

# KusakiDB Tools: Search System

Select Taxon Level:  
Species or Family

Select Species or Family

Filter your results

Search: glutathione transferase

Column visibility

kusakiDB\_id

KusakiDB\_validation

Species

Family

Phylum

Superkingdom

## KusakiDB v1.0

A Novel Approach for Validation and Completeness of Protein Orthologous Groups

Search KusakiDB

OG Assessment

OG Management

OG Management User Data

Taxon Level

Species

Taxon Name

Phaseolus vulgaris

Column visibility

OrthoDB\_Group

Accession

Accession\_Protein\_Name

OrthoDB\_Protein\_Name

178219at33090	V7BG38	GST C-terminal domain-containing protein	Glutathione S-transferase, N-terminal
178219at33090	V7CKA2	GST C-terminal domain-containing protein	Glutathione S-transferase, N-terminal
184679at33090	V7AMD3	GST C-terminal domain-containing protein	Glutathione S-transferase, C-terminal
167195at33090	V7AUM5	GST N-terminal domain-containing protein	Glutathione S-transferase, C-terminal
167328at33090	H2FH46	Glutathione transferase	Glutathione S-transferase, C-terminal
183963at33090	H2FH47	Glutathione transferase	Glutathione S-transferase, N-terminal
184679at33090	H2FH48	Glutathione transferase	Glutathione S-transferase, C-terminal
173030at33090	T2DNZ6	In2-1 protein	Glutathione S-transferase, C-terminal domain superfamily
140562at33090	V7CVX8	Uncharacterized Protein	Glutathione S-transferase, N-terminal
140562at33090	V7AXY2	Uncharacterized Protein	Glutathione S-transferase, N-terminal

Showing 1 to 10 of 57 entries (filtered from 26,364 total entries)

Previous

1

2

3

4

5

6

Next

Results of KusakiDB data

# KusakiDB Tools: OG Assessment

Select a Family  
from KusakiDB

Choose Family

Family Name

P

- Poaceae
- Phrymaceae
- Pedaliaceae
- Papaveraceae
- Apiaceae
- Asparagaceae
- Cephalotaceae
- Euphorbiaceae

Results of KusakiDB data

## KusakiDB v1.0

A Novel Approach for Validation and Completeness of Protein Orthologous Groups

Search KusakiDB OG Assessment OG Management OG Management User Data

Choose Family

Family Name

Poaceae

Or analyse your own data. [Click here to create Hayai-annotation file](#)

Upload Hayai\_annotation\_v2.0.tsv

Browse... No file selected

Submit

User Data

KusakiDB Complete Genomes Data

Family_name	Scientific_name	N_of_clusters	kusakiDB_validated	Median_Family	Total_Family	Median_Species
Poaceae	Oryza brachyantha	11801	96.95	90.91	11	83.76
Poaceae	Triticum urartu	10512	91.83	90.91	11	82.91
Poaceae	Brachypodium distachyon	12833	95.11	90.91	11	82.91
Poaceae	Zea mays	13021	96.84	90.91	11	82.05
Poaceae	Oryza sativa Japonica Group	13128	96.95	90.91	11	82.05
Poaceae	Sorghum bicolor	13314	94.88	90.91	11	82.05
Poaceae	Dichanthelium oligosanthos	11988	93.45	90.91	11	82.05
Poaceae	Panicum hallii	13708	92.08	90.91	11	82.05
Poaceae	Setaria italica	13190	93.49	90.91	11	82.05
Poaceae	Aegilops tauschii	15967	80.14	81.82	11	78.63
Poaceae	Triticum aestivum	10932	59.80	45.45	11	9.40

# KusakiDB OG Assessment accepts user's data

## KusakiDB Tools: OG Assessment

Upload your own data annotated by  
Hayai-annotation v2.0

Users data result:  
KusakiDB predicts family and  
calculate number of validated OGs,  
median of OG within same family  
and median of OG within all species

**KusakiDB v1.0**

A Novel Approach for Validation and Completeness of Protein Orthologous Groups

Search KusakiDB **OG Assessment** OG Management OG Management User Data

Choose Family

Family Name

Poaceae

Or analyse your own data. Click here to create Hayai-annotation file

Upload Hayai\_annotation\_v2.0.tsv

Browse... Hayai\_annotation\_v2.0.tsv

Upload complete

Submit

User Data

Family_name	Species_source	N_of_clusters	kusakiDB_validated	Median_Family	Total_Family	Median_Species
Brassicaceae	User_data	13648	98.98	98.98	8	98.98

# OG Assessment Equations

Median of OGs within a family (median family):

$$\text{Median\_Family (\%)} = \text{median}\left(\frac{\text{rep}_{\text{fam}}}{\text{Total}_{\text{family}}} \times 100\right), \text{ where:}$$

$\text{rep}_{\text{fam}}$  is the count of unique sequences in an OG, per species, within a family.

$\text{Total}_{\text{family}}$  is the total number of species in each family.

