Practical Session: Rapid Genotyping of *B. pertussis* Using BIGSdb

Valérie Bouchez & Martin Rethoret-Pasty

GenEpi Webinar :

Focus on the Agents of Whooping Cough
A virtual training workshop within GenEpi-BioTrain.

November 4th – 5th 2024

Intended Learning Objectives

https://bigsdb.pasteur.fr/bordetella/

Demos on the BIGSdb Bordetella base (Valérie Bouchez)

- **Exploration of BIGSdb Bordetella Allele & profiles databases**
- **Overview of BIGSdb Bordetella Isolates database, genotyping schemes and public projects**
- **Analysis on BIGSdb Bordetella Isolates database (genome comparator, GrapeTree…)**

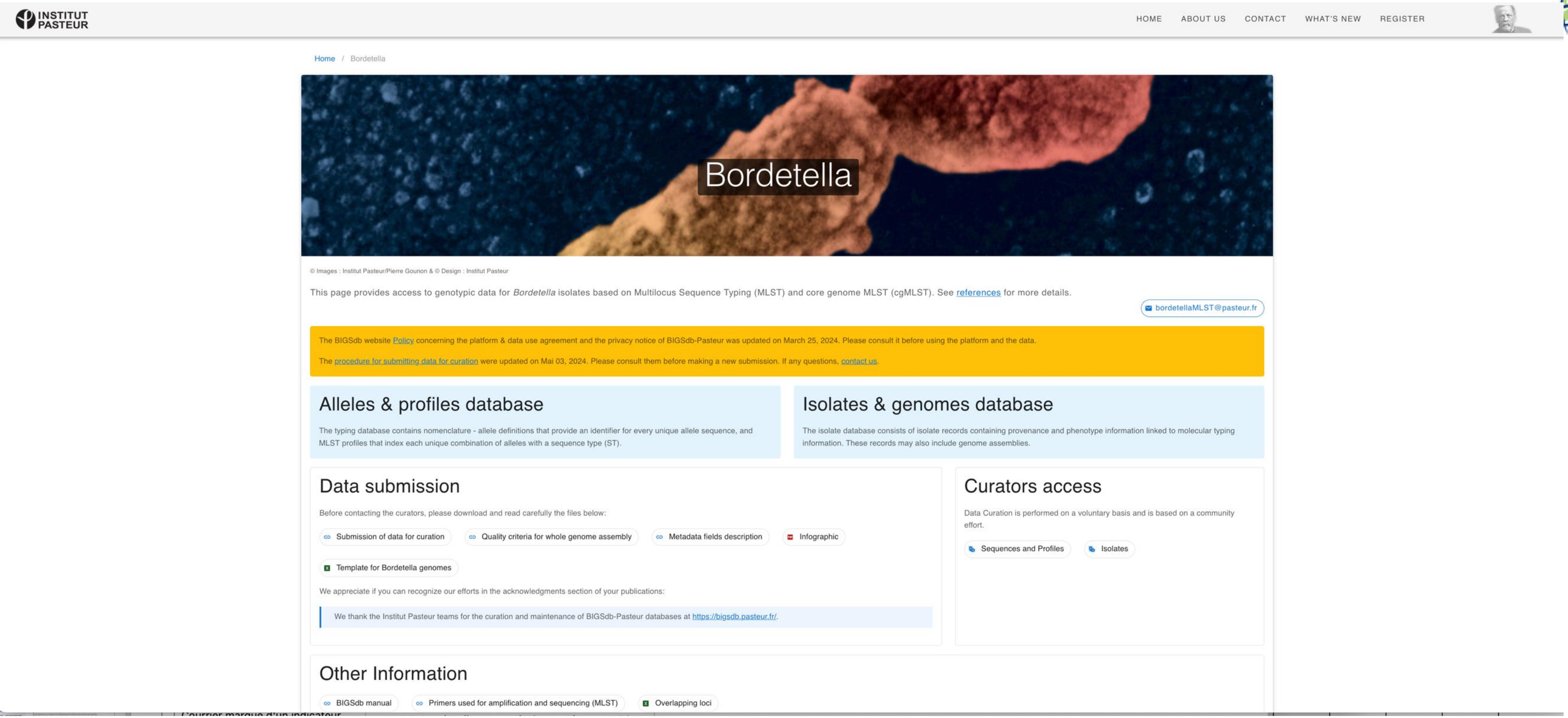
Presentation of the exercises to be carried out on public projects

Session of November 5th, 2024 (Valérie Bouchez & Martin Rethoret-Pasty)
Correction of the practical exercises

Exploration of BIGSdb Bordetella Allele and profiles databases

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



© Images : Institut Pasteur/Pierre Gounon & © Design : Institut Pasteur

This page provides access to genotypic data for *Bordetella* isolates based on Multilocus Sequence Typing (MLST) and core genome MLST (cgMLST). See references for more details.

The BIGSdb website Policy concerning the platform & data use agreement and the privacy notice of BIGSdb-Pasteur was updated on March 25, 2024. Please consult it before using the platform and the data.

The procedure for submitting data for curation were updated on Mai 03, 2024. Please consult them before making a new submission. If any questions, contact us.

