

# KusakiDB Tools: Search System

KusakiDB v1.0

A novel approach for assessing existence and completeness of orthogroups in plant species

Search KusakiDB OG Assessment OG Management OG Management User Data Help

Search proteins within species or family level

Select Taxon Level: Species or Family

Select Species or Family

Filter your results

Search:

Column visibility

OrthoDB_Group	OrthoDB_Protein_Name	Accession	Accession_Protein_Name	Evidence_existence
<a href="#">201674at33090</a>	RNA-binding domain, S1	<a href="#">P46618</a>	Translation initiation factor IF-1, chloroplastic	Yes
<a href="#">80872at33090</a>	EGF-like calcium-binding, conserved site	<a href="#">A0A0A7ENQ5</a>	Wall-associated kinase	Yes
<a href="#">70014at33090</a>	SecY/SEC61-alpha family	<a href="#">B8A1R6</a>	SecY protein transport family protein	Yes
<a href="#">187354at33090</a>	Ubiquitin-conjugating enzyme E2	<a href="#">A0A3L6DUL2</a>	Ubiquitin-conjugating enzyme E2 36	Yes
<a href="#">184484at33090</a>	Ubiquitin-conjugating enzyme E2	<a href="#">A0A3L6FR20</a>	Ubiquitin-conjugating enzyme E2 2	Yes
<a href="#">218158at33090</a>	Sm-like protein LSM2	<a href="#">A0A3L6F5V8</a>	Sm-like protein LSM2	Yes
<a href="#">155756at33090</a>	Small GTPase	<a href="#">B4FE06</a>	Uncharacterized Protein	Yes
<a href="#">234012at33090</a>	RNA polymerase archaeal subunit P/eukaryotic subunit RPABC4	<a href="#">A0A3L6E0I6</a>	DNA-directed RNA polymerases II, IV and V subunit 12	Yes
<a href="#">209311at33090</a>	Histone H2A	<a href="#">A0A3L6ET91</a>	Histone H2A	Yes
<a href="#">157046at33090</a>	DNA recombination and repair protein Rad51-like, C-terminal	<a href="#">B4FMC1</a>	Meiotic recombination protein DMC1	Yes

Column visibility

Evidence\_existe...

Species

Family

kusakiDB\_id

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 assessing existence and completeness of orthogroups in plant species

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

Assess conservation of orthogroups within complete genomes

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

Users' Data

KusakiDB Complete Genomes Data

Family	Scientific_name	Number_of_OG_Groups	Evidence_existence (%)	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

User's 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 assessing existence and completeness of orthogroups in plant species

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

**Assess conservation of orthogroups within complete genomes**

Choose Family  
Family Name  
Brassicaceae

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

**Users' Data**

Inferred_family	Species_source	Number_of_OG_Groups	Evidence_existence (%)	Median_Family (%)	Total_Family	Median_Species (%)
Brassicaceae	User_data	13595	98.99	100	8	82.05

**KusakiDB Complete Genomes Data**

Family	Scientific_name	Number_of_OG_Groups	Evidence_existence (%)	Median_Family (%)	Total_Family	Median_Species (%)
Brassicaceae	Arabis alpina	9914	96.65	100.00	8	83.76
Brassicaceae	Eutrema salsugineum	12789	95.75	100.00	8	82.91
Brassicaceae	Arabidopsis lyrata subsp. lyrata	13754	96.06	100.00	8	82.05
Brassicaceae	Capsella rubella	13107	95.90	100.00	8	82.05
Brassicaceae	Brassica rapa	13666	93.90	100.00	8	82.05
Brassicaceae	Arabidopsis thaliana	13501	99.17	100.00	8	82.05
Brassicaceae	Raphanus sativus	13944	92.86	100.00	8	81.20
Brassicaceae	Camelina sativa	14771	91.44	100.00	8	80.34

# OG Assessment Equations

Median of OGs within a family (median family):

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

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

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

Median of OGs within all species (median species):

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

rep is the count of unique sequences in an OG, per species.

