RepetDB: A TE database

Nicolas Francillonne, Mariène Wan, Nathalie Choisne, Françoise Alfama, Raphaël Flores, Joelle Amselem, Johann Confais, Hadi Quesneville

RepetDB : a TE database

Nicolas FRANCILLONNE [1,2], Mariène WAN [1,2], Nathalie CHOISNE [1,2], Françoise ALFAMA [1,2], Raphaël FLORES [1,2], Joelle AMSELEM [1,2], Johann CONFAIS [1,2] and **Hadi QUESNEVILLE [1,2]**

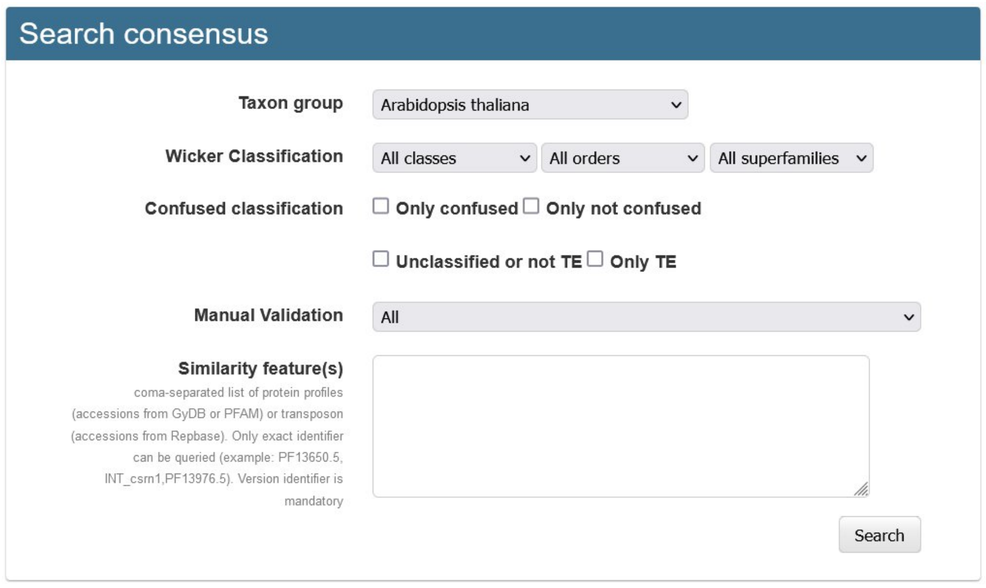
[1] Université Paris-Saclay, INRAE, URGI, 78026, Versailles, France

[2] Université Paris-Saclay, INRAE, BioinfOmics, Plant bioinformatics facility, 78026,Versailles, France

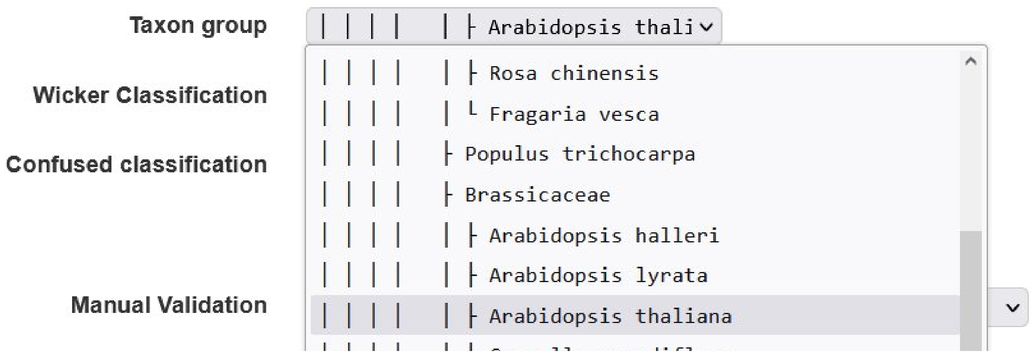
Transposable elements (TEs) are major players of structure and evolution of eukaryote genomes. Thanks to their ability to move around and to replicate within genomes, they are probably the most important contributors to genome plasticity. **Their detection and annotation are considered essential and must be undertaken in any genome sequencing project.**

To annotate transposable elements in genome, consensus are needed and RepetDB aims to make available these sequence to make easier the annotation of closely related genomes. You can select consensus of interest in the database and download fasta sequence.

