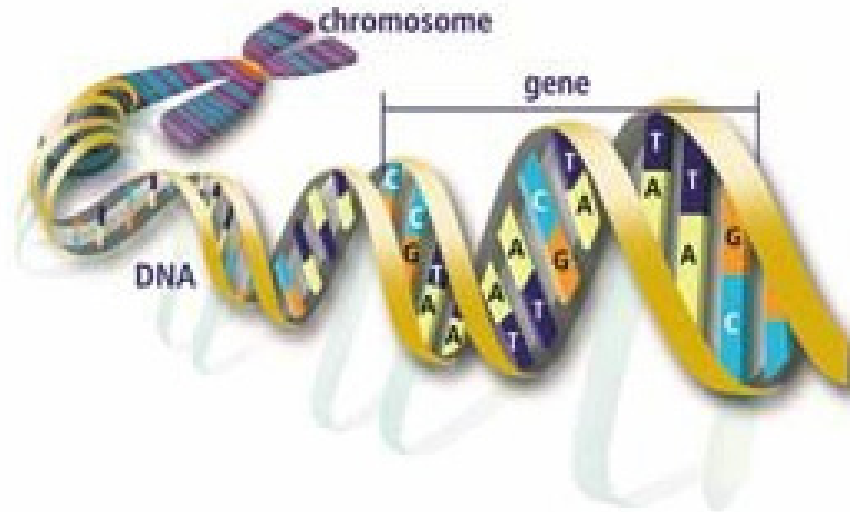
- Peter Langridge – Wheat – At PAG reported 5-7 years reduced time from first cross to crop release.
  - With a 5 fold increase in success of choices.



# Genotype

DNA molecules have a unique double-helical structure.





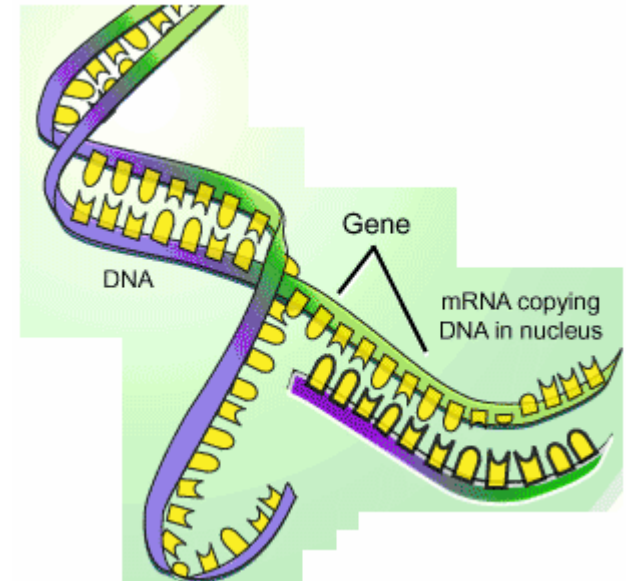
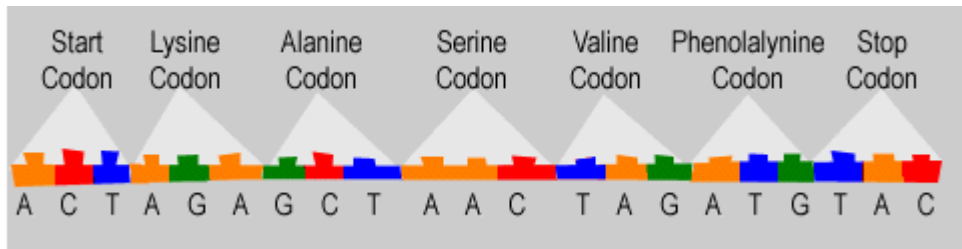
# Genes

**LOC\_Os11g47740**

```
ATGTCAGTGCAAACCTCTCCTTTCTTCTTCTCCTCCTCCCTCTCTCTCTCCCTTCTCTCTCTCCCGTGGCCTCT
CTCTCCACCTTGGCCAGAGCAATGGCGGCAGCGGTGGCAGACGACGTCGGAGCATCGGTGCGCCGTCGT
ACCAAAGTAGCGAGAGCACGCAGGACGTGGAGAAGGGGAGAGAGAGATTGAGAGAGTCGTCCTCGTCA
CCGCCGGTGCCACGCGCAAGAAGAAAAGGAAACCGAGGATAAGGGGAGAGAGAGAGTCGTCCCGTCA
CTGCCGGTGCCGCCGTCCTAGAGATCCCTCGCGAGCTGTGTCCCGCCACCACCGCCGTCCACCACCGG
TCGCCGTCCCGTGCTCCTTCGCTGTCCCCGGGCGGAGAAGAAAGGAAGACCGGGGAGAGAGGGGGCA
GAGAGAGCCATCCGTCACGGCGAGCTGCGCCGCCCTGGAGCGTTGCGATGCTACTTCTGGCTTCTTCT
GCGATGCTACTGAGGATGAAGAAAGCCTCTGTCTAGGTAATTCTTTTTCTTCTTTTCTCTCACTCACTCCAT
CAGCTGCTAGCTTTGCTAGCTAGGAGTAGCTAGCTGCTTGTGCTCTTGCTGTGCGAGAGGAGGAAGAAAG
ACTCGAGGTGGTTGGCTGTGCTGTTGTGCATGGCTGCATTCAGAGCCCCGACGGTGATGTCATCGACAG
CGTGCCGCTGCACCTGCAACCGGCGTTTCGACCACCCGAAGCTCATTCAATTTTGTTCATTTTGCAGTGA
TCGTTTCTATCTCCAGCGACGTTTCTTGATCGTGTTTGTGATTGTGTGCGCGTGACGGTTGAGCCGGAGC
CAGAGGAGAGGCCCAAGGTCGGTGGCGCGGCGGCGGCATGGGGAGGCGGCGGAGGAGGAGGTGCG
TGTTCCCGATGGCGTGGACGGACGACGACGAGTTGTGCCCGGAGGGGACGGTGCCGGTGCCGGCAGAC
GACGAAGCGCGACGTGCTGAGGTCCAGCTCCTCTCCCTGTTTGGGGATGAAGCAGCCTCGTGCCGGCG
TGCCGCTGGTGTGCTCCGCGTGA
```

# Codons

Using the 4 nucleotides A, T, C, G  
(**A**denine, **G**uanine, **T**hymine,  
**C**ytosine, of DNA, there are 64  
possible codons



Transcription by RNA: T → U (Uracil)



Table 1: RNA codon table

This table shows the 64 codons and the amino acid each codon codes for.

		2nd base			
		U	C	A	G
1st base	U	UUU (Phe/F)Phenylalanine	UCU (Ser/S)Serine	UAU (Tyr/Y)Tyrosine	UGU (Cys/C)Cysteine
		UUC (Phe/F)Phenylalanine	UCC (Ser/S)Serine	UAC (Tyr/Y)Tyrosine	UGC (Cys/C)Cysteine
		UUA (Leu/L)Leucine	UCA (Ser/S)Serine	UAA Ochre (Stop)	UGA Opal (Stop)
		UUG (Leu/L)Leucine, Start	UCG (Ser/S)Serine	UAG Amber (Stop)	UGG (Trp/W)Tryptophan
	C	CUU (Leu/L)Leucine	CCU (Pro/P)Proline	CAU (His/H)Histidine	CGU (Arg/R)Arginine
		CUC (Leu/L)Leucine	CCC (Pro/P)Proline	CAC (His/H)Histidine	CGC (Arg/R)Arginine
		CUA (Leu/L)Leucine	CCA (Pro/P)Proline	CAA (Gln/Q)Glutamine	CGA (Arg/R)Arginine
		CUG (Leu/L)Leucine, Start	CCG (Pro/P)Proline	CAG (Gln/Q)Glutamine	CGG (Arg/R)Arginine
	A	AUU (Ile/I)Isoleucine, Start <sup>2</sup>	ACU (Thr/T)Threonine	AAU (Asn/N)Asparagine	AGU (Ser/S)Serine
		AUC (Ile/I)Isoleucine	ACC (Thr/T)Threonine	AAC (Asn/N)Asparagine	AGC (Ser/S)Serine
		AUA (Ile/I)Isoleucine	ACA (Thr/T)Threonine	AAA (Lys/K)Lysine	AGA (Arg/R)Arginine
		AUG (Met/M)Methionine, Start <sup>1</sup>	ACG (Thr/T)Threonine	AAG (Lys/K)Lysine	AGG (Arg/R)Arginine
	G	GUU (Val/V)Valine	GCU (Ala/A)Alanine	GAU (Asp/D)Aspartic acid	GGU (Gly/G)Glycine
		GUC (Val/V)Valine	GCC (Ala/A)Alanine	GAC (Asp/D)Aspartic acid	GGC (Gly/G)Glycine
		GUA (Val/V)Valine	GCA (Ala/A)Alanine	GAA (Glu/E)Glutamic acid	GGA (Gly/G)Glycine
		GUG (Val/V)Valine, Start <sup>2</sup>	GCG (Ala/A)Alanine	GAG (Glu/E)Glutamic acid	GGG (Gly/G)Glycine

