



# GrTEdb: the first web-based database of transposable elements in cotton (*Gossypium raimondii*)

Although several diploid and tetraploid *Gossypium* species genomes have been sequenced, the well annotated web-based transposable elements (TEs) database is lacking. To better understand the roles of TEs in structural, functional and evolutionary dynamics of the cotton genome, a comprehensive, specific, and user-friendly web-based database, *Gossypium raimondii* transposable elements database (GrTEdb), was constructed.

Meanwhile, the web-based sequence browsing, searching, downloading and blast tool were implemented to help users easily and effectively to annotate the TEs or TE fragments in genomic sequences from *G. raimondii* and other closely related *Gossypium* species.

GrTEdb provides resources and information related with TEs in *G. raimondii*, and will facilitate gene and genome analyses within or across *Gossypium* species, evaluating the impact of TEs on their host genomes, and investigating the potential interaction between TEs and protein-coding genes in *Gossypium* species.

**Zhenzhen Xu**

**Jiangsu Academy of Agricultural Sciences, China**

# Results

## 1. Construction and content of the database.

14 332 TEs were structurally annotated and clearly categorized in the *G. raimondii* genome, including 2929 *Copia*-like elements, 10 368 *Gypsy*-like elements, 299 *L1*, 12 *Mutators*, 435 *PIF-Harbingers*, 275 *CACTAs* and 14 *Helitrons* (Table 1).

## 2. User interface.

The GrTEdb database organization is navigated by six sections: Home, Browse, Search and Download, Blast, Links and Contact (Figure 1A).

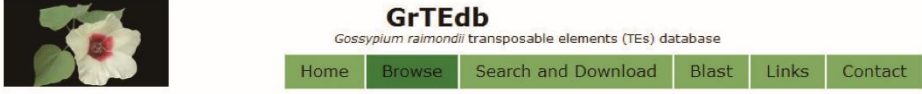
## 3. Browse.

In the browsing interface, the classification structures of TEs deposited in GrTEdb were showed. Users can download the whole TEs sequences, and can browse any one superfamily of interest by the hyperlinks provided. The detailed information of each superfamily can be retrieved and downloaded by clicking the corresponding entry (Figure 1B).

Table 1. Summary of the identified TEs in *G. raimondii*

Class	Order	Superfamily	Copy numbers
Retrotransposons	LTR	<i>Copia</i>	2929
		<i>Gypsy</i>	10 368
		<i>L1</i>	299
DNA transposons	TIR	<i>Mutator</i>	12
		<i>PIF-Harbinger</i>	435
		<i>CACTA</i>	275
		<i>Helitron</i>	14
Total			14 332

**A**



**B**

**Browse**

Click the hyperlink to view or download each superfamily and its members. The statistical information of TEs belong to *Copia*.

Download all sequences in GrTEdb ↓

GrTEdb Summary

Class	Order	Superfamily	Copy numbers
Retrotransposons	LTR	<i>Copia</i>	2929
		<i>Gypsy</i>	10368
		<i>L1</i>	299
DNA transposons	TIR	<i>Mutator</i>	12
		<i>PIF-Harbinger</i>	435
		<i>CACTA</i>	275
		<i>Helitron</i>	14
Total			14332

Download all sequences belonged to *Copia* ↓

A total of 199 records

Superfamily	Family	Number	View	Download
<i>Copia</i>	RLC_1	21	View	Download ↓
	RLC_10	2	View	Download ↓
	RLC_100	4	View	Download ↓
	RLC_101	1	View	Download ↓
	RLC_102	2	View	Download ↓
	RLC_103	2	View	Download ↓
	RLC_104	1	View	Download ↓
	RLC_105	20	View	Download ↓
	RLC_106	2	View	Download ↓
	RLC_107	1	View	Download ↓
RLC_108	1	View	Download ↓	

Figure 1. (A) The top menu of GrTEdb. (B) The user interface of browsing in GrTEdb.

# Results

## 4. The searching interface of GrTEdb.

In the searching and downloading interface, users can use a keyword to search the GrTEdb (e.g. TE ID, Class, Order, Superfamily and Family) to locate specific TEs quickly. The search results can be viewed and downloaded by clicking the hyperlinks provided on the page (Figure 2).

Search TEs by ID **Please enter sequence ID:**  
 Search TE sequences by inputting single or multiple sequence ID (one per line; examples: #RLC\_1\_1\_Gr #RLC\_1\_10\_Gr #RLC\_1\_11\_Gr)

RLC\_1\_1\_Gr  
RLC\_1\_10\_Gr

SUBMIT RESET

A total of 2 records

ID	Class	Order	Superfamily	Family	Description	Length	Chromosome	Start	End	Strand	View	Download
RLC_1_1_Gr	Retrotransposons	LTR	Copia	RLC_1	INTACT-LTR	6216	6	5377540	5383755	-	View	Download
RLC_1_10_Gr	Retrotransposons	LTR	Copia	RLC_1	Solo LTR	484	7	53742478	53742561	-	View	Download
Total												Download all

Search TE sequences by different categories

Search TEs by family

Enter a key name: RLC\_2 searched in Family

Search TE sequences by inputting a class, order, superfamily or family name (e.g (Class) Retrotransposons)

