

PlantPathMarks

Manual

PlantPathMarks Main-Bar Overview

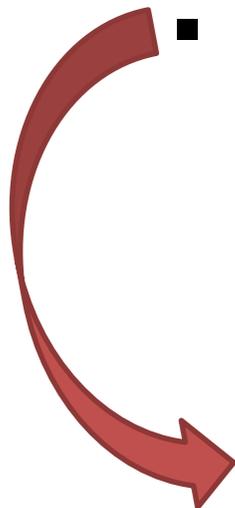


- **Home:** Navigation Bar; Database Search; Map Browse; JBrowse; Statistics; Data Resources; Bulk Download; Species Comparisons; Manual .

Home Page

The “**Home**” page introduces **PPMdb** as a comprehensive database along with its PGM pipeline by describing the PPMdb database sections, in addition to providing users with a ‘PPMdb Quick-Access; Analyzed Genomes in PPMdb Database Section’ option for all analyzed plant genomes. In the ‘PPMdb Quick-Access’ section, plant genomes analyzed in the PPMdb database are categorized according to plant type/class (dicot plants or monocot plants).

Then under each group, interactively, the plants were sorted alphabetically to enable straightforward access and searches within each plant genome independently through ‘**Map Browse**’ or ‘**JBrowse**’ quick-access links

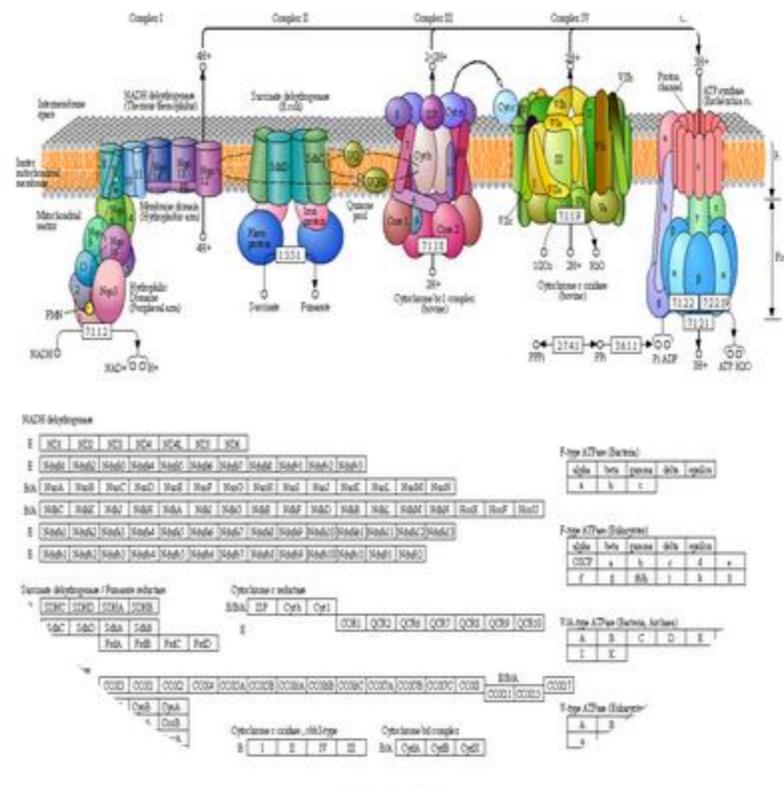


Navigation Bar

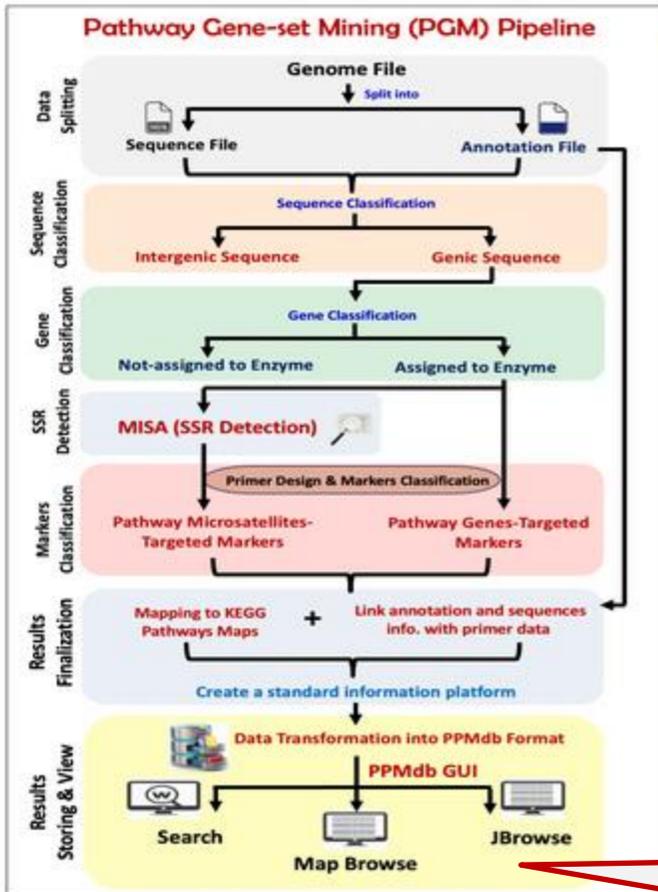
Home Page



PlantPathMarks (PPMdb) A hub for Plant Pathways-Based Markers



Valuable Resources



- Valuable Resources
- NCBI
 - SSRome
 - VFODB
 - MISAweb
 - Primer3web
 - EMBL-EBI
 - Phytozome
 - TAIR
- Useful Links
- CPAN
 - MySQL
 - KEGG