# Amino Acids

<b>G</b>	Glycine	Gly	<b>P</b>	Proline	Pro
<b>A</b>	Alanine	Ala	<b>V</b>	Valine	Val
<b>L</b>	Leucine	Leu	<b>I</b>	Isoleucine	Ile
<b>M</b>	Methionine	Met	<b>C</b>	Cysteine	Cys
<b>F</b>	Phenylalanine	Phe	<b>Y</b>	Tyrosine	Tyr
<b>W</b>	Tryptophan	Trp	<b>H</b>	Histidine	His
<b>K</b>	Lysine	Lys	<b>R</b>	Arginine	Arg
<b>Q</b>	Glutamine	Gln	<b>N</b>	Asparagine	Asn
<b>E</b>	Glutamic Acid	Glu	<b>D</b>	Aspartic Acid	Asp
<b>S</b>	Serine	Ser	<b>T</b>	Threonine	Thr

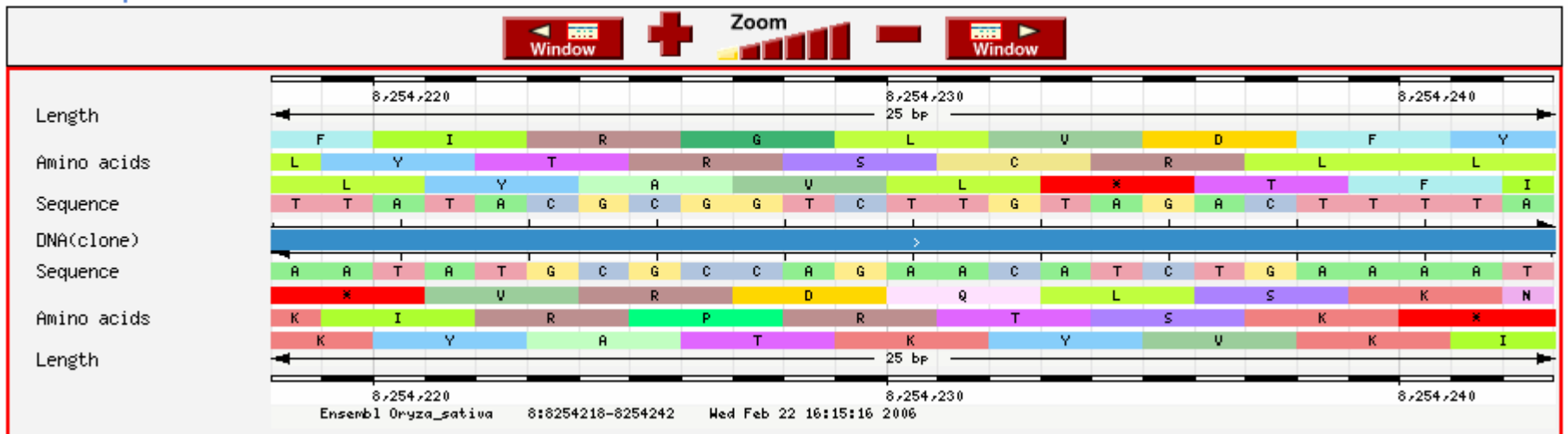
**LOC\_Os11g47740**

MSVQTLISFFLLPLSLPSLSPVASLSTLARAMAAVADDVGASVAVVEPEPEERPKVGGAAAAHG  
EAAEEEVVFPMAWTDDELCPEGTPVVRQTTKRDLRSSSSPCLGMKQPRAGVPLVSSA

Peptide Sequence

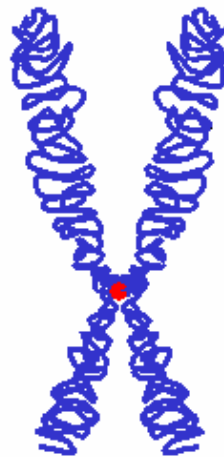


[-] Basepair view

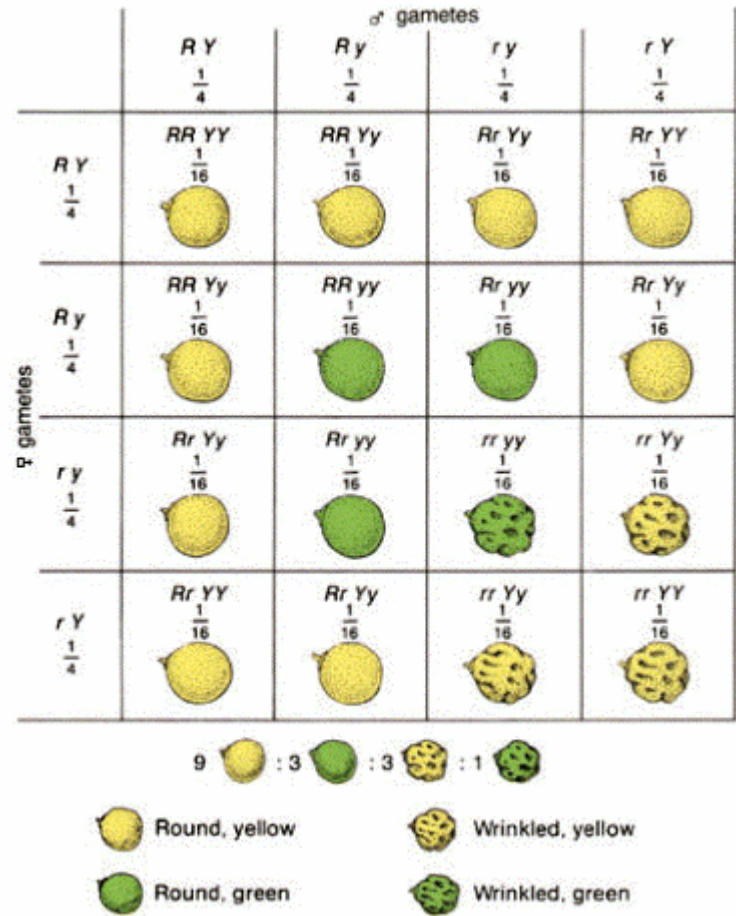


# Trait Inheritance - Alleles

- Genes come in pairs, and have alternative forms called alleles, some which are dominant.



<http://en.wikipedia.org/wiki/Chromosome>



<http://www.emc.maricopa.edu/faculty/farabee/biobk/BioBookgenintro.html>

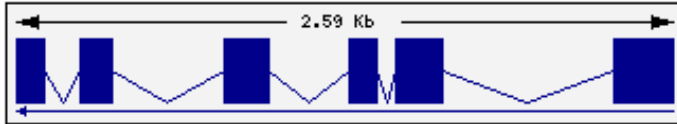


# Proteins

- Any change (mutation) in the sequence will alter the proteins produced, thus possibly resulting in different phenotypes beyond the expected Punnett square.

# Coding and non-coding sequences

Transcript Structure

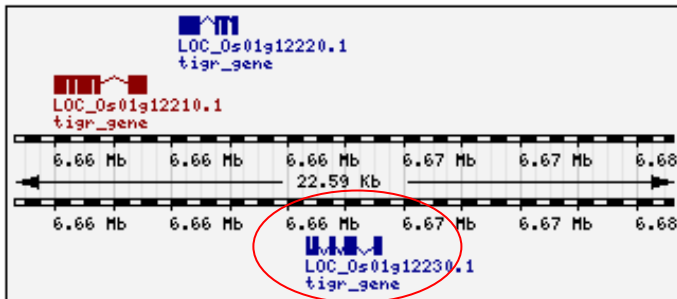


```

ATGGGGCGGCGCGTGGACCATGGCGGGCGGGGAGCTGGACGTGTTCCGATGCCGGGCGG
TCCGGCGGCGGTGGGTACGAGCTGGGGGTCCCGTGGGGCGGAGGTTCCGGCGAGGCCATC
AGGACAGGATGAGCGGCGACGCGCTCTCTGGCGGGCGGCTCTCCCGTTCGGTCCGAC
CGCGCGGGCGGCGCGTGGTGGACCGGCTCCGGCAGCCCAACCGCCGAGGAAGGAGCTC
CTGCCGTTTCATCACAGAGGAGGAAGATCAACACCACCACCGGGGAAGACGAAGCCGCGCC
GTCGCGCGCGGACGCGCGACGACGACTGCTCCGACGTCCTGATCGTCGGCGAGTCCGGCGCC
ATCGCCGCGCACACAGGAGACGCCAACGTCGGCGCTCTCCGGCCACACCTACCTGGTGAAG
CCACCTCCCGGGGACGGCTCTGCTCTTCCGCGCTACACCTACGCGGGGAGCTCCCG
ACCTCCGCTTCCGGCTTCAACAGCAACCGGAGTCAAGGTTGAGCTCGCCGGCCATGTCGGTC
GGCCACAGCTACAACCTGATGGAGCTCCGGCGCCGGAGGATCGTCAACGTCGAGACCGCC
TCCGGCAACCGCTTCTCCGTCGCGGAGGCCCGCCCGCGCCCTTCTTCCACGCCAACATG
TACCGCCATCTCCAGGTCAACCAGCTGCAAGGACGAGAATCCATGACGAGGACAGGAGA
CGCGCGGAGCTCTCCCGGACACCAAGCAGAAAGCGCTCTCCGCTCCGGGACACCGCCG
GATGACAAGTACCGGATCTACATGACAGGTCCAACACTGTACACTCTGTGTCACCGCTTGTG
GTTGATCTCGACGAGGCGACGATGACCATCTACAAGGGGAATCCGAAGAACAGAGATGCA
GTTCCAGTGTCCGATGCTGTGA
    
```