Median of OGs within all species (median species):

$$\text{Median\_Species (\%)} = \text{median}\left(\frac{\text{rep}}{\text{Total}_{\text{species}}} \times 100\right), \text{ where:}$$

$\text{rep}$  is the count of unique sequences in an OG, per species.

$\text{Total}_{\text{species}}$  is the total number of species in KusakiDB.

# KusakiDB Tools: OG Management

Users can select the parameters to compare OGs among all species in KusakiDB, such as:

- Validation tag
- Number of species in each family
- Percentage of species in each family
- Percentage within all species

## KusakiDB Validation

- ☒ Validated  
☐ Not validated

### KusakiDB Validation

Validated: existence of at least one protein, in an OG, at a transcript or protein level

Pop-up

User can download the results

The results with the selected parameters are shown in two table:

- List of species and number of OGs
- List of protein names and correspondent frequency

**KusakiDB v1.0**  
A Novel Approach for Validation and Completeness of Protein Orthologous Groups

Search KusakiDB   OG Assessment   **OG Management**   OG Management User Data

**KusakiDB Validation**  
☒ Validated  
☐ Not validated

**Total number of species in a family**  
5

**Percentage of Species in a Family**  
1 11 21 31 41 51 61 70 81 91 100

**Percentage of Total Species**  
1 11 21 31 41 50 61 71 81 91 100

Submit

Download

**List of Species with selected OG parameters**

Scientific_name	Freq
Sorghum bicolor	1610
Panicum hallii	1597
Aegilops tauschii	1587
Oryza sativa Japonica Group	1578
Setaria italica	1577
Camelina sativa	1566
Arabidopsis lyrata subsp. lyrata	1548
Brachypodium distachyon	1531
Capsella rubella	1519
Brassica rapa	1518

**List of Protein names with selected OG parameters**

Protein_name	Freq
Zinc finger, RING/FYVE/PHD-type	463
Leucine-rich repeat	375
Pentatricopeptide repeat	327
Peroxidase	243
NAC domain	238
Glycosyltransferase	233
F-box domain	212
Transcription factor, MADS-box	208
transmembrane protein, putative	207
Cytochrome P450	198

# KusakiDB Tools: OG Management User Data

Users can select the parameters to compare OGs among all species in KusakiDB, such as:

- Validation tag
- Number of species in each family
- Percentage of species in each family
- Percentage within all species

Users can upload the functional annotation performed by Hayai-annotation v2.0

The results show a “filter” of the genes that are selected under the conditions regarding the conservation level of each OG.

**KusakiDB v1.0**  
A Novel Approach for Validation and Completeness of Protein Orthologous Groups

Search KusakiDB   OG Assessment   OG Management   **OG Management User Data**

**User Data Analysis**

**KusakiDB Validation**  
☒ Validated  
☐ Not validated

**Total number of species in a family**  
5

**Percentage of Species in a Family**  
1 11 21 31 41 51 61 71 81 91 100

**Percentage of Total Species**  
1 11 21 31 41 51 61 71 81 91 100

[Click here to create Hayai-annotation file](#)  
**Upload Hayai\_annotation\_v2.0.tsv**

Browse... Hayai\_annotation\_v2.0.tsv  
Upload complete

Submit

Download

Filter your results

Search: flower

Show 10 entries

	queryid_pep	pub_og_id	uniprot_protein_name	uniprot_accession	kusakiDB_curation	Perc	Protein_name
6549	AT1G26310.1	163301at33090	Transcription factor CAULIFLOWER	Q39081	V	82.05	Transcription factor, MADS-box
1679	AT1G72390.1	11593at33090	Protein PHYTOCHROME-DEPENDENT LATE-FLORERING	F4IDB2	V	84.62	histone-lysine N-methyltransferase 2D isoform X1
15130	AT2G25930.1	24394at33090	Protein EARLY FLOWERING 3	O82804	V	85.47	protein EARLY FLOWERING 3
14057	AT2G29950.1	228137at33090	Protein ELF4-LIKE 1	O80877	V	66.67	protein EARLY FLOWERING 4-like
16999	AT2G35670.1	44305at33090	Polycomb group protein FERTILIZATION-INDEPENDENT SEED 2	P0DKJ8	V	92.31	polycomb group protein EMBRYONIC FLOWER 2-like isoform X1
14056	AT2G40080.1	228137at33090	Protein EARLY FLOWERING 4	O04211	V	66.67	protein EARLY FLOWERING 4-like
910	AT3G04610.1	108737at33090	Flowering locus K homology domain	Q9SR13	V	91.45	K Homology domain, type 1
16870	AT3G10390.3	42805at33090	Protein FLOWERING locus D-like protein	A0A1I9LNW6	V	99.15	lysine-specific histone demethylase 1 homolog 2
15129	AT3G21320.1	24394at33090	EARLY FLOWERING protein	Q5Q0C8	V	85.47	protein EARLY FLOWERING 3
17001	AT4G16810.1	44305at33090	VEFS-Box of polycomb protein	B3H7I2	V	92.31	polycomb group protein EMBRYONIC FLOWER 2-like isoform X1

Showing 1 to 10 of 19 entries (filtered from 22,124 total entries)

Previous 1 2 Next