The workflow of Pathway Gene-Set Mining (PGM) pipeline

About PPMdb

The PlantPathMarks (PPMdb) portal is an extensive website providing interactive and convenient access and mapping to thousands of pathways-based primers/markers in 82 plant genomes. Here, we also developed a novel approach called the Pathway Gene-set Mining (PGM). Besides, new marker systems were developed based on their pathways-based mapping of developed markers. The systems were named Pathway Microsatellites-Targeted Markers (PMTM) and Pathway Genes-Targeted Markers (PGTM); as a kind of annotation-based marker. In the PlantPathMarks hub, 9894 marker-panels (comprising 2,690,742 markers) were developed across eighty-two plant genomes. This huge number of markers includes 1,999,187 PMTM and 691,555 PGTM markers

Dicot Plants

Monocot Plants

Analyzed Genomes in PPMdb Database

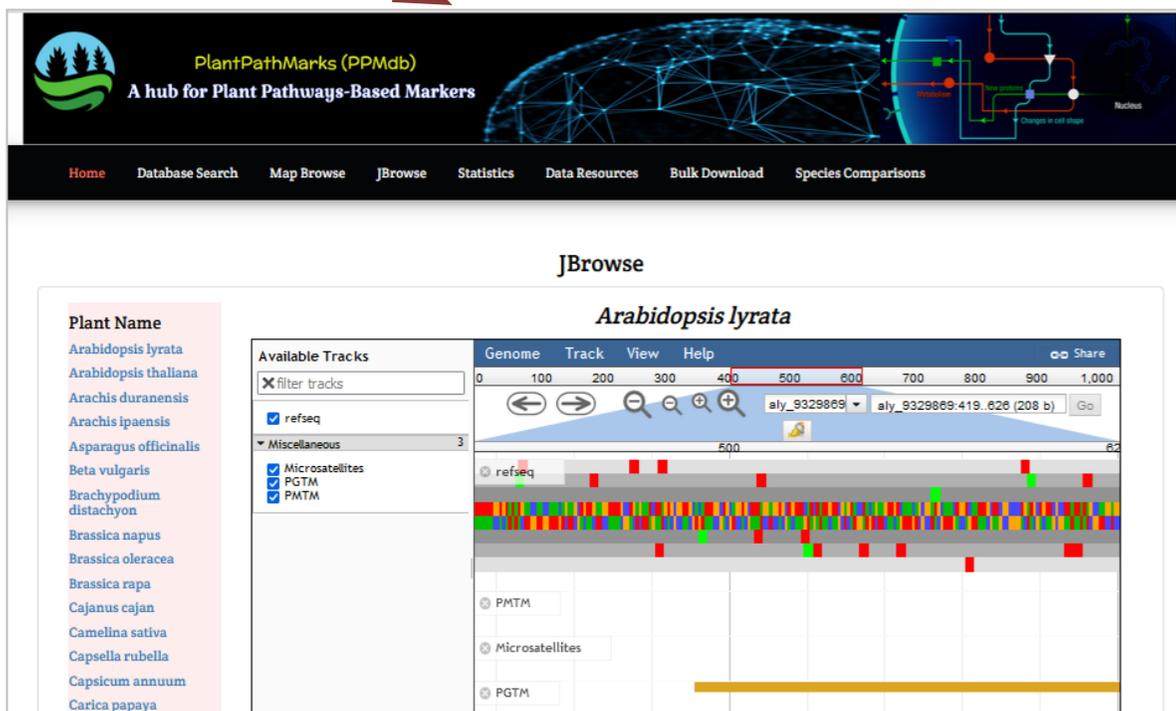
Plant Scientific Name	Map Browse Link	JBrowse Link	Plant Image
Dicot Plants			
Monocot Plants			
<i>Arabidopsis lyrata</i>	Map Browse	JBrowse	
<i>Arabidopsis thaliana</i>	Map Browse	JBrowse	
<i>Arachis duranensis</i>	Map Browse	JBrowse	

Plant Scientific Name

Map Browse Link

JBrowse Link

Plant Image



JBrowse

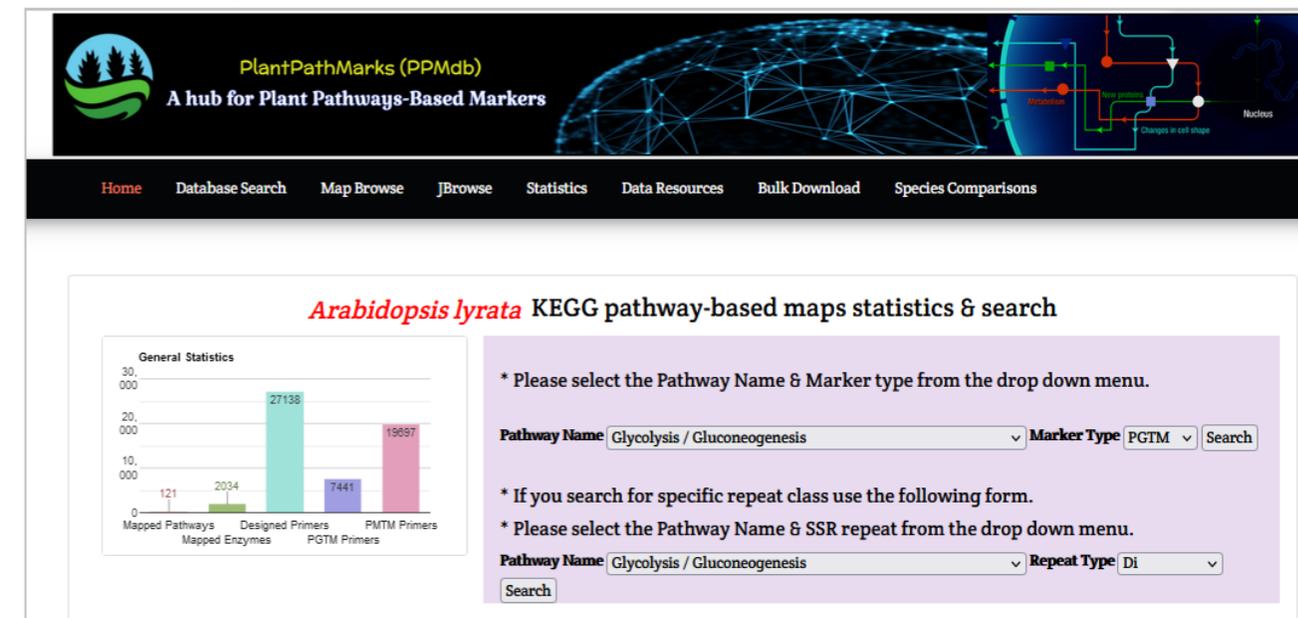
Arabidopsis lyrata

Available Tracks: refseq, Miscellaneous (Microsatellites, PGTM, PMTM)