[Exons are shown in alternating blue/black] - [View exon information]

Transcript Neighbourhood



No.Exon / Intron	Chr	Strand	Start	End	Length	Sequence
5' upstream sequence						.....cttgatccttctgcccattccaggctcatatagcacacacatagc
1 11667_e05582	1	-1	6666995	6667224	230 bp	ATGGGGCGGCGCGTGGACCATGGCGGGCGGGGAGCTGGACGTGTTCCGATGCCGGGCGG TCCGGCGGCGGTGGGTACGAGCTGGGGGTCCCGTGGGGCGGAGGTTCCGGCGAGGCCATC AGGACAGGATGAGCGGCGACGCGCTCTCTGGCGGGCGGCTCTCCCGTTCGGTCCGCTCGACG CGCGCGGGCGGCGCGTGGTGGACCGGCTCCGGCAGCCCAACCGCCGAGGAAGGAGCTC CTGCCGTTTCATCACAGAGGAGGAAGATCAACACCACCACCGGGGAAGACGAAGCCGCGCC GTCGCGCGCGGACGCGCGACGACGACTGCTCCGACGTCCTGATCGTCGGCGAGTCCGGCGCC ATCGCCGCGCACACAGGAGACGCCAACGTCGGCGCTCTCCGGCCACACCTACCTGGTGAAG CCACCTCCCGGGGACGGCTCTGCTCTTCCGCGCTACACCTACGCGGGGAGCTCCCG ACCTCCGCTTCCGGCTTCAACAGCAACCGGAGTCAAGGTTGAGCTCGCCGGCCATGTCGGTC GGCCACAGCTACAACCTGATGGAGCTCCGGCGCCGGAGGATCGTCAACGTCGAGACCGCC TCCGGCAACCGCTTCTCCGTCGCGGAGGCCCGCCCGCGCCCTTCTTCCACGCCAACATG TACCGCCATCTCCAGGTCAACCAGCTGCAAGGACGAGAATCCATGACGAGGACAGGAGA CGCGCGGAGCTCTCCCGGACACCAAGCAGAAAGCGCTCTCCGCTCCGGGACACCGCCG GATGACAAGTACCGGATCTACATGACAGGTCCAACACTGTACACTCTGTGTCACCGCTTGTG GTTGATCTCGACGAGGCGACGATGACCATCTACAAGGGGAATCCGAAGAACAGAGATGCA GTTCCAGTGTCCGATGCTGTGA
Intron 1-2	1	-1	6666315	6666994	680 bp	gtaccgcgctactgggacagatg.....aacagtgattctagtaactttag
2 11667_e05581	1	-1	6666138	6666314	177 bp	GAAAGGAGCTCTGCGGTTTCATCACAGAGGAGGAAGATCAACACCACCACCGGGGAAGACGA AGCCGCCCGGCTGCCCGCCGACCGCGACGACTGCTCCGACGCTCTGATCGTCGGCGA GTCGGCGCGGATCGCCGCGCACACAGGAGGACGCCAACGTCGCGCTCTCCGGCCACAC gtacgcactcccactctctctct.....cgtgaatccctgggttttccag
Intron 2-3	1	-1	6666059	6666137	79 bp	gtacgcactcccactctctctct.....cgtgaatccctgggttttccag
3 11667_e05580	1	-1	6665953	6666058	106 bp	CTACGTGGTGAAGGCCACCTCGCCGACGGCTGCTCGTCTCACCGCTACACCTACCC CGCGGAGCTCCCACCTGCGCTTCCGGCTTCAACAGCAACGGAGTG gtactgtagcgcgatcgagccggca.....ttggtgatctctgctgagcagcag
Intron 3-4	1	-1	6665632	6665952	321 bp	gtactgtagcgcgatcgagccggca.....ttggtgatctctgctgagcagcag
4 11667_e05579	1	-1	6665461	6665631	171 bp	AGGGTGAGCTCGCCGGCCATGTCGGTCCGGCCACAGCTACAACCTGAGGACTCGCGGCGC CGGAGGATCGTCAACGTCGAGACCGCTCCGGCAACCGCTTCTCCGTCGCGGAGGCCCGCC GCCGCGCCCTTCTTCCAGCCCAACATGTACCGCCATCTCCAGGTCACACAG gtaattaatcaagcagataggaaag.....taactaaacctgtgaatttgcag
Intron 4-5	1	-1	6665013	6665460	448 bp	gtaattaatcaagcagataggaaag.....taactaaacctgtgaatttgcag
5 11667_e05578	1	-1	6664889	6665012	124 bp	GTGCAGGACGAGAATCCATGAGCAGGAGGAGAGAGCGCGGAGCTCTCCGGGACACC AAGGAGAAGGCGCTGCTGCTGCTCGCGGACACCGCGGATGACAAGTACCCGATCTACATG ACAG gtgaggtgattgtttcatctttca.....ctgatgaactgagcattgattcag
Intron 5-6	1	-1	6664752	6664888	137 bp	gtgaggtgattgtttcatctttca.....ctgatgaactgagcattgattcag
6 11667_e05577	1	-1	6664636	6664751	116 bp	GTCCAACACTGTACACTCTGTGCACCGCTTGTGTTGATCTCGACGAGGCGACGATGACCA TCTACAAGGGGAATCCGAAGAACAGAGATGACGATTCGAGTGTCCGATGCTGTGA aactagtagaacagataagcttattagctctagaagtaagagattagca.....
3' downstream sequence						aactagtagaacagataagcttattagctctagaagtaagagattagca.....



How important is it to know the identity of the gene(s) underlying a trait of interest?

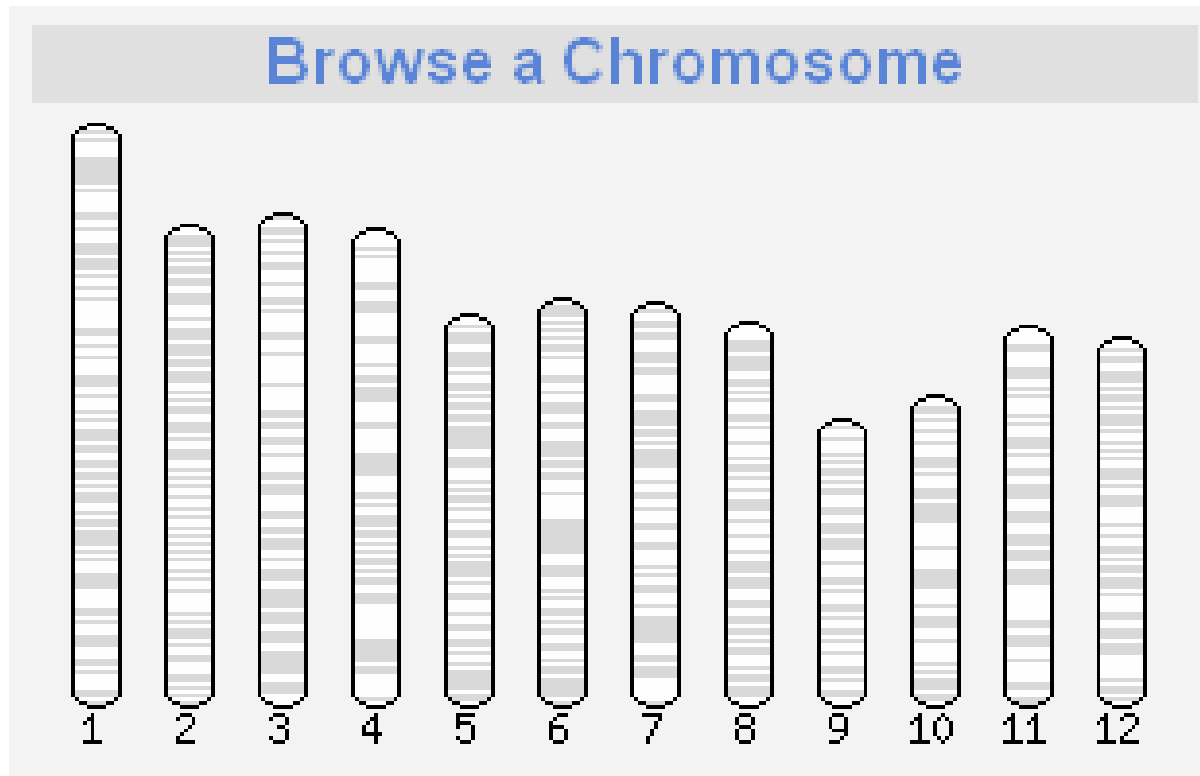


# How does Gramene Depict the Genotype?

The rice genome has 12 chromosomes, with 259 -485 mbase pairs

It has the smallest genome of the grasses.

For sorghum it is ~1190 mbp, and for maize it is ~2670 mbp.



- **A biological database is:**
  - [extensive] collection of biological data in computerized format . . .
  - for updating, querying and retrieving information. (communication)
- HCI (Human-Computer interaction)



- Using information in well-characterized organisms to better understand similar sequences in other organisms.
- To be useful to researchers:
  - Must have easy access to the information
  - Must have a method for extracting only that pertinent information.
- Database developers must understand both human and computer functioning.