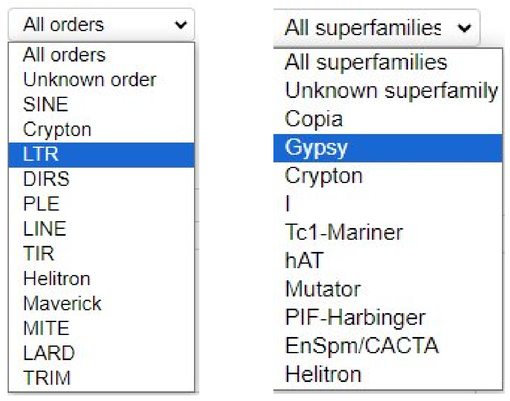
- I wish to make an annotation of a Brassicaceae species close from *Arabidopsis thaliana.* - I go to repetdb website and check the related species of interest https://urgi.versailles.inrae.fr/repetdb



- I can also select multiple species, from example all *Brassicaceae* or just one close related species like *Arabidopsis thaliana*



- Selection of specifics class, order or superfamily is also available to refine a search



Consensus search result

Back to home screen

Search criteria:

- **Taxon group**: Arabidopsis thaliana

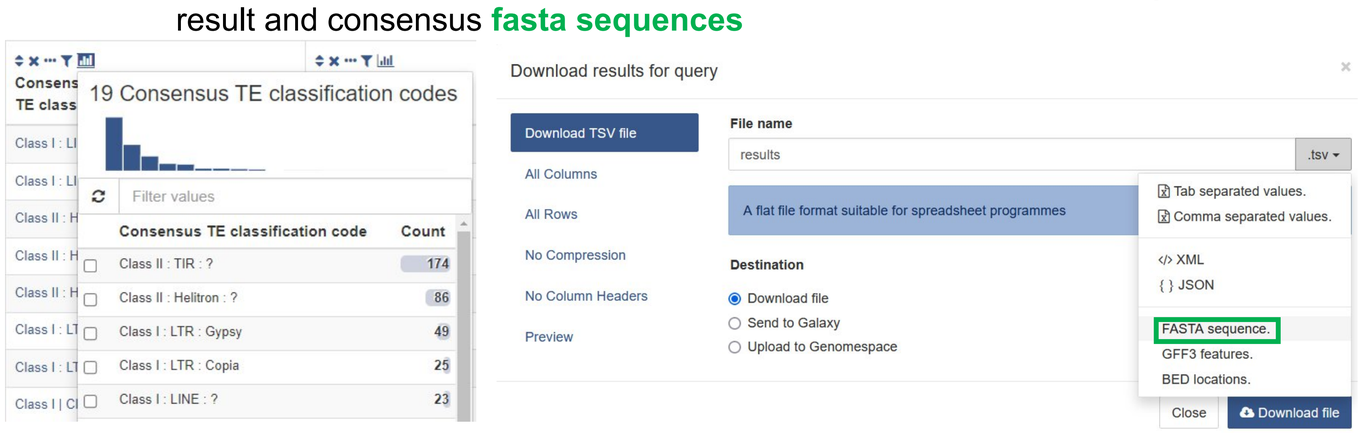
Manage Columns | Manage Filters | Manage Relationships | Save as List | Generate Python code | Export

Showing 1 to 25 of 406 rows

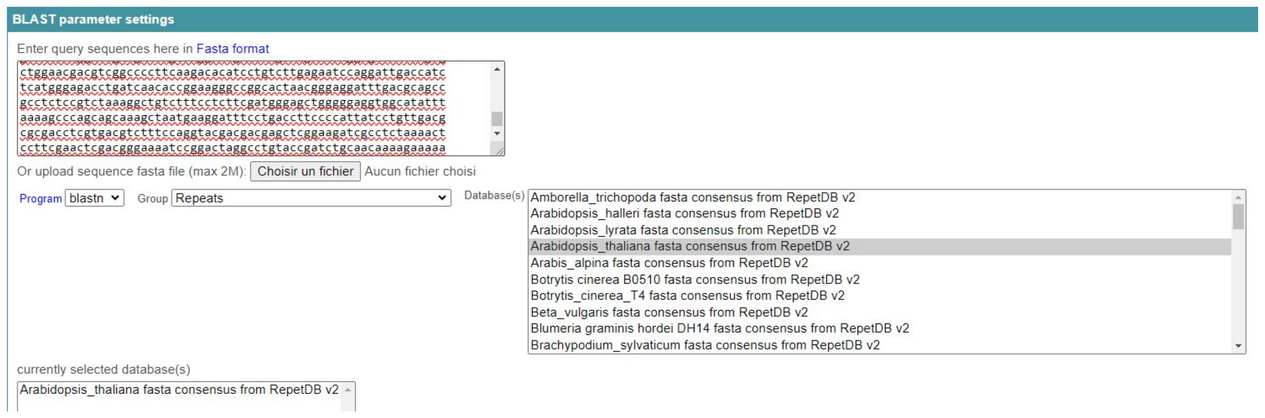
Rows per page: 25 | page 1

Consensus Consensus identifier	Consensus Copies	Consensus Full-length copies	Consensus Length	Consensus TE classification code	Consensus Manual validation	Consensus Confused	Consensus Comments
MCL102_AthaCol0_TEdenovo-B-G1208-Map3	20	6	471	Class I : LINE : ?	classification	no	NO VALUE
MCL102_AthaCol0_TEdenovo-B-R178-Map3	122	8	2017	Class I : LINE : ?	classification	no	NO VALUE
MCL103_AthaCol0_TEdenovo-B-G1214-Map20	23	11	449	Class II : Helitron : ?	classification	no	NO VALUE
MCL103_AthaCol0_TEdenovo-B-P161.100-Map16	21	11	452	Class II : Helitron : ?	classification	no	NO VALUE