Genome Track: 0 100 200 300 400 500 600 700 800 900 1,000

Microsatellites

PGTM



Arabidopsis lyrata KEGG pathway-based maps statistics & search

General Statistics

Category	Count
Mapped Pathways	121
Mapped Enzymes	2034
Designed Primers	27138
PMTM Primers	7441
PMTM Primers	19657

* Please select the Pathway Name & Marker type from the drop down menu.

Pathway Name: Marker Type: Search

* If you search for specific repeat class use the following form.

* Please select the Pathway Name & SSR repeat from the drop down menu.

Pathway Name: Repeat Type: Search

Database Search

The “**Database Search**” page provides users the searching utility within two main sections:

- a) The ‘**Search Section**’, which enables the user to obtain results by selecting the following types of interest keywords: Pathway Name (e.g., “Glycolysis/Gluconeogenesis”), Marker Type (e.g., PGTM) and Organism Name (e.g., “Arabidopsis lyrata”), from the available drop-down menus to effortlessly access and retrieve all markers and maps data stored in the PPMdb database.

- b) The ‘KEGG pathway maps overview’ section, which provides the user with the necessary information for each pathway such as Pathway ID (hyperlinked to KEGG Ref. Pathways); pathway class (e.g., Metabolism), pathway sub-class (e.g., carbohydrate metabolism), and pathway map. The search results of this section are designed in a user-friendly style containing essential descriptive information for each marker (e.g., Enzyme ID, Marker Type, Repeat type & sequence (this info. exclusive to PMTM), Primer Sequence, Primer annealing temp., Primer position within CDS, Product size, JBrowse View link, NCBI Gene Accession Number & CDS Sequence, and Gene Information).

- c) Additionally, general statistic table are presented to summarize the total number of designed primers, number of PGTM primers/number of PMTM primers within the selected pathway of a particular plant. All search results can download freely for each plant straightforwardly as a “CSV” file

Database Search

Search Section



How to search in database (instructions)

Pathway Name

Pathway Name

Marker Type

Repeat type

Organism Name

PlantPathMarks Search Page

* Please select the Pathway Name from Pathway Name drop down menu.
 * Please select Marker type (PGTM, PMTM) from Marker Type drop down menu.
 * Please select the Organism Name from Organism Name drop down menu.

Pathway Name: Glycolysis / Gluconeogenesis | Marker Type: PGTM | Organism Name: Arabidopsis lyrata | Search

* If you search for specific repeat class use the following form.
 * Please select the Pathway Name & Organism Name & SSR repeat from the drop down menu.

Pathway Name: Glycolysis / Gluconeogenesis | Repeat Type: Di | Organism Name: Arabidopsis lyrata | Search

Statistics

Total number of designed primers in this pathway: **407**
 Number of PGTM primers: **122**

[Download Search Result](#)

Enzyme Id: K03841

Marker Type	PGTM
Repeat type & sequence	Repeat type: Repeat sequence:
Repeat start & end within CDS	Repeat start: Repeat end:
Forward primer	Primer sequence: <i>GGACCACACTGATGAGCCAA</i> Tm(°C): 59.963 GC (%): 55 Size: 20
Reverse primer	Primer sequence: <i>TTTCCCGTAAAGGCCTGACC</i> Tm(°C): 59.963 GC (%): 55 Size: 20
Primer start, end within sequence and product size	Start: 431 End: 904 Product size (bp): 474
JBrowse View	JBrowse
NCBI Gene Accession Number & CDS Sequence	Accession Number: 9330022 CDS Sequence: 9330022
Gene Information	Gene Information

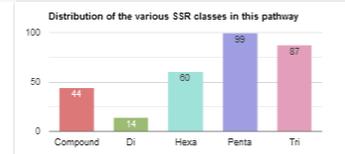
Statistics

Total number of designed primers in this pathway: **407**
 Number of PMTM primers: **285**
 Number of Di repeat primers: **13**
 Number of Di repeats Failed to design primers: **1**

[Download Search Result](#)

Enzyme Id: K00627

Marker Type	PMTM
Repeat type & sequence	Repeat type: <i>DI</i> Repeat sequence: <i>(TC)₄</i>
Repeat start & end within CDS	Repeat start: 349 Repeat end: 356
Forward primer	Primer sequence: <i>AGTAGTCAATGGGAAGGCC</i> Tm(°C): 60.034 GC (%): 55 Size: 20
Reverse primer	Primer sequence: <i>CCAACCTGCAATGTTACCC</i> Tm(°C): 59.96 GC (%): 55 Size: 20
Primer start, end within sequence and product size	Start: 234 End: 389 Product size (bp): 156
JBrowse View	JBrowse
NCBI Gene Accession Number & CDS Sequence	Accession Number: 9330544 CDS Sequence: 9330544
Gene Information	Gene Information