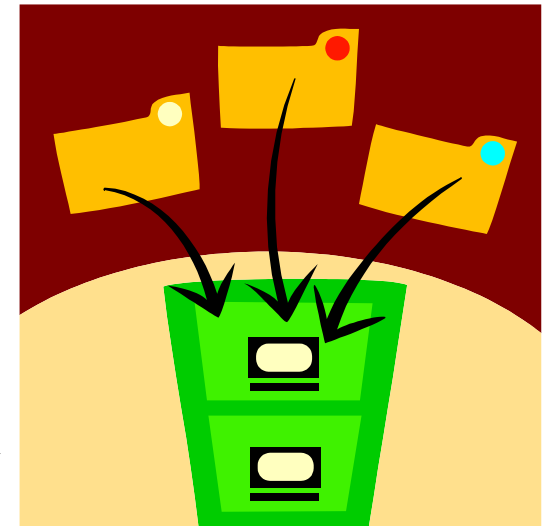


# Human-Human

## Tendency for Organization

This common vocabulary is arranged in a structured order or network based on their relationships to each other.

Kingdom: Animalia  
Phylum: Chordata  
Subphylum: Vertebrata  
Class: Mammalia  
Subclass: Theria  
Infraclass: Eutheria  
Order: Primates  
Suborder: Anthroipoidea  
Superfamily: Hominoidea  
Family: Hominidae  
Genus: Homo  
Species: sapiens

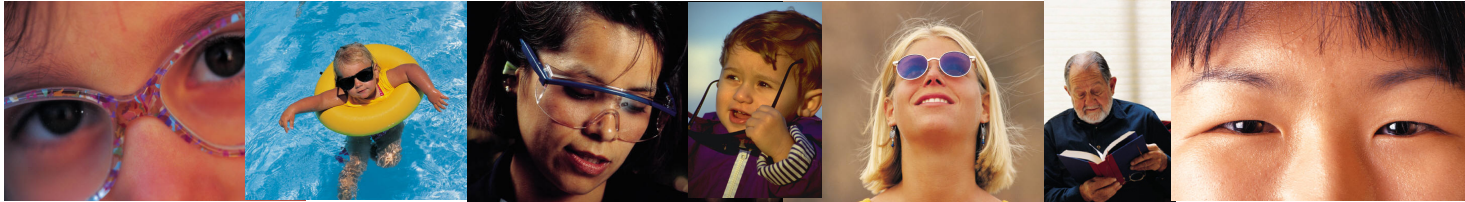


Facilitates global communication



# Perspectives

- When there is no pre-existing organizational template, our training and experience shape our language.



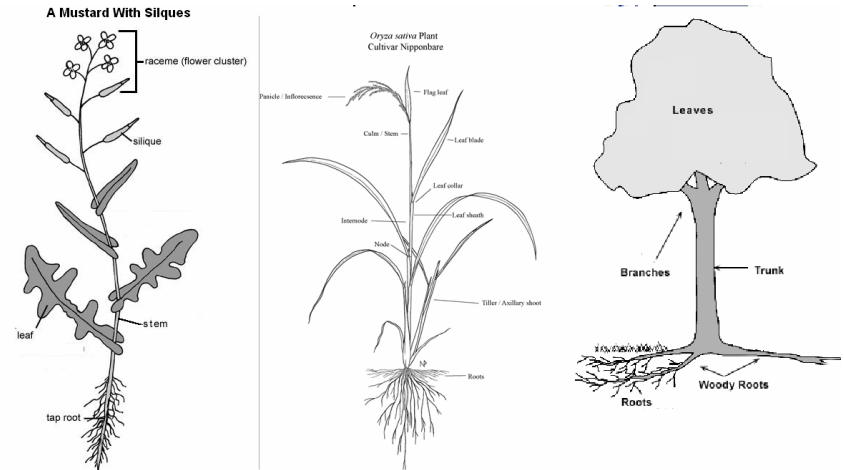
Term Name	Synonym
photoperiodism, <b>flowering</b>	photoperiodic control of <b>flowering</b> time, photoperiodic control of inflorescence development, response to daylength, <b>flowering</b> , response to nightlength, <b>flowering</b> , response to photoperiod, <b>flowering</b> .
long-day photoperiodism, <b>flowering</b>	long-day photoperiodic control of <b>flowering</b> , long-day photoperiodic control of <b>flowering</b> time, long-day photoperiodic control of inflorescence development, response to long-day, <b>flowering</b> , response to short-night, <b>flowering</b> , short-night photoperiodism, <b>flowering</b> .
short-day photoperiodism, <b>flowering</b>	long-night photoperiodism, <b>flowering</b> , response to long-night, <b>flowering</b> , response to short-day, <b>flowering</b> , short-day photoperiodic control of <b>flowering</b> , short-day photoperiodic control of <b>flowering</b> time, short-day photoperiodic control of inflorescence development.

Many authors use variations of terminology when discussing the same or similar concepts (synonyms).



# Human – Computer interaction

- Computers can manage and interpret data.
  - Classify data
  - Store reasoning tools.
  - Infer new knowledge



Term Name	Synonym	Definition
fruit	achene, berry, capsule, caryopsis, circumcissile capsule, cypsela, drupe, follicle, <b>grain</b> , kernel, legume, loculicidal capsule, lomentum, nut, PO:0020067, PO:0020068, PO:0020069, PO:0020070, PO:0020071, PO:0020072, PO:0020073, PO:0020074, PO:0020076, PO:0020077, PO:0020078, PO:0020079, PO:0020080, PO:0020082, PO:0020083, PO:0020087, PO:0020107, pod, poricidal capsule, schizocarp, septicidal capsule, septifragal capsule, silicula, siliqua, silique.	The seed-bearing structure in angiosperms, formed from the ovary after flowering.



# Ontologies

- An Ontology is a “common vocabulary” used to facilitate communication between researchers and databases.
- An ontology matches classification and reasoning methods of the computer.
- Ontologies can be indexed “objectively” by a computer.



# Ontologies at Gramene

- Plant Ontology
  - Plant Structure(PO)
  - Growth stages (GRO)
- Gene Ontology (GO)
  - Molecular Function
  - Biological Process
  - Cellular Component
- Environment Ontology (EO)
- Gramene Taxonomy (Gr\_tax)
- Trait Ontology (TO)



# Trait ontology (TO)

- Anatomy and Morphology related traits (root #, depth, penetration)
  - Stature or Vigor Trait (height, germination)
  - Stress trait (environmental factors, drought, disease)
  - Biochemical trait (enzyme activity, fat/oil content)
  - Growth and development trait (days to maturity, ratooning ability)
  - Sterility or Fertility trait (pollen fertility, male sterility)
  - Quality Trait (flavor, tenderness, seed quality)
  - Yield Trait (grain yield, panicle weight)
  - Misc. Trait
- (In TO there is some overlap (ie root penetration is both anatomy and stress, Ratooning is both growth and yield))



- Through the ontology database terms at Gramene you can link traits to associations with various objects such as QTL, phenotype, gene, proteins and Ensembl rice genes.





# You Can Participate in Creating Ontologies

- You may submit a Gene Ontology (GO) term or Plant Ontology (PO) term by going to “resources” and selecting “Submit a Term.”

The screenshot displays the GRAMENE Ontologies website interface. At the top, the logo 'GRAMENE Ontologies' is visible, followed by a navigation bar with links for Search, Genomes, Download, Resources, About, and Help. A vertical menu on the right side of the page lists 'Species Pages', 'Glossary', 'Submit a Term' (highlighted in red), and 'Links Pages'. The main content area is divided into two primary sections: 'Controlled Vocabulary Suggestion' and 'Basic Rice Gene Information'.

**Controlled Vocabulary Suggestion**

- \* Submission Type: [Dropdown menu]
- If correction, please specify original term: [Text input]
- \* Suggested Term: [Text input]
- Synonym: *If more than one, please separate with semi-colons* [Text input]
- \* Associated Ontology: [Dropdown menu]
- Taxonomic Class: [Dropdown menu] Please specify if "other" [Text input]
- \* Definition of the Suggested Term: [Text input]

**Basic Rice Gene Information**

- \* Gene Name: [Text input]
- \* Phenotypic Description: [Text area]
- Other Description: [Text area]
- Reference for the Rice Gene:
  - Ref. Category: [Dropdown menu]
  - Curator: [Dropdown menu] (internal use only)
  - Ref. Source: [Dropdown menu]
  - \* Reference Detail: *please provide the complete citation for the reference(s)* [Text area]

The GRAMENE logo is also present in the bottom right corner of the page.

# Gramene database - 9 modules

**Genomes:** Browse and search genes, markers, expressed sequences, etc. on the Rice-Japonica, Maize and Arabidopsis genomes, as well as sequences from, sorghum, barley and wheat and other cereals mapped on rice.

**BLAST:** Search sequences for similarity matches; select the best target sequence database and alignment parameters for your search.

**Maps:** Search and view mapped genes, markers, QTL and clones using various types of maps (including genetic, physical and sequence); Compare maps; Examine genetic co-linearity between species.

**Markers:** Find a marker based upon name, type or species; View its detail.

**Proteins:** Find a protein and its sequence; Determine it's cellular location and function; Explore protein families.

**Genes:** Learn about genes and alleles associated with important phenotypes and functions.

**QTL:** Find qualitative trait locus (QTL) from major cereal crops associated with traits.

**Ontologies:** Find keywords for plant structure, growth stages, traits, function, process, cellular component, environment and taxonomy.