SUBMIT CLEAR

A total of 9 records

ID	Class	Order	Superfamily	Family	Description	Length	Chromosome	Start	End	Strand	View	Download
RLC_2_1_Gr	Retrotransposons	LTR	Copia	RLC_2	INTACT-LTR	4957	8	31678581	31710844	+	View	Download
RLC_2_2_Gr	Retrotransposons	LTR	Copia	RLC_2	INTACT-LTR	5577	6	31197956	31203531	+	View	Download
RLC_2_3_Gr	Retrotransposons	LTR	Copia	RLC_2	INTACT-LTR	4051	2	39534915	39538665	+	View	Download
RLC_2_4_Gr	Retrotransposons	LTR	Copia	RLC_2	INTACT-LTR	4436	6	38278210	38282675	+	View	Download
RLC_2_5_Gr	Retrotransposons	LTR	Copia	RLC_2	INTACT-LTR	4994	2	55070789	55075782	-	View	Download
RLC_2_6_Gr	Retrotransposons	LTR	Copia	RLC_2	INTACT-LTR	4244	2	58618134	58622377	-	View	Download
RLC_2_7_Gr	Retrotransposons	LTR	Copia	RLC_2	Solo-LTR	219	6	44595478	44595696	+	View	Download
RLC_2_8_Gr	Retrotransposons	LTR	Copia	RLC_2	INTACT-LTR	4392	2	55287558	55291649	+	View	Download
RLC_2_9_Gr	Retrotransposons	LTR	Copia	RLC_2	INTACT-LTR	4958	13	8894925	8895922	+	View	Download
Total												Download all

Figure 2. The searching interface of GrTEdb.

## 5. The chromosomal region search page.

In the chromosomal region search page, users can retrieve the TEs for any one entire chromosome or in a defined window around either a chromosomal position or a gene model, and the detailed information of each retrieved TEs can be viewed and downloaded by clicking the hyperlinks provided on the page (Figure 3).

**A**

Search TE sequences by a chromosomal position

1. Choose a chromosome from the *Gossypium raimondii* genomic sequences  
01

2. Enter a chromosomal position 654787 bp

3. Select a search window size around your position to search TEs.  
100 kb

SUBMIT QUERY RESET QUERY

A total of 2 records

ID	Class	Order	Superfamily	Family	Description	Length	Chromosome	Start	End	Strand	View	Download
RLG_111_34_Gr	Retrotransposons	LTR	Gypsy	RLG_111	INTACT-LTR	2369	1	682560	684928	+	View	Download
RLG_74_963_Gr	Retrotransposons	LTR	Gypsy	RLG_74	Solo-LTR	2366	1	604360	606731	-	View	Download
Total												Download all

**B**

Search TE sequences by a gene name

1. Enter a cotton gene name from *Gossypium raimondii* genomic sequences (e.g. Goral.001G001100) Goral.001G001100

2. Select a search window size around your gene to search TEs.  
100 kb

SUBMIT QUERY RESET QUERY

A total of 5 records

ID	Class	Order	Superfamily	Family	Description	Length	Chromosome	Start	End	Strand	View	Download
RLC_148_38_Gr	Retrotransposons	LTR	Copia	RLC_148	INTACT-LTR	3991	1	64254	71491	+	View	Download
RLC_188_12_Gr	Retrotransposons	LTR	Copia	RLC_188	INTACT-LTR	4996	1	46892	51887	+	View	Download
RLG_134_365_Gr	Retrotransposons	LTR	Gypsy	RLG_134	Solo-LTR	3101	1	34778	10976	-	View	Download
RLG_136_251_Gr	Retrotransposons	LTR	Gypsy	RLG_136	Solo-LTR	3093	1	37696	49762	+	View	Download
RLG_71_194_Gr	Retrotransposons	LTR	Gypsy	RLG_71	INTACT-LTR	7842	1	100996	108837	+	View	Download
Total												Download all

Figure 3. The chromosomal region search page.

# Results

## 6. Blast.

In the BLAST search page, users can handy and quickly compare their sequences with the cotton TEs deposited in GrTEdb (Figure 4).

# Conclusions

We have generated a web-based GrTEdb, and it provides researchers with not only resources and information related to different TEs in the cotton genome but also tools for performing data analysis.



Figure 4. The BLAST interface (left) and a sample of BLASTn results (right) provided in GrTEdb.



Database tool

## GrTEdb: the first web-based database of transposable elements in cotton (*Gossypium raimondii*)

Zhenzhen Xu<sup>1,†</sup>, Jing Liu<sup>2,†</sup>, Wanchao Ni<sup>1</sup>, Zhen Peng<sup>2</sup>, Yue Guo<sup>2</sup>, Wuwei Ye<sup>3</sup>, Fang Huang<sup>1</sup>, Xianggui Zhang<sup>1</sup>, Peng Xu<sup>1</sup>, Qi Guo<sup>1</sup>, Xinlian Shen<sup>1,\*</sup> and Jianchang Du<sup>2,\*</sup>

Database, 2017, 1–7  
doi: 10.1093/database/bax013  
Database tool