Alleles & profiles database

The typing database contains nomenclature - allele definitions that provide an identifier for every unique allele sequence, and MLST profiles that index each unique combination of alleles with a sequence type (ST).

Isolates & genomes database

The isolate database consists of isolate records containing provenance and phenotype information linked to molecular typing information. These records may also include genome assemblies.

Data submission

Before contacting the curators, please download and read carefully the files below:

We appreciate if you can recognize our efforts in the acknowledgments section of your publications:

> We thank the Institut Pasteur teams for the curation and maintenance of BIGSdb-Pasteur databases at https://bigsdb.pasteur.fr/.

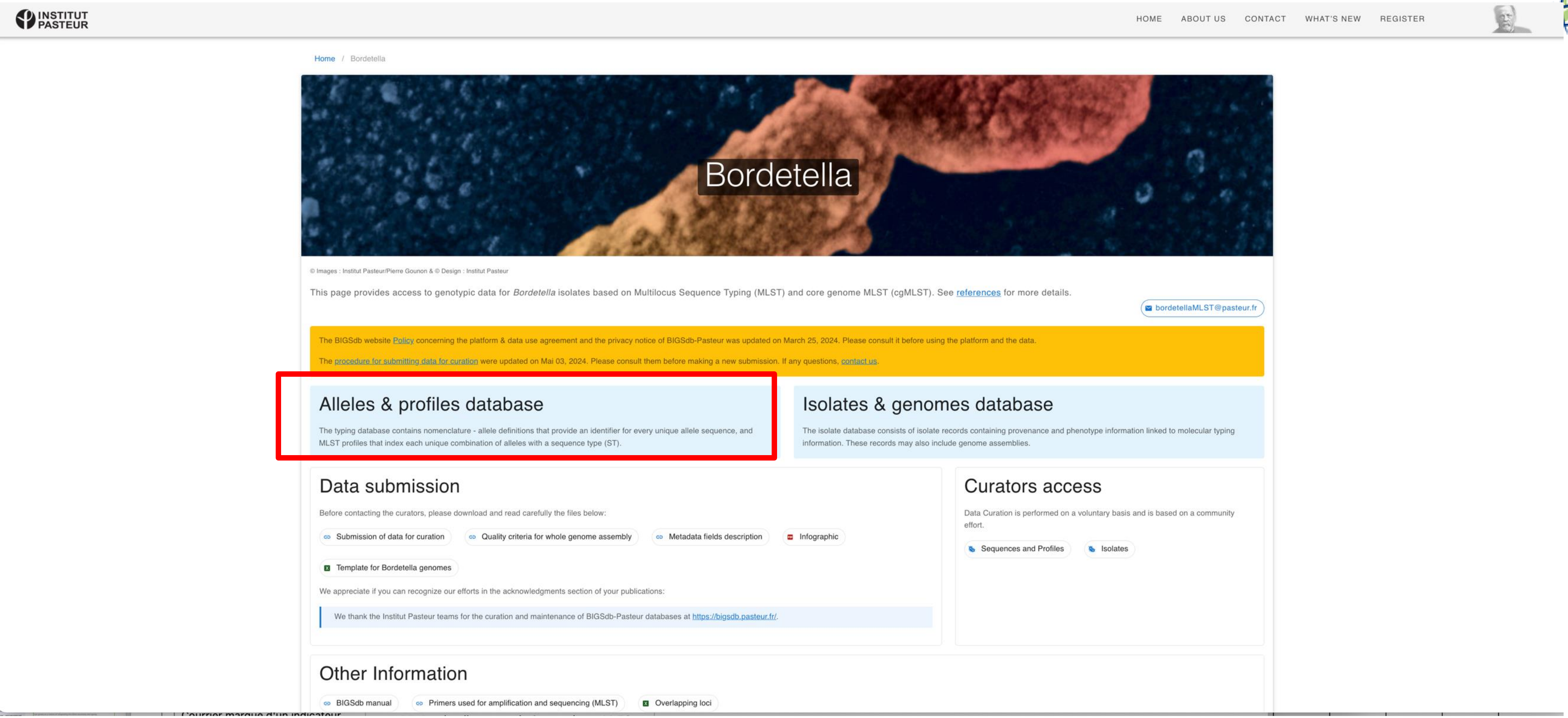
Curators access

Data Curation is performed on a voluntary basis and is based on a community effort.

Other Information

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



© Images : Institut Pasteur/Pierre Gounon & © Design : Institut Pasteur

This page provides access to genotypic data for *Bordetella* isolates based on Multilocus Sequence Typing (MLST) and core genome MLST (cgMLST). See references for more details.

bordetellaMLST@pasteur.fr

The BIGSdb website Policy concerning the platform & data use agreement and the privacy notice of BIGSdb-Pasteur was updated on March 25, 2024. Please consult it before using the platform and the data.

The procedure for submitting data for curation were updated on Mai 03, 2024. Please consult them before making a new submission. If any questions, contact us.

Alleles & profiles database

The typing database contains nomenclature - allele definitions that provide an identifier for every unique allele sequence, and MLST profiles that index each unique combination of alleles with a sequence type (ST).