**Literature:** Find articles about genes, proteins, QTL, markers, or ontologies.



## Quick Search

Find anything

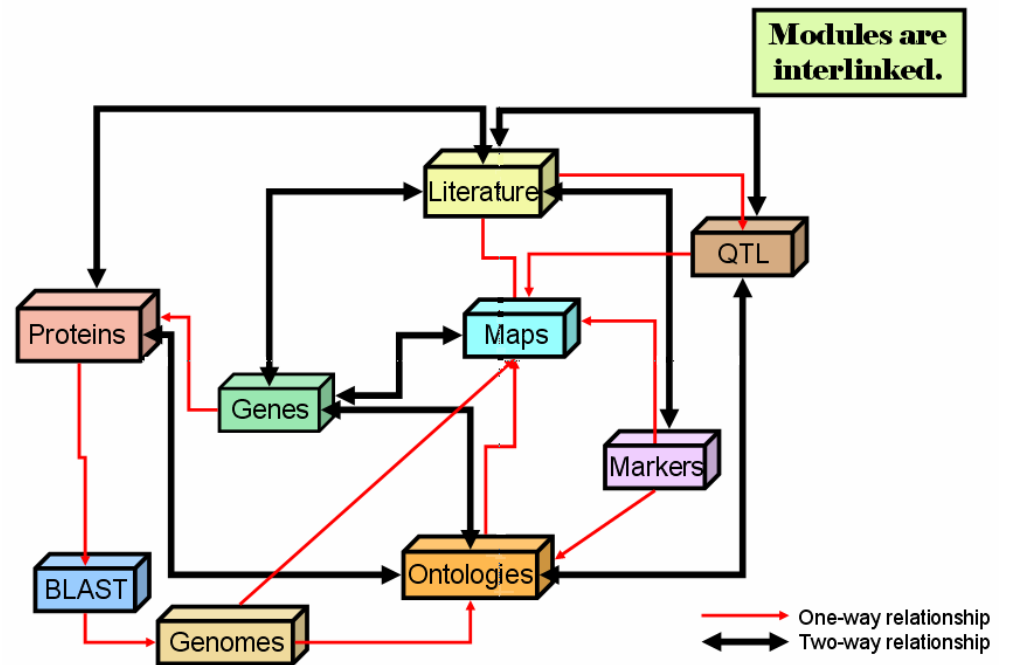
Search

- Genomes-Ensembl
- Maps-CMap
- Markers
- QTL
- Genes
- Proteins
- Ontologies
- Literature
- Sequences-BLAST
- All-GrameneMart

## Quick Start

Accessed genomes for [Rice](#), [Maize](#) & [Arabidopsis](#).  
 Refine your search with [GrameneMart](#); Start  
 your search by [Gene Ontology](#).

**Modules are interlinked.**



Map Type: **Genetic**

Accession ID: genetic

Map Units: cM

Map Type: **QTL**

Accession ID: qtl

Map Units: cM

Map Type: **Physical**

Map Units: bands

Is Relational Only

Map Type: **Deletion**

Accession ID: deletion

Map Units: arm fraction

Map Type: **Sequence**

Accession ID: sequence

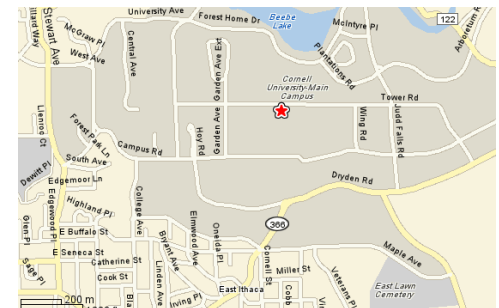
Map Units: bp

# Map Types

**genetic maps** - indirect estimate of the distance and order of the sequence; use markers; provide framework for genome sequencing.



**physical maps**- estimate of the true distance (base pairs); define sequence between markers; used for gene identification.



# Features

- Everything that can be mapped
- Gene – sequence of DNA that codes for a product that in turn has a predicted or demonstrated impact on phenotype.
  - (Some genes not yet associated with sequences)

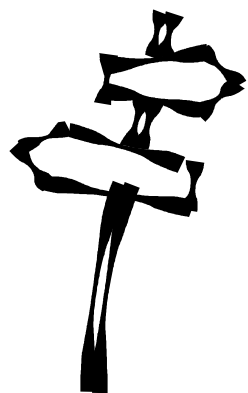


# Molecular Markers

- Molecular markers are the genetic signposts (dna segments) that flag the presence of genes that control particular traits. Once we know about a useful molecular marker, plant breeders can test the seeds or leaves of a plant to make sure it has the allele they are looking for.
- Molecular markers can be generated by a variety of different techniques, and information surrounding these markers will vary quite a bit depending on the marker type.
- When we know what alleles a plant has at loci associated with critical traits of agronomic importance and what those alleles do. We can make predictions about whether it can thrive in certain conditions.
- Measuring such characteristics by conventional methods is much more difficult, time-consuming, or expensive, since it requires the organism to grow to maturity.



# Gramene Markers



## Currently in Marker Database

Marker Type	Count
AFLP	950
BAC end sequence	676,959
Clone	602,304
EST	2,530,661
EST Cluster	601,935
Gene Primer	19
mRNA	91,074
Primer	34
RAPD	135
RFLP	7,791
SSR	16,835
STS	65
Tos17	18,023
Undefined	823
4,547,608 total	

# Markers Screenshot

Oryza sp. mRNA "S49967"

Marker ID	5153045
Marker Name	S49967
Synonyms (2)	259136 S49967.1
Type	mRNA
Species	<a href="#">Oryza sp. (Oryza)</a>
Germplasm	N/A
Description	oryzacystatin=cysteine protease inhibitor [Oryza=rice, mRNA, 643 nt].
Mappings (0)	No map positions.
Correspondences	No correspondences.
Sequence Length	643
Sequence	GCATTCGCTAGCCACGCCGTCCTCCGCTCAGGCCGAGGGCCATCGCGCAGGGG GAGAAGGGGAGGAGAAGATGTTCGAGCGACGGAGGGCCGGTGTGGCGGC GTCGAGCCCGTGGGAACGAGAACGACCTCCACCTCGTCGACCTCGCCCG CTTCGCCGTCACCGAGCACAAACAAGAAGGCCAATTCTCTGCTGGAGTTGG AGAAGCTTGTGAGTGTGAAGCAGCAAGTTGTCGCTGGCACTTTGACTAT TTCACAATTGAGGTGAAGGAAGGGGATGCCAAGAAGCTCTATGAAAGCTAA GGTCTGGGAGAAACCATGGATGGACTTCAAGGAGCTCCAGGAGTTCAAGC CTGTCGATGCCAGTGCAAATGCCTAAGGCCATCTCGTATCCTATGTGTA TCAAGTTATCAAGAAGATGGGGAATAATATGGTGGATATAGCTATTGG ACATGTTAATTATCCACATGATAATATGGCTTGGATATAAGGATCTCACA CGATAATATGGCTTGGATATATAGCTATTAAAGATTTTACCTATGGCATA TTCAATGTGATTAGTACTAAGTAAGAATGATTGCAAGGTGATTAACT ACAAATATTGCAATAAAAGTCCCTGTTACTACAACCTTACAAGG
Date Created	1993-05-08
Date Updated	1993-05-08

Mol T	
Not	
Numl	
Orig	
Phenotype	
Plasmid	
Pop Variant	
Product	oryzacystatin
Protein Id	AAB24010.1
Pseudo	
Rearranged	
Ref Authors	Chen,M.S., Johnson,B., Wen,L., Muthukrishnan,S., Kramer,K.J., Morgan,T.D. and Reeck,G.R.
Ref Location	Protein Expr. Purif. 3 (1), 41-49 (1992)
Ref Pubmed	1422207
Ref Title	Rice cystatin: bacterial expression, purification, cysteine proteinase inhibitory activity, and insect growth suppressing activity of a truncated form of the protein
Ref Year	1992

Translation	MSSDGGPVLGGV KQQVVAGTLYYFTIEVKEGDAKKLYEAKVWEKPMDFKELQEFKPVDA NA
Transposon	
Variety	
Genome Positions	No known genome positions.
Images	No images.