Database Search

KEGG pathway maps overview

Pathway Class



PlantPathMarks KEGG pathway-based maps overview

Metabolism

Genetic Information Processing

Environmental Information Processing

Cellular Processes

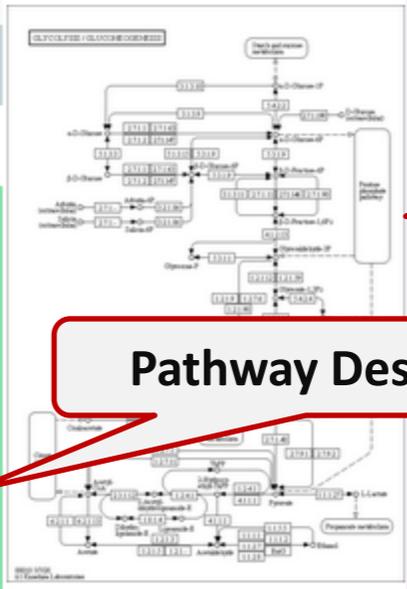
Organismal Systems

Pathway Name: Glycolysis / Gluconeogenesis

Pathway ID: [ko00010](#)

Class & Sub-Class: Class: *Metabolism* Sub-Class: *Carbohydrate metabolism*

Description: Glycolysis is the process of converting glucose into pyruvate and generating small amounts of ATP (energy) and NADH (reducing power). It is a central pathway that produces important precursor metabolites: six-carbon compounds of glucose-6P and fructose-6P and three-carbon compounds of glycerone-P; glyceraldehyde-3P; glycerate-3P; phosphoenolpyruvate; and pyruvate [MD:M00001]. Acetyl-CoA; another important precursor metabolite; is produced by oxidative decarboxylation of pyruvate [MD:M00307]. When the enzyme genes of this pathway are examined in completely sequenced genomes; the reaction steps of three-carbon compounds from glycerone-P to pyruvate form a conserved core module [MD:M00002]; which is found in almost all organisms and which sometimes contains operon structures in bacterial genomes. Gluconeogenesis is a synthesis pathway of glucose from noncarbohydrate precursors. It is essentially a reversal of glycolysis with minor variations of alternative paths [MD:M00003].

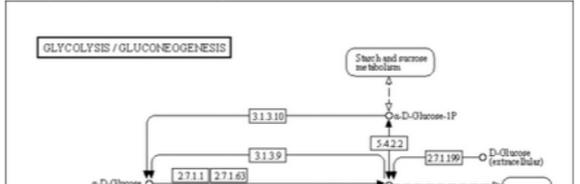


Pathway Map

Pathway Description

KEGG PATHWAY: ko00010

Entry	ko00010	Pathway
Name	Glycolysis / Gluconeogenesis	
Description	Glycolysis is the process of converting glucose into pyruvate and generating small amounts of ATP (energy) and NADH (reducing power). It is a central pathway that produces important precursor metabolites: six-carbon compounds of glucose-6P and fructose-6P and three-carbon compounds of glycerone-P, glyceraldehyde-3P, glycerate-3P, phosphoenolpyruvate, and pyruvate [MD:M00001]. Acetyl-CoA, another important precursor metabolite, is produced by oxidative decarboxylation of pyruvate [MD:M00307]. When the enzyme genes of this pathway are examined in completely sequenced genomes, the reaction steps of three-carbon compounds from glycerone-P to pyruvate form a conserved core module [MD:M00002], which is found in almost all organisms and which sometimes contains operon structures in bacterial genomes. Gluconeogenesis is a synthesis pathway of glucose from noncarbohydrate precursors. It is essentially a reversal of glycolysis with minor variations of alternative paths [MD:M00003].	
Class	Metabolism; Carbohydrate metabolism BRITE hierarchy	
Pathway map	ko00010	Glycolysis / Gluconeogenesis



Map Browse

The “Map Browse” page offers the users with searching utility in two styles inside the same page:

- a) selecting the plant of interest from the page-side plant list for direct access of the pathways and markers.
- b) selecting the plant of interest from drop-down menus to directly access and obtain all markers and maps.

In both styles, the search results will be presented in a convenient, visualization-supported and well-dissected manner involving essential information for each pathway such as; pathway ID (hyperlinked to KEGG Ref. Pathways), pathway class (e.g., Metabolism), pathway sub-class (e.g., carbohydrate metabolism), pathway map browse (hyperlinked to interactive pathway map localizing all developed PMTM and PGTM markers on the pathway map), pathway description and pathway map

Map Browse

Statistics & Search Section



Plant Name Links

Statistics Image

PlantPathMarks Map Browse

Plant Name

- Arabidopsis lyrata
- Arabidopsis thaliana
- Arachis duranensis
- Arachis ipaensis
- Asparagus officinalis
- Beta vulgaris
- Brachypodium distachyon
- Brassica napus
- Brassica oleracea
- Brassica rapa

General Statistics

Mapped Pathways	Mapped Enzymes	Designed Primers	PGTM Primers	PMTM Primers
120	2099	28598	7318	21282

Arabidopsis thaliana Statistics & Search

* Please select the Pathway Name & Marker type from the drop down menu.

Pathway Name: Glycolysis / Gluconeogenesis

Marker Type: PGTM Search

* If you search for specific repeat class use the following form.

* Please select the Pathway Name & SSR repeat from the drop down menu.

Pathway Name: Glycolysis / Gluconeogenesis

Repeat Type: Di Search



PlantPathMarks Map Browse

Plant Name

- Arabidopsis lyrata
- Arabidopsis thaliana
- Arachis duranensis
- Arachis ipaensis
- Asparagus officinalis
- Beta vulgaris
- Brachypodium distachyon
- Brassica napus
- Brassica oleracea
- Brassica rapa

General Statistics

Mapped Pathways	Mapped Enzymes	Designed Primers	PGTM Primers	PMTM Primers
121	2010	29338	8893	22448

Brachypodium distachyon Statistics & Search

* Please select the Pathway Name & Marker type from the drop down menu.