Isolates & genomes database

The isolate database consists of isolate records containing provenance and phenotype information linked to molecular typing information. These records may also include genome assemblies.

Data submission

Before contacting the curators, please download and read carefully the files below:

We appreciate if you can recognize our efforts in the acknowledgments section of your publications:

> We thank the Institut Pasteur teams for the curation and maintenance of BIGSdb-Pasteur databases at https://bigsdb.pasteur.fr/.

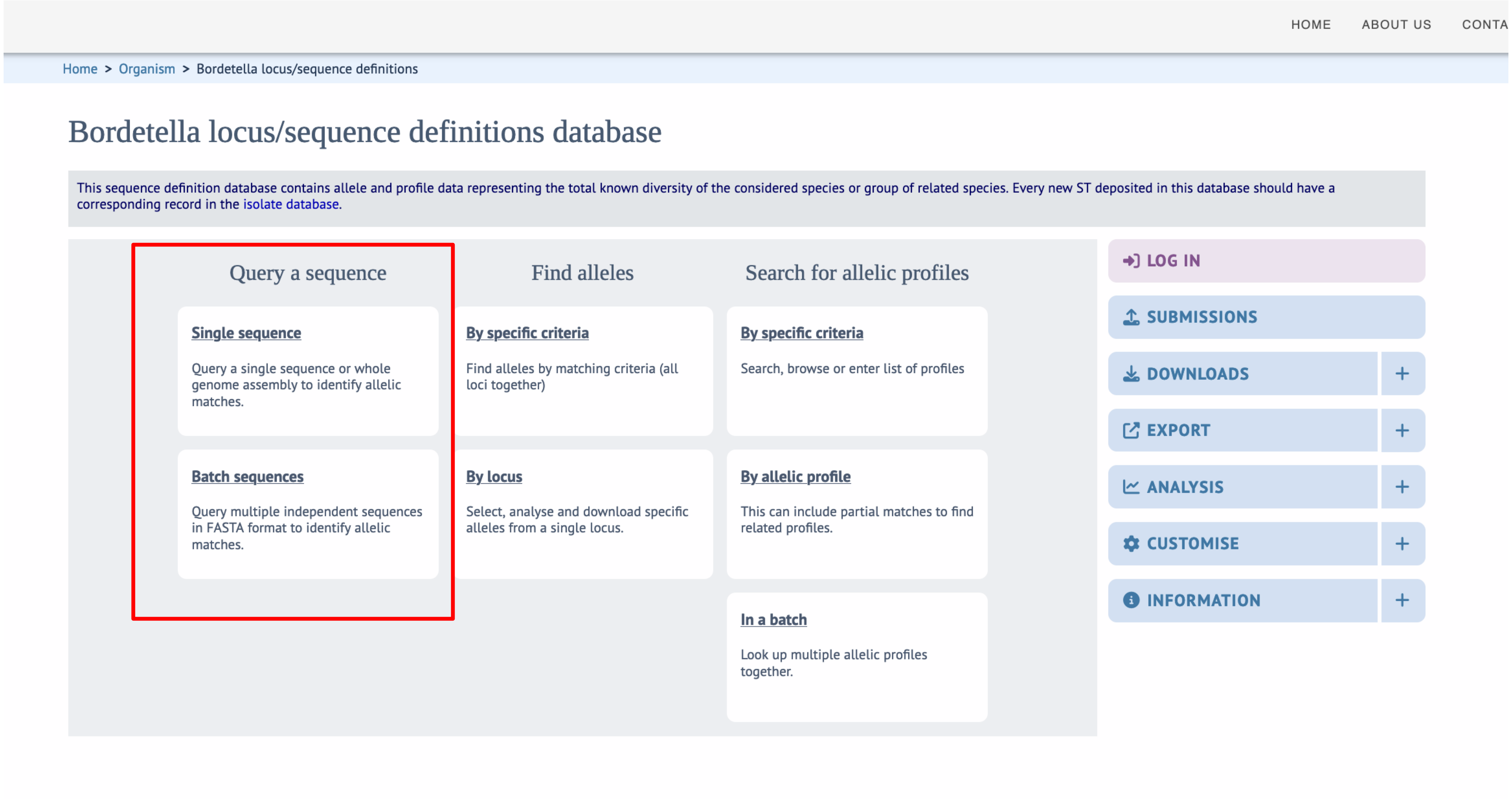
Curators access

Data Curation is performed on a voluntary basis and is based on a community effort.

Other Information

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



Bordetella locus/sequence definitions database

This sequence definition database contains allele and profile data representing the total known diversity of the considered species or group of related species. Every new ST deposited in this database should have a corresponding record in the isolate database.

Query a sequence

Single sequence

Query a single sequence or whole genome assembly to identify allelic matches.

Batch sequences

Query multiple independent sequences in FASTA format to identify allelic matches.

Find alleles

By specific criteria

Find alleles by matching criteria (all loci together)

By locus

Select, analyse and download specific alleles from a single locus.

Search for allelic profiles

By specific criteria

Search, browse or enter list of profiles

By allelic profile