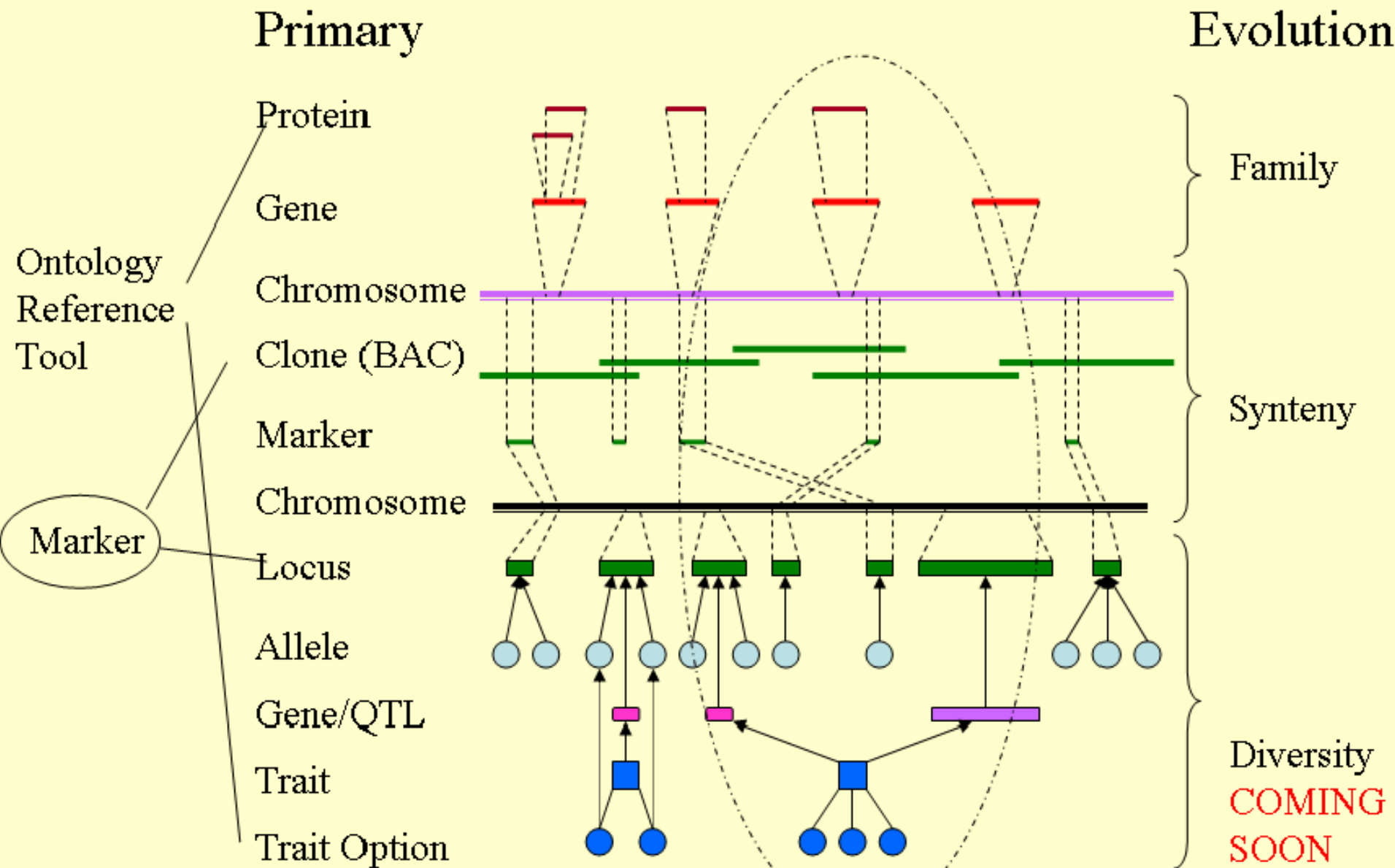


# QTL

- A QTL (or quantitative trait locus) refers to a particular region of DNA that is hypothetically associated with a particular trait.
- Based on the idea that some traits are controlled by multiple genes.



# Overview of concepts in Gramene



# Mapped Features

## –All Feature Types–

AFLP

BAC

Breakpoint interval

Centromere

Clone

EST

GenBank Marker

MMP Unigene

Maize Bin

Maize EST

Maize Marker

Marker

Phenotype

QTL

RFLP

Rice Marker

Rice SSR

SOG

SSR

STS

Sorghum GSS

Sorghum Marker

Tos17 Insertion

Wheat EST Marker

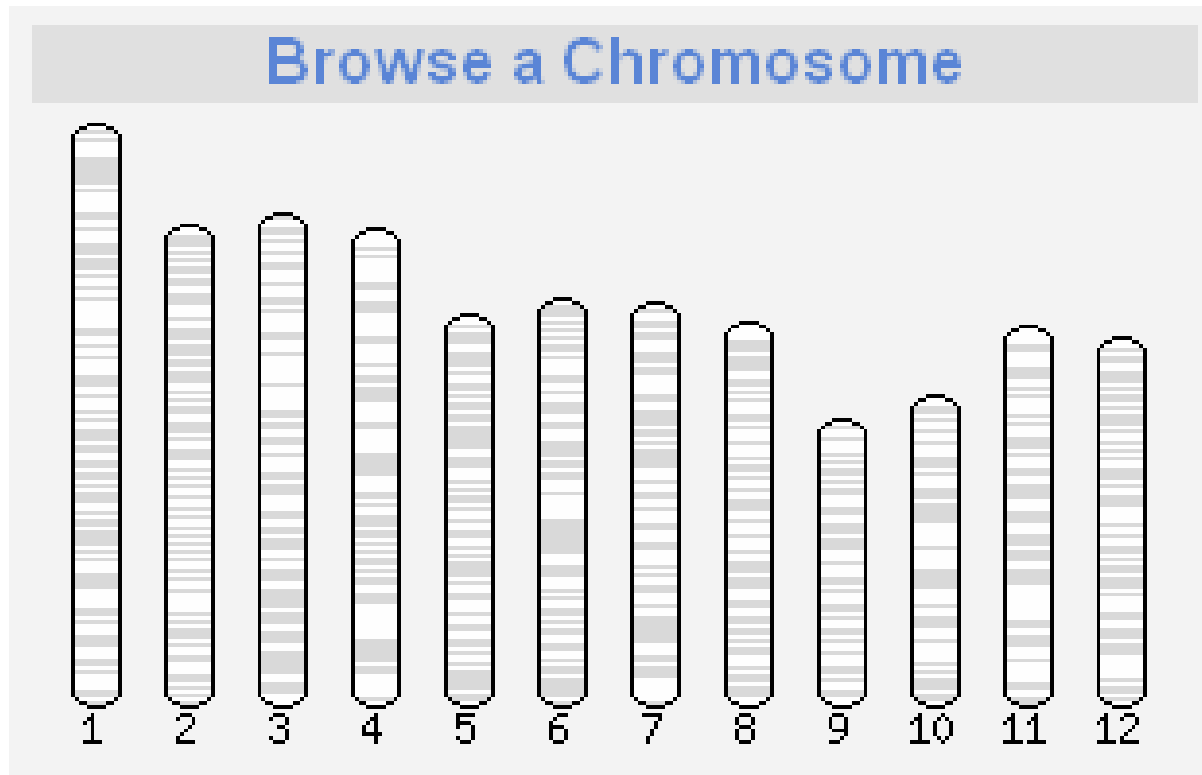
dd

interpolated phenotype

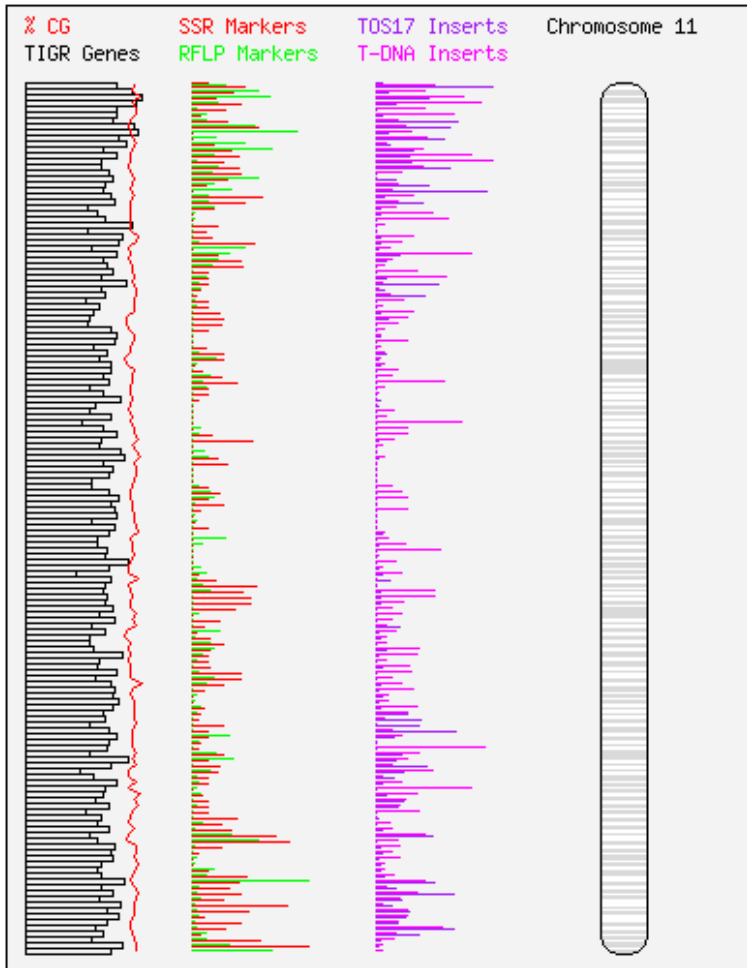
low LOD marker



# Rice Genome



# Map of Chromosome 11



Change Chromosome

Chromosome: 11

Chromosome 11

Length:	28369397 bps
GeneModel_Submission:	9
Rice_GeneModel_FGENESH:	4234
Rice_GeneModel_TIGR:	4367
Rice_ArrayConsensus_Affy57K:	9193
Rice_ArrayOligo_NSF20K:	676
Rice_BAC:	25
Rice_CDS:	323
Rice_EST:	29674
Rice_ESTcluster_PlantGDB:	11576
Rice_ESTcluster_TIGR:	5978
Rice_FST-TDNA:	375
Rice_FSTtos17:	1095
Rice_FSTtransposon:	187
Rice_MPSStag_Meyers:	4718
Rice_MarkerRFLP:	288
Rice_MarkerSSR:	176
Rice_SAGETag_MGOS:	4909
RiceAlta_BACend_OMAP:	12735
RiceAustraliensis_BACend_OMAP:	14383
RiceBrachyantha_BACend_OMAP:	4021
RiceCoarctata_BACend_OMAP:	9652
RiceGlaberrima_BACend_OMAP:	3387
RiceIndica_EST_BGI:	8009
RiceIndica_ESTcluster_BGI:	2374