- I can obtain statistics and filter some values out or download my result and consensus **fasta sequences**



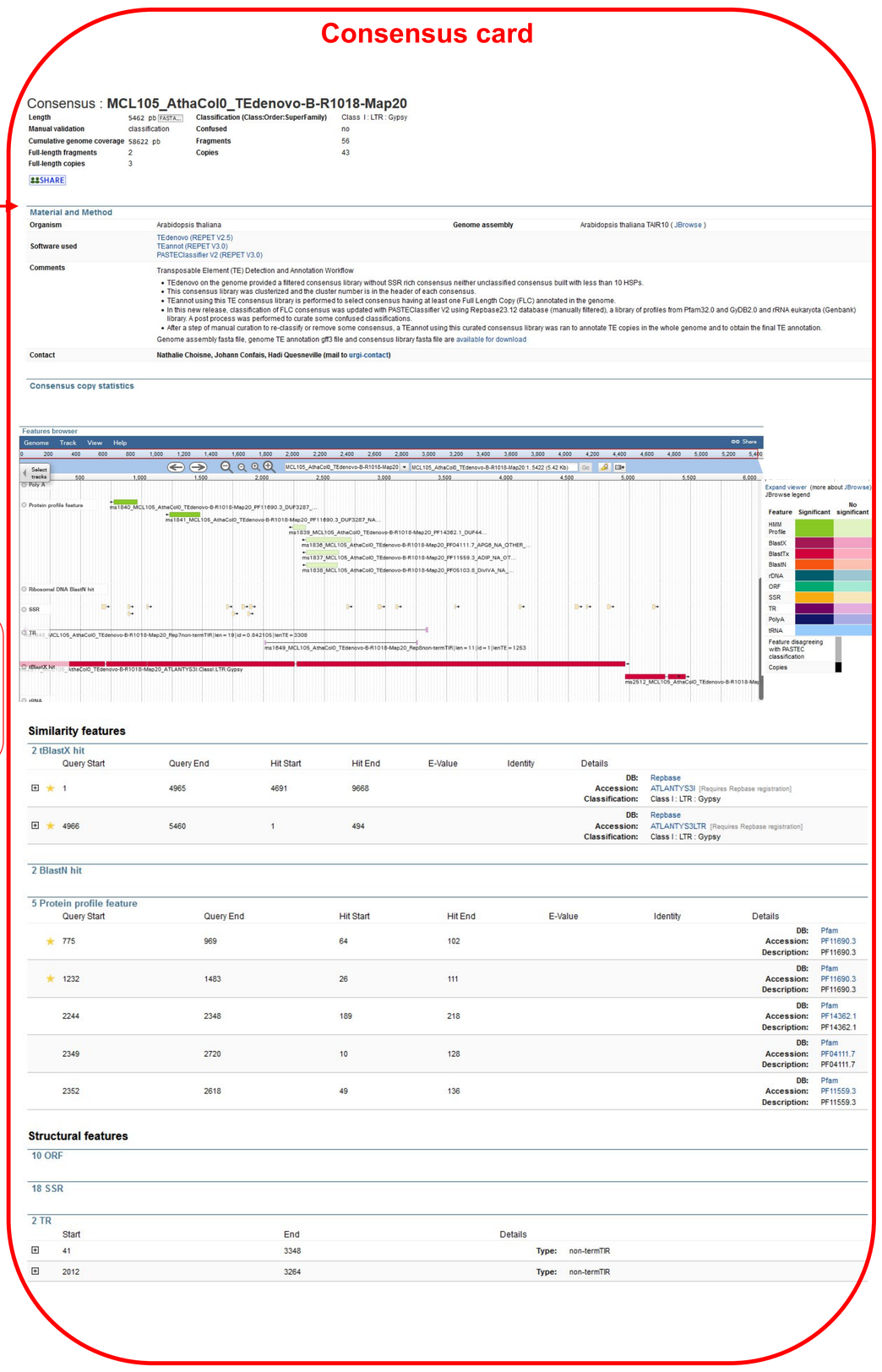
- I have an interesting sequence that I suspect is a transposable element and I wish to check if a match exist in the database. I go to PlantbioinfoPF blast services https://urgi.versailles.inrae.fr/blast/ and select in group « *Repeats* » a close related species



- If a match is found I get a direct link to the consensus card

By selecting one consensus I can access a consensus card describing its classification

Consensus card



Visit RepetDB : https://urgi.versailles.inrae.fr/repetdb/begin.do

Use URGI Blast : https://urgi.versailles.inrae.fr/blast/

Supplementary information and tutorials : https://urgi.versailles.inrae.fr/Data/Transposable-elements/REPETDB