Pathway Name: Glycolysis / Gluconeogenesis

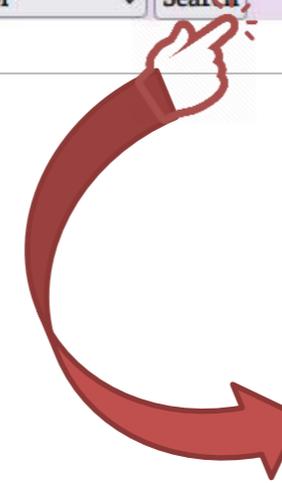
Marker Type: PGTM Search

* If you search for specific repeat class use the following form.

* Please select the Pathway Name & SSR repeat from the drop down menu.

Pathway Name: Glycolysis / Gluconeogenesis

Repeat Type: Di Search



Statistics

Total number of designed primers in this pathway:	425
Number of PMTM primers:	308
Number of Di repeat primers:	21
Number of Di repeats Failed to design primers:	2

Distribution of the various SSR classes in this pathway

Compound	Di	Hexa	Penta	Tetra	Tri
23	23	73	114	1	97

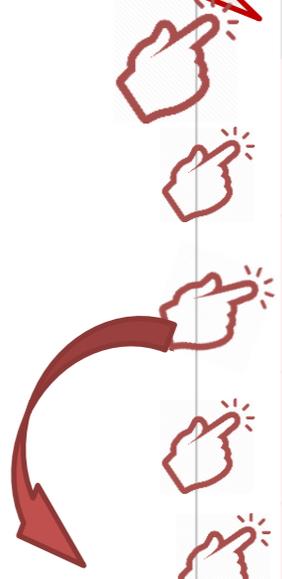
Enzyme Id: Koo895

Marker Type	PMTM
Repeat type & sequence	Repeat type: Di Repeat sequence: (GA) ₄
Repeat start & end within CDS	Repeat start: 1174 Repeat end: 1181
Forward primer	Primer sequence: AGAGGTTTCTGCCAGAAGC Tm(°C): 59.96 GC(%): 55 Size: 20
Reverse primer	Primer sequence: AGCAACTGGACATTCCGT Tm(°C): 59.89 GC(%): 50 Size: 20
Primer start, end within sequence and product size	Start: 893 End: 1205 Product size (bp): 313
JBrowse View	JBrowse
NCBI Gene Accession Number & CDS Sequence	Accession Number: ATG12000 CDS Sequence: ATG12000
Gene Information	Gene Information

Map Browse

Maps Overview Section

Pathway Class



Brachypodium distachyon Maps overview

- Metabolism
- Genetic Information Processing
- Environmental Information Processing
- Cellular Processes
- Organismal Systems

Pathway Name: Glycolysis / Gluconeogenesis

Pathway ID: [k000010](#)

Class & Sub-Class: Class: [Metabolism](#) Sub-Class: [Carbohydrate metabolism](#)

Pathway Map Browse: [Browse](#)

Description: [Read description](#)

Glycolysis / Gluconeogenesis

Markers Statistics

Pathway Id: [k000010](#)

Class & Sub-Class: Class: [Metabolism](#) Sub-Class: [Carbohydrate metabolism](#)

Description: [Read description](#)

Select enzymes Id from the checkbox menu to highlight in the map or click on enzyme direct in the map to highlight if it successfully designed primer (you can get all information about the enzyme and it's designed primers by click direct enzyme Id in the left menu **(Mapped Enzymes menu)** or click the mouse in your interest enzyme within the pathway map and click mouse right and open it in new tab)

Mapped Enzymes

- K00002
- K00016
- K00131
- K00134
- K00182
- K00382
- K00627
- K00844
- K00850
- K00873
- K00895
- K00927
- K00844
- K00850
- K00873
- K00895
- K00927

Brachypodium distachyon Maps overview

- Metabolism
- Genetic Information Processing
- Environmental Information Processing
- Cellular Processes
- Organismal Systems

Pathway Name: ABC transporters

Pathway ID: [k002010](#)

Class & Sub-Class: Class: [Environmental Information Processing](#) Sub-Class: [Membrane transport](#)

Pathway Map Browse: [Browse](#)

Description: The ATP-binding cassette (ABC) transporters are one of the largest known protein families widespread in bacteria, archaea; They couple ATP hydrolysis to active transport of a wide variety of substrates such as lipids, sterols, peptides; proteins, nucleic acids, and drugs. The structure of a prokaryotic ABC transporter consists of three components; two transmembrane proteins each having two transmembrane segments; two peripheral proteins that bind and hydrolyze ATP; and a central nucleotide-binding domain.

Statistics

Number of enzymes: 1
Total Number of designed primers: 6
Number of PGTM primers: 5
Number of PNTM primers: 1

Enzyme Id: [K00002](#)

Gene Information	
NCBI Gene Accession Number & Symbol	Accession Number: 100843184 NCBI Gene Symbol: LOC100843184
Gene Aliases	BRADI_3937327
Gene description & Other designations	Description: NADPH-dependent aldo-keto reductase; chloroplastic Other designations: NADPH-dependent aldo-keto reductase; chloroplastic
Chromosome, Strand & Exon count	Chromosome: 3 Strand: plus Exon count: 8
Gene Location within genomic sequence	Genomic accession No. NC_016133.3 Gene Start and end within genomic accession: 39477677 39483261
CDS Sequence	100843184

Marker Information	
Marker Type	PNTM
Repeat type & sequence	Repeat type: <i>Penta</i> Repeat sequence: (GCCGCT)₂
Repeat start & end within CDS	Repeat start: 754 Repeat end: 763
Forward primer	Primer sequence: TCAGTCGCAGAGAAGCTGG Tm(°C): 60.038 GC (%): 55 Size: 20
Reverse primer	Primer sequence: CACTCACGCTCTTGGGAGT Tm(°C): 59.966 GC (%): 55 Size: 20
Primer start, end within sequence and product size	Start: 74 End: 807 Product size (bp): 94

JBrowse

The “**JBrowse**” page enables the users to visualize and map all identified microsatellite motifs besides all developed PMTM and PGTM markers against the Refseq CDS for each plant genome by selecting the plant of interest from the page-side plant list. The mapped microsatellite motifs and developed PMTM/PGTM markers were linked to its essential information