# Screenshot of Browser

## Chromosome 11

### Chromosome 11

Chr. 11

### Overview

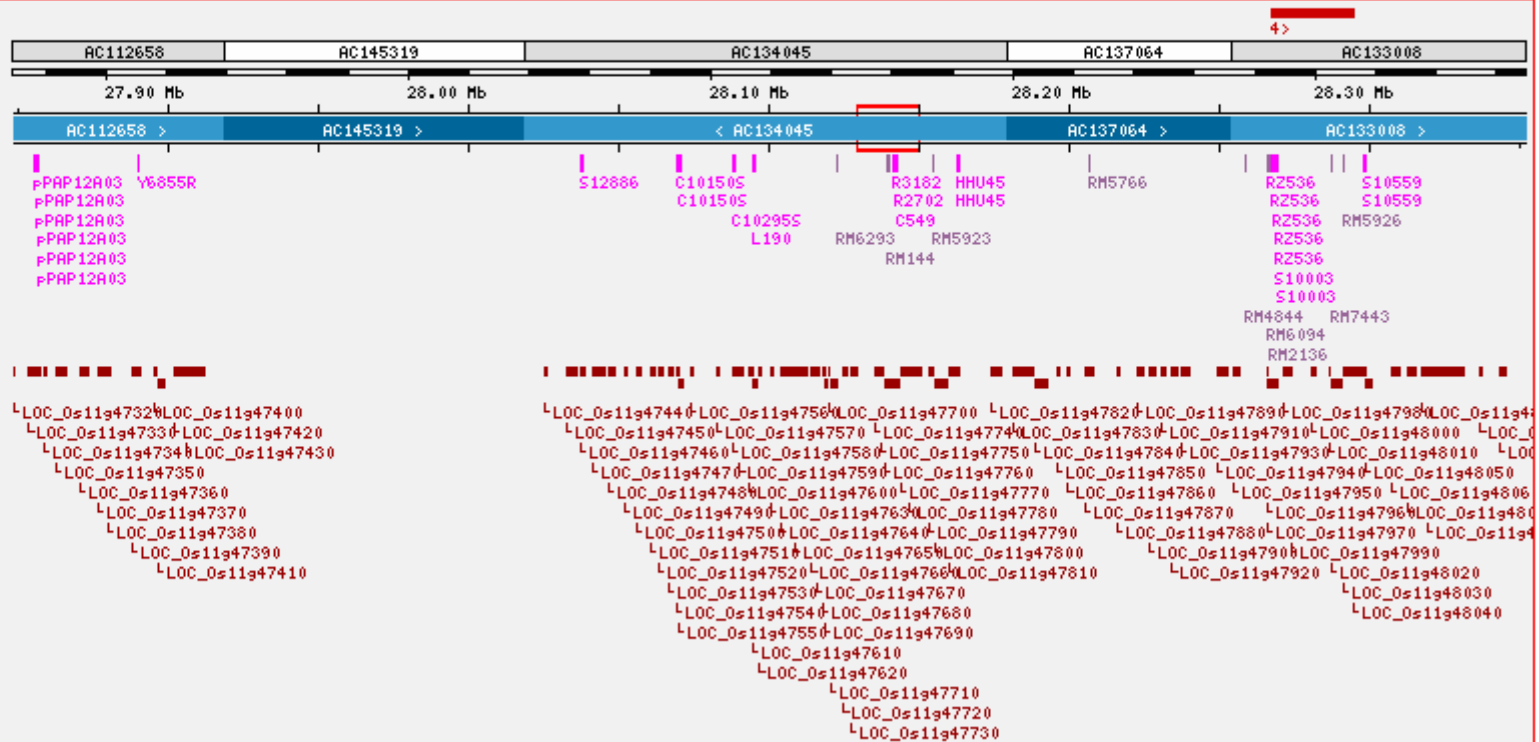
maize synteny  
Chr. 11 band

DNA(clone)

Rice\_Markers

GeneModel TIGR

Gene legend







# Detailed view

Jump to region:  bp  to

**Refresh**

**<< 2 Mb**

**< 1 Mb**

**Window**

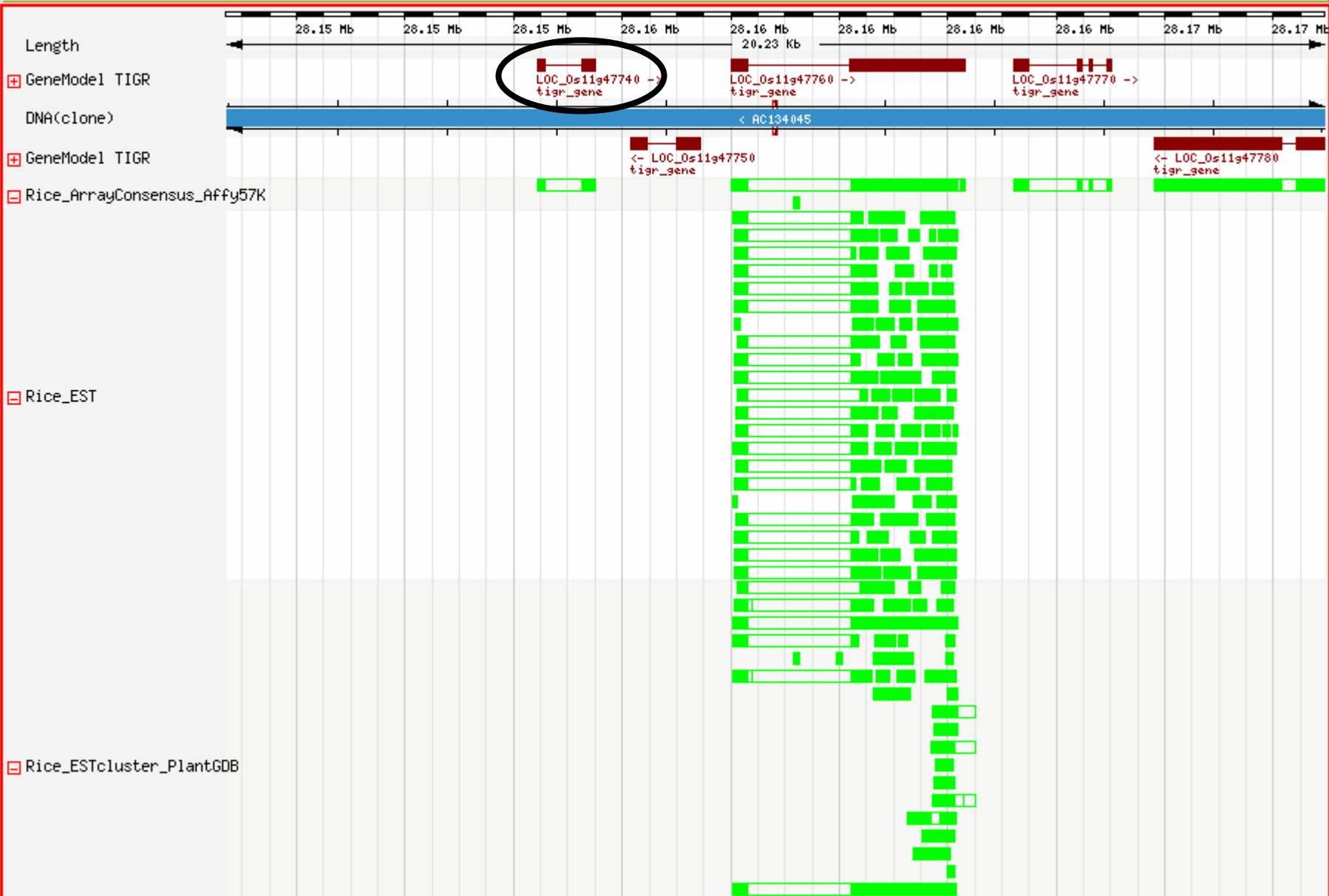


**Window**

**1 Mb >**

**2 Mb >>**

Features ▾ ESTs ▾ GSSs ▾ FSTs ▾ Markers ▾ Arrays ▾ Decorations ▾ Export ▾ Jump to ▾ Width ▾ Help ▾



# Customizing the Detail View

Jump to region:  bp  to  Refresh

<< 2 Mb < 1 Mb Window Zoom Window 1 Mb 2 Mb >>

Features ▾ ESTs ▾ GSSs ▾ FSTs ▾ Markers ▾ Arrays ▾ Decorations ▾ Export ▾ Jump to ▾ Width ▾

**Features**

- SNPs
- Rice\_GeneModel\_TIGR
- Rice\_GeneModel\_Submitted
- genbank\_trna
- Rice\_GeneModel\_FGENESH
- Rice\_EST
- Rice\_CDS
- Maize\_CDS
- Maize\_Methylation
- Sorghum\_Methylation