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

# KusakiDB Tools: OG Management

Select the parameters to compare OGs among all species in KusakiDB, such as:

- Evidence existence tag
- Number of species in each family
- Percentage of species in each family
- Percentage within all species (Percentage of Total Species)

## Evidence existence

- ☒ Yes  
☐ No

## Evidence existence

Yes: 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 correspondent number of OGs/species
- List of protein names and correspondent number of species/OrthoDB protein name

KusakiDB v1.0

A novel approach for assessing existence and completeness of orthogroups in plant species

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

Manage parameters to assess conservation of orthogroups within complete genomes

**Evidence existence**  
☒ Yes  
☐ No

**Total number of species in a family**  
Slider: 1 to 11 (set to 5)

**Percentage of Species in a Family**  
Slider: 1 to 100 (set to 100)

**Percentage of Total Species**  
Slider: 1 to 100 (set to 30)

Submit

Download

List of Species with selected OG parameters

Show 10 entries Search:

Scientific_name	Freq
Arabidopsis lyrata subsp. lyrata	553
Arabidopsis thaliana	553
Arabis alpina	553
Brassica rapa	553
Camelina sativa	553
Capsella rubella	553
Eutrema salsugineum	553
Raphanus sativus	553
Aegilops tauschii	167
Brachypodium distachyon	167

Showing 1 to 10 of 50 entries

Previous 1 2 3 4 5 Next

List of Protein names with selected OG parameters

Show 10 entries Search:

OrthoDB_Group	OrthoDB_Protein_Name	Freq
226059at33090	Bowman-Birk type proteinase inhibitor	21
146849at33090	BTB/POZ domain-containing protein	20
68458at33090	Proton-dependent oligopeptide transporter family	20
177228at33090	uncharacterized protein LOC104594483	19
39676at33090	Ionotropic glutamate receptor	19
18311at33090	Beta-galactosidase	18
225713at33090	VQ	18
69952at33090	3-ketoacyl-CoA synthase	18
102253at33090	Membrane transport protein	16
15336at33090	Cation/H+ exchanger	16

Showing 1 to 10 of 1,222 entries

Previous 1 2 3 4 5 ... 123 Next

# KusakiDB Tools: OG Management User Data

Select the parameters to compare OGs among all species in KusakiDB, such as:

- Evidence existence tag
- Number of species in each family
- Percentage of species in each family
- Percentage within all species (Percentage of Total Species)

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 for each OG.

**KusakiDB v1.0**

A novel approach for assessing existence and completeness of orthogroups in plant species

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

Analyse conservation of orthogroups within complete genomes using your own data

Evidence existence  
☒ Yes  
☐ No

Total number of species in a family  
Slider: 1 to 11 (set to 5)

Percentage of Species in a Family  
Slider: 1 to 100 (set to 50)

Percentage of Total Species  
Slider: 1 to 100 (set to 40)

[Click here to create Hayai-annotation file](#)

Upload Hayai\_annotation\_v2.0.tsv

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

Upload complete

Submit

Download

Column visibility

Search: **protease**

Query	OrthoDB_Group	OrthoDB_Protein_Name	Accession	Accession_Protein_Name	Evidence_existence	Perc_Family
AT1G06260.1	<a href="#">107271at33090</a>	Cysteine peptidase, cysteine active site	<a href="#">Q9LNC1</a>	Cysteine proteinases superfamily protein	Yes	100
AT3G19400.2	<a href="#">107271at33090</a>	Cysteine peptidase, cysteine active site	<a href="#">F4JB71</a>	Cysteine proteinases superfamily protein	Yes	100
AT3G48350.2	<a href="#">119997at33090</a>	Cysteine peptidase, cysteine active site	<a href="#">A0A1I9LSH0</a>	Cysteine proteinases superfamily protein	Yes	100
AT3G12490.2	<a href="#">167069at33090</a>	Cysteine proteinase inhibitor	<a href="#">Q8H0X6</a>	Cysteine proteinase inhibitor 6	Yes	100
AT2G40880.1	<a href="#">167069at33090</a>	Cysteine proteinase inhibitor	<a href="#">Q41906</a>	Cysteine proteinase inhibitor 3	Yes	100
AT3G12490.1	<a href="#">167069at33090</a>	Cysteine proteinase inhibitor	<a href="#">A0A2H1ZEH9</a>	Cysteine proteinase inhibitor	Yes	100
AT5G05110.1	<a href="#">167069at33090</a>	Cysteine proteinase inhibitor	<a href="#">Q8LC76</a>	Cysteine proteinase inhibitor 7	Yes	100
AT5G12140.1	<a href="#">167069at33090</a>	Cysteine proteinase inhibitor	<a href="#">Q945Q1</a>	Cysteine proteinase inhibitor 1	Yes	100
AT5G47550.1	<a href="#">200804at33090</a>	Cysteine proteinase inhibitor	<a href="#">Q41916</a>	Cysteine proteinase inhibitor 5	Yes	100
AT4G16500.1	<a href="#">200804at33090</a>	Cysteine proteinase inhibitor	<a href="#">A0A178V035</a>	CYS4	Yes	100