JBrowse

The image shows the JBrowse interface for *Arabidopsis thaliana*. The title "Arabidopsis thaliana" is centered at the top. On the left, a "Plant Name" list includes: Arabidopsis lyrata, Arabidopsis thaliana, Arachis duranensis, Arachis ipaensis, Asparagus officinalis, Beta vulgaris, Brachypodium distachyon, Brassica napus, Brassica oleracea, Brassica rapa, Cajanus cajan, Camelina sativa, and Capsella rubella. The "Available Tracks" panel on the left shows: filter tracks, refseq, Microsatellites, PGTM, and PMTM. The main visualization area shows a genomic track for the region ath_AT1G01020:296..443 (14). The track includes: a RefSeq track with red and green blocks; a track with a dense multi-colored bar representing various motifs; a PMTM track with a long yellow bar; and a Microsatellites track with a small yellow bar. The top of the track area has a scale from 0 to 700 and navigation controls.

JBrowse

PlantPathMarks (PPMdb)
A hub for Plant Pathways-Based Markers

Home Database Search Map Browse **JBrowse** Statistics Data Resources Bulk Download Species Comparisons Manual

Plant Name Links

- Plant Name
- Arabidopsis lyrata
- Arabidopsis thaliana
- Arachis duranensis
- Arachis ipaensis
- Asparagus officinalis
- Beta vulgaris
- Brachypodium distachyon
- Brassica napus
- Brassica oleracea
- Brassica rapa
- Cajanus cajan
- Camelina sativa
- Capsella

JBrowse

Arabidopsis thaliana

Genome Track View Help Share

0 50 100 150 200 250 300 350 400 450 500 550 600 650 700

ath_AT1G01020 ath_AT1G01020:296..443 (14) Go

Available Tracks

- filter tracks
- refseq
- Miscellaneous 3
 - Microsatellites
 - PGTM
 - PMTM

refseq

PMTM

Tri

Primary Data

- Type: Tri
- Position: ath_AT1G01020:439..447
- Length: 9 bp

Attributes

- repeat_sequence: (AGA)3
- seq_id: ath_AT1G01020
- source: SSR_Repeat
- uniqueID: offset-2769956

Region sequence

PMTM

Primary Data

- Type: PMTM
- Position: ath_AT1G01020:124..419
- Length: 296 bp

Attributes

- forward_primer_gc%: 50
- forward_primer_sequence: ACATCGAGTGTGAACGCATG
- forward_primer_size: 20
- forward_primer_tm: 58
- reverse_primer_gc%: 50
- reverse_primer_sequence: ACCCTTAGTCGCAATGGCAA

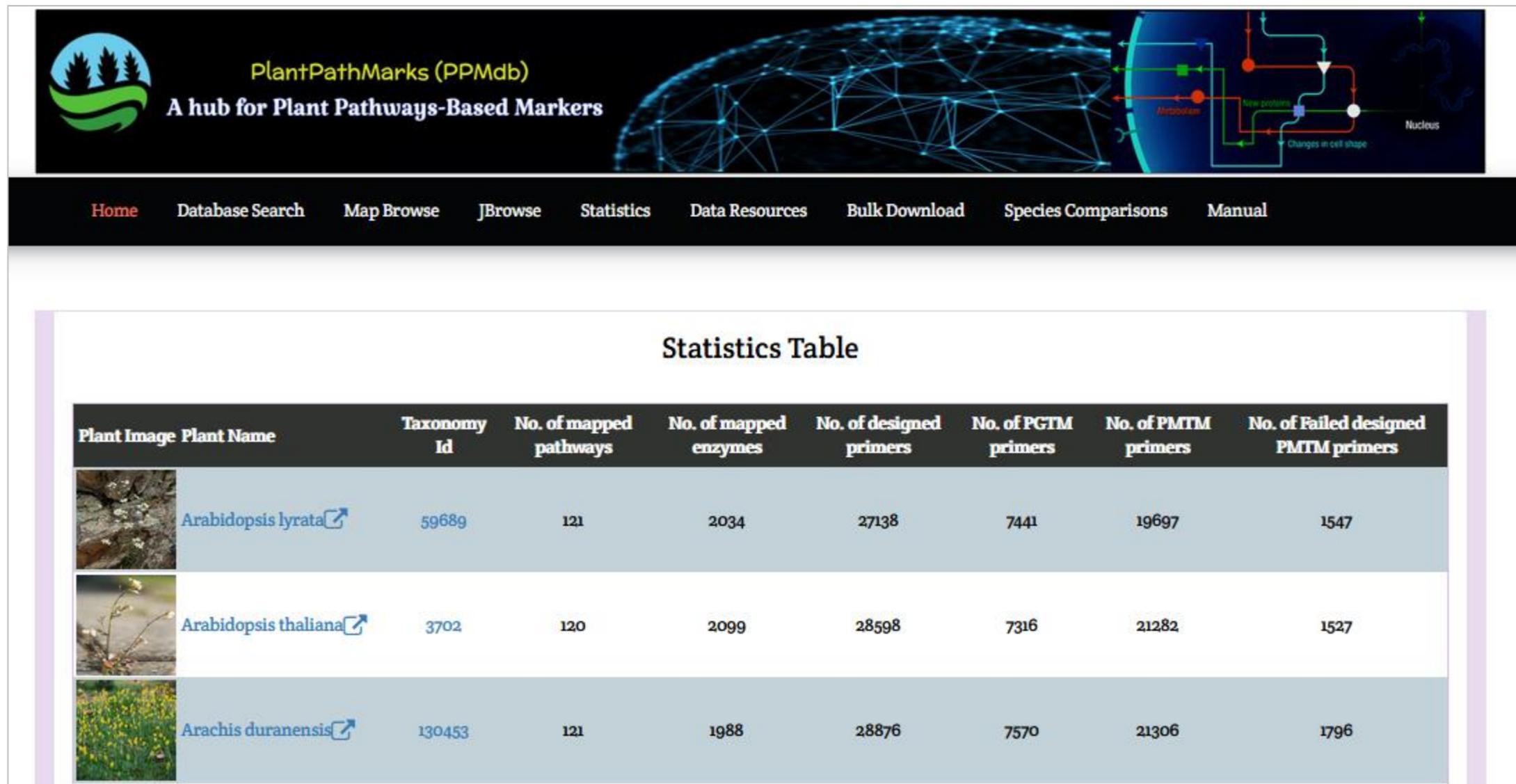
FASTA

```
>ath AT1G01020 ath_AT1G01020:124..419 class=PMTM length=296
ACATCGAGTGTGAACGCATGATTATTTTCATCGATTTAATCCTTCACAGACCAAAGGTATATA
GACACGTCCTCTACAATGCAATTAATCCAGCAACTGTCAATATTCAGCATCTGTTGTGGAAGTT
GGTCTTCGCCTATCTTCTTAGACTGTTATAGAAGCTTGCTACTGAGAAAAAGTGATGAAGAA
TCGAGCTTTTCTGATAGCCCTGTTCTTCTATCTATAAAGGTTCTGATTGGTGTCTTATCTGCAA
ACGCTGCATTTATCATCTCTTTTGCCATTGCGACTAAGGG
```