**ESTs**

- Rice\_EST
- Rice\_ESTCluster
- Rice\_ESTCluster\_TIGR
- RiceIndica\_EST
- RiceIndica\_ESTCluster
- RiceJaponica\_EST
- RiceJaponica\_ESTCluster
- Barley\_EST
- Barley\_ESTCluster
- Barley\_ESTCluster\_TIGR
- Maize\_EST
- Maize\_ESTCluster
- Maize\_ESTCluster\_TIGR
- Maize\_ESTCluster\_TUG
- Millet\_EST
- Sorghum\_EST
- Sorghum\_ESTCluster\_Pratt
- Sorghum\_ESTCluster\_TGI
- Sorghum\_ESTCluster\_TUG
- Sugarcane\_EST
- wheat\_EST
- wheat\_ESTCluster\_TGI
- wheat\_ESTCluster\_TUG

**GSSs**

- RiceAlta\_BACend\_OMAP
- RiceAustraliensis\_BACend\_OMAP
- RiceBrachyantha\_BACend\_OMAP
- RiceGlaberrima\_BACend\_OMAP
- RiceJaponica\_BACend\_IRGSP
- RiceNivara\_BACend\_OMAP
- RicePunctata\_BACend\_OMAP
- RiceRufipogon\_BACend\_OMAP
- Maize\_BACend
- Maize\_HiCot\_Bennetzen
- Maize\_HiCotCluster\_TIGR
- Maize\_HiCotMethylFilterCluster\_TIGR
- Maize\_MethylFilter\_CSHL
- Maize\_MethylFilter\_Orion
- Ryegrass\_MethylFilter\_Orion
- Ryegrass\_MethylFilterCluster\_Orion
- Sorghum\_GSS\_Klein
- Sorghum\_MethylFilter\_Orion

**FSTs**

- Rice\_FST\_Ds
- Rice\_FST\_IS
- Rice\_FST\_DNA
- Rice\_FST\_Tos17
- Maize\_FST\_Mu

**Markers**

- Rice\_Marker\_BFLP
- Rice\_Marker\_SSR
- Maize\_Markers
- Sorghum\_Markers
- wheat\_markers

**Arrays**

- Rice\_ArrayOligo\_NSFP
- Rice\_ArrayTarget\_Affinity
- Rice\_Mapping
- Barley\_ArrayTarget\_Affinity
- Maize\_ArrayTarget\_Affinity
- wheat\_ArrayTarget\_Affinity

**Decorations**

- Sequence
- Codons
- Start/Stop Codons
- Contigs
- Ruler
- Scale Bar
- %GC
- Show register lines
- Show empty tracks
- Show popup menus
- ... popup on click
- Half-height gaps
- Concise labels
- Restriction Enzymes
- Information tracks
- Gene legend
- SNP legend
- Reset options

**Export**


- Include PDF links
- Include SVG links
- Include POSTSCRIPT links
- FASTA
- Flat

**Jump To**

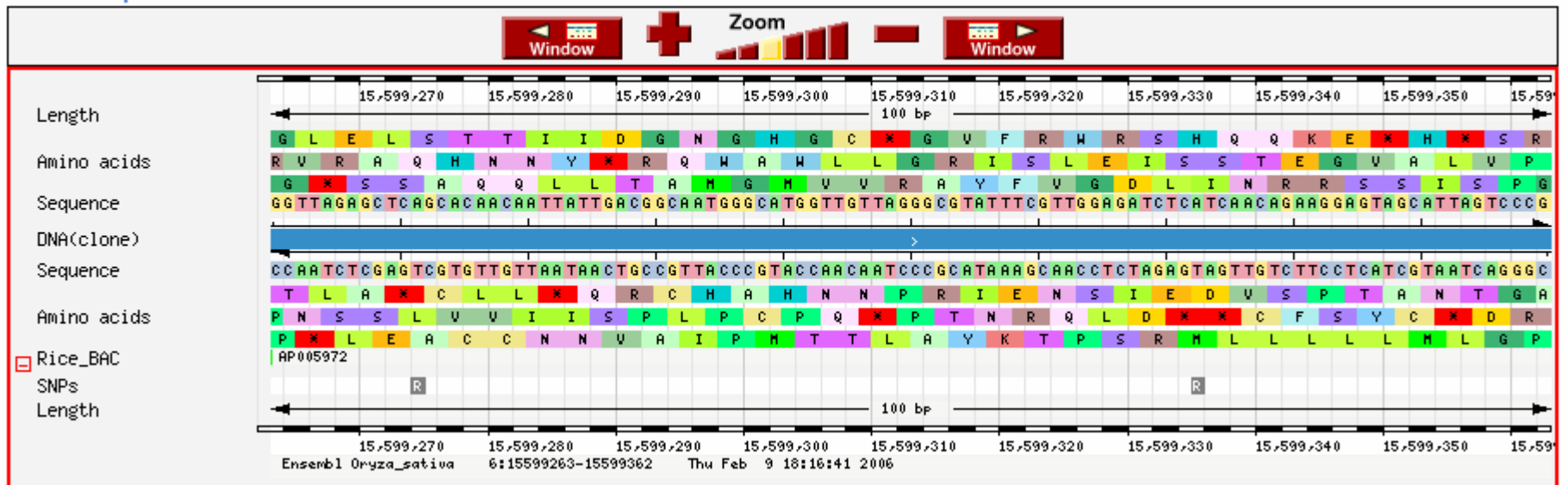
- SyntenyView (Zea\_mays)
- MapView
- CytoView
- CMAA TIGR Assembly

**Width**

- Width 700px
- \* Width 900px
- Width 1000px
- Width 1500px
- Width 2000px



## Basepair view





# Help

Don't get overwhelmed

.....Get help!



# Future Plans for Gramene

- QTL to sequence Map
- Diversity
- Pathways



# Gramene Contributors

## Controlled Vocabulary

- TAIR
- IRRI
- MaizeGDB
- Gene Ontology Consortium
- ICIS
- INGER
- CIMMYT
- Oryzabase
- Scottish Crop Research Institute
- Plant Ontology Consortium
- Monsanto
- Pioneer

## Maps

- Cari Soderland, University of Arizona
- Rod Wing, University of Arizona
- John Mullet, Texas A & M
- Patricia Klein, Texas A & M
- Robert Klein, USDA-ARS
- Andrew Paterson, Univ of GA
- Dave Matthews, USDA-ARS
- Mary Polacco, USDA-ARS
- Maize GDB

## Sequences and Proteins

- National Center for Gene Research Chinese Academy of Sciences
- NCBI
- TIGR
- Uniprot/SPTremble
- LGB
- BGI
- Interpro
- MaizeGDB
- BarleyBase
- Orion Genomics
- NASC for the arabidopsis genome view in Ensembl.
- Cari Soderland SAGE
- GuoLiang Wang SAGE
- Blake Meyers MPSS

## QTL

- MaizeGDB
- GrainGenes

## Genes

- Oryzabase
- MaizeGDB

## Reference Databases

- Albert Mann Library, Cornell University
- TEEAL

## Software Development

- Ensembl
- Gene Ontology Consortium
- GMOD
- Pioneer

## Germplasm

- NPGS/GRIN

## Collaborators

- Ben Faga – CMap
- Bonnie Hurwitz – OMap

## SAB

- Anna M McClung
- Georgia Davis
- James H. Oard
- David Marshall
- Patricia Klein

# Gramene Personnel

<b>PI Team</b>	<b>Lincoln Stein, Ph.D</b>	<b>PI, Database design and software development</b>
	<b>Susan R. McCouch , Ph.D.</b>	<b>Co-PI, rice data curation</b>
	<b>Doreen Ware, Ph.D.</b>	<b>Co-PI, Comparative genome analysis</b>
	<b>Pankaj Jaiswal, Ph.D</b>	<b>Co-PI; Proteins, Ontologies, QTL</b>
	<b>Ed Buckler, Ph.D</b>	<b>Co-PI, Diversity, Maize</b>
<b>Biological Curation</b>	<b>Junjian Ni, Ph.D.</b>	<b>Genes, QTL</b>
	<b>Immanuel Yap, Ph.D.</b>	<b>Maps, Markers, QTL</b>
	<b>Isaak Yosief Tecle, Ph.D.</b>	<b>Germplasm, Genetic Diversity</b>
	<b>Dean Ravenscroft, Ph.D.</b>	<b>Pathways</b>
<b>Bioinformatics Developers</b>	<b>Steven Schmidt</b>	<b>Gramene Genome Database &amp; Browser</b>
	<b>Ken Youens-Clark</b>	<b>Comparative Maps, QTL, Markers, Diversity</b>
	<b>Shulamit Avraham (Shuly)</b>	<b>Database Management</b>
	<b>Liya Ren</b>	<b>Proteins, Literature, Ontologies, Genes</b>
	<b>William Spooner, Ph.D.</b>	<b>Genome Browser, Gramene Blast &amp; Mart</b>
	<b>Payan Canaran</b>	<b>Protein Annotation Pipeline</b>
	<b>Sharon Wei</b>	<b>Gramene Pipeline and Genome Browser</b>
	<b>Terry Casstevens</b>	<b>Diversity</b>
<b>Coordinators</b>	<b>Chengzhi Liang, Ph.D</b>	<b>Project Management, Gene Build</b>
	<b>Claire Hebbard</b>	<b>Outreach Coordinator</b>