OK

Statistics

The 'Statistics' page was designed to provide users a primary indication of the taxonomy ID, number of mapped pathways, number of mapped enzymes, number of designed primers, number of PGTM primers, number of PMTM primers, and number of failed designed PMTM primers for each plant species stored in the PPMdb database.



PlantPathMarks (PPMdb)
A hub for Plant Pathways-Based Markers

Home Database Search Map Browse JBrowse Statistics Data Resources Bulk Download Species Comparisons Manual

Statistics Table

Plant Image	Plant Name	Taxonomy Id	No. of mapped pathways	No. of mapped enzymes	No. of designed primers	No. of PGTM primers	No. of PMTM primers	No. of Failed designed PMTM primers
	Arabidopsis lyrata	59689	121	2034	27138	7441	19697	1547
	Arabidopsis thaliana	3702	120	2099	28598	7316	21282	1527
	Arachis duranensis	130453	121	1988	28876	7570	21306	1796

Data Resources

The “**Data Resources**” page provide users with hyperlinks of all types/formats of the data utilized to build the PPMdb database. The page includes hyperlinks of Taxonomy Id, Genbank File, Fasta File, GFF File, and KEGG Genes Annotation for each plant species analyzed within the PPMdb database.

PlantPathMarks (PPMdb)
A hub for Plant Pathways-Based Markers

Home Database Search Map Browse JBrowse Statistics Data Resources Bulk Download Species Comparisons Manual

Data Resources

Plant Image	Plant Name	Taxonomy Id	Genbank File	Fasta File	GFF File	KEGG Genes Annotation
	Arabidopsis lyrata	59689	Genbank File	Fasta File	GFF File	KEGG Annotation
	Arabidopsis thaliana	3702	Genbank File	Fasta File	GFF File	KEGG Annotation
	Arachis duranensis	130453	Genbank File	Fasta File	GFF File	KEGG Annotation
	Arachis ipaensis	130454	Genbank File	Fasta File	GFF File	KEGG Annotation

Bulk Download

The “**Bulk Download**” page enables the users to download all developed primers and their secondary information under a specific pathway in the organism of interest through the selection of Organism Name (e.g., “*Arabidopsis lyrata*”), Pathway Name (e.g., “Glycolysis/Gluconeogenesis”), and Marker Type (e.g., PGTM), from the drop-down menus straightforwardly in a convenient manner.

PlantPathMarks (PPMdb)
A hub for Plant Pathways-Based Markers

Home Database Search Map Browse JBrowse Statistics Data Resources Bulk Download Species Comparisons Manual

Organism Name Arabidopsis lyrata

Pathway Name Glycolysis / Gluconeogenesis

Marker Type PGTM

Search

Download Result

Thanks For Use PlantPathMarks Database

Opening ko00010.PGTM.aly.csv.gz

You have chosen to open:
ko00010.PGTM.aly.csv.gz
which is: WinRAR archive (3.5 KB)
from: http://ppmdb.easyomics.org

What should Firefox do with this file?

Open with WinRAR archiver (default)

Save File

Do this automatically for files like this from now on.

OK Cancel

Species Comparisons

The “**Species Comparisons**” page offers the users powerful utility to compare two or three organisms in a particular pathway. The comparison results are provided under three main sections: general statistics, markers statistics, and distribution of the various SSR classes.

The general statistics section layout the no. of mapped pathways, no. of mapped enzymes, no. of total designed primers, no. of PGTM designed primers, and no. of PMTM designed primers under particular pathway between the organisms of interest.

Also, the page offers a simple comparison of the distribution of the various SSR classes within the pathway of interest.

All comparisons are offered in a simple and effective visualization style

Species Comparisons



First Species

Second Species

Third Species

Pathway Name

Species Comparisons Page

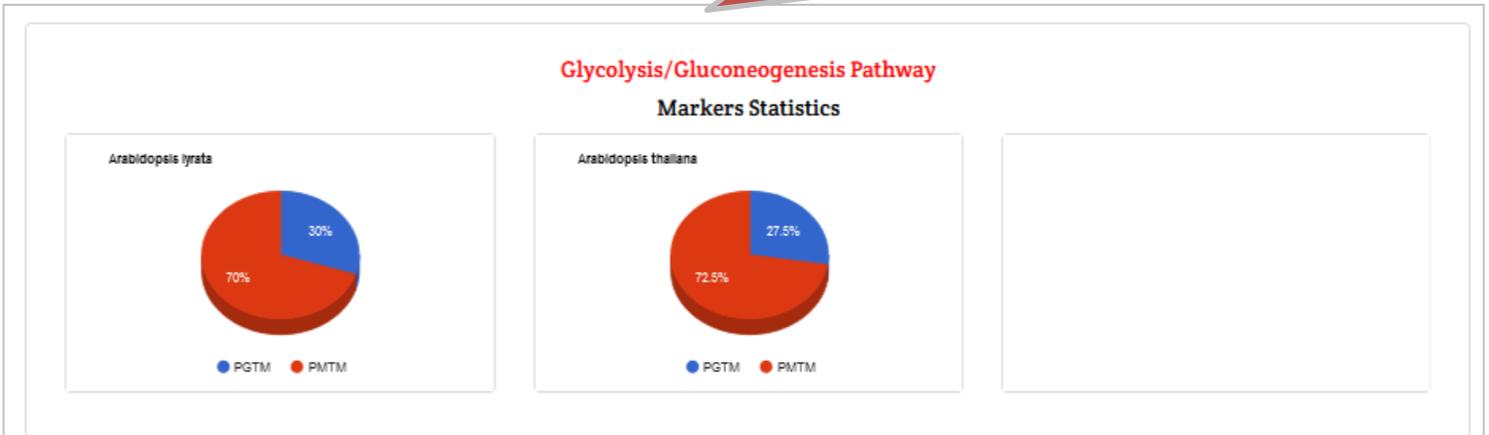
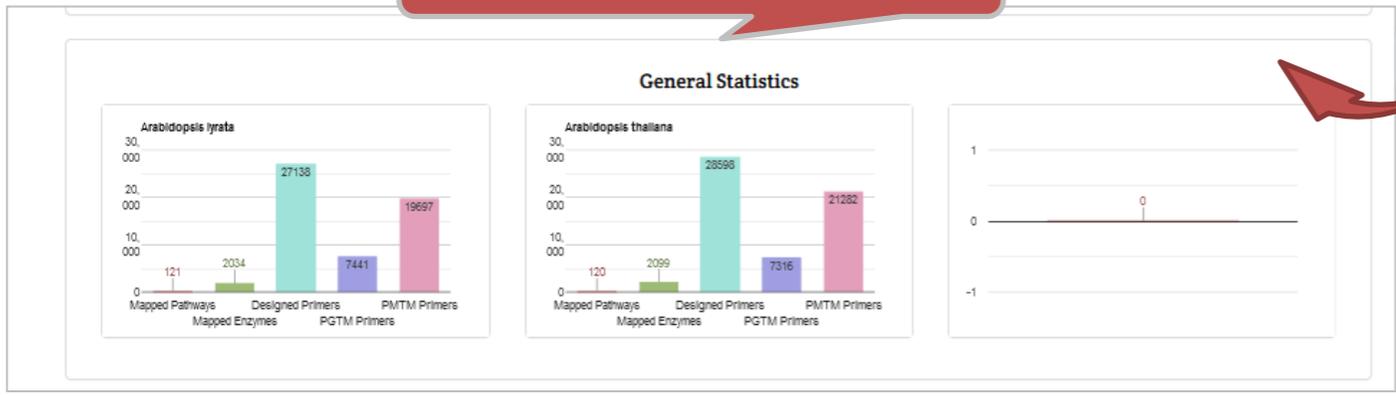
* Please Select Species & Pathway name and Press Submit Query

First Species
Second Species
Third Species

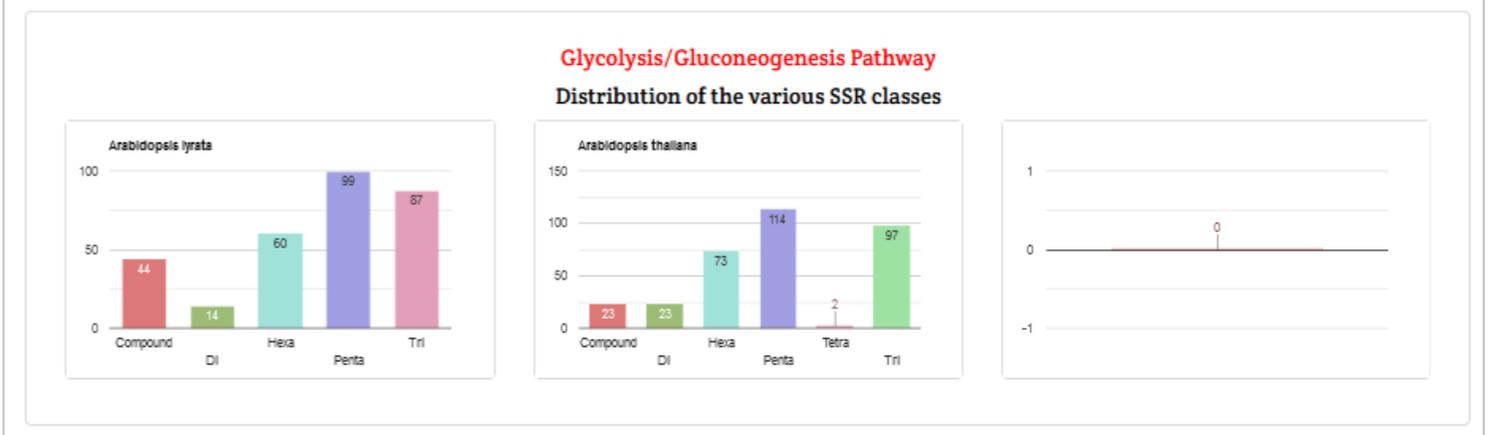
Pathway Name

General Statistics

Markers Statistics



Glycolysis/Gluconeogenesis Pathway Browse



For any further questions please contact us

Prof. Dr. Mohamed A.M. Atia: matia@ageri.sci.eg

Dr. Morad M. Mokhtar: morad.mokhtar@ageri.sci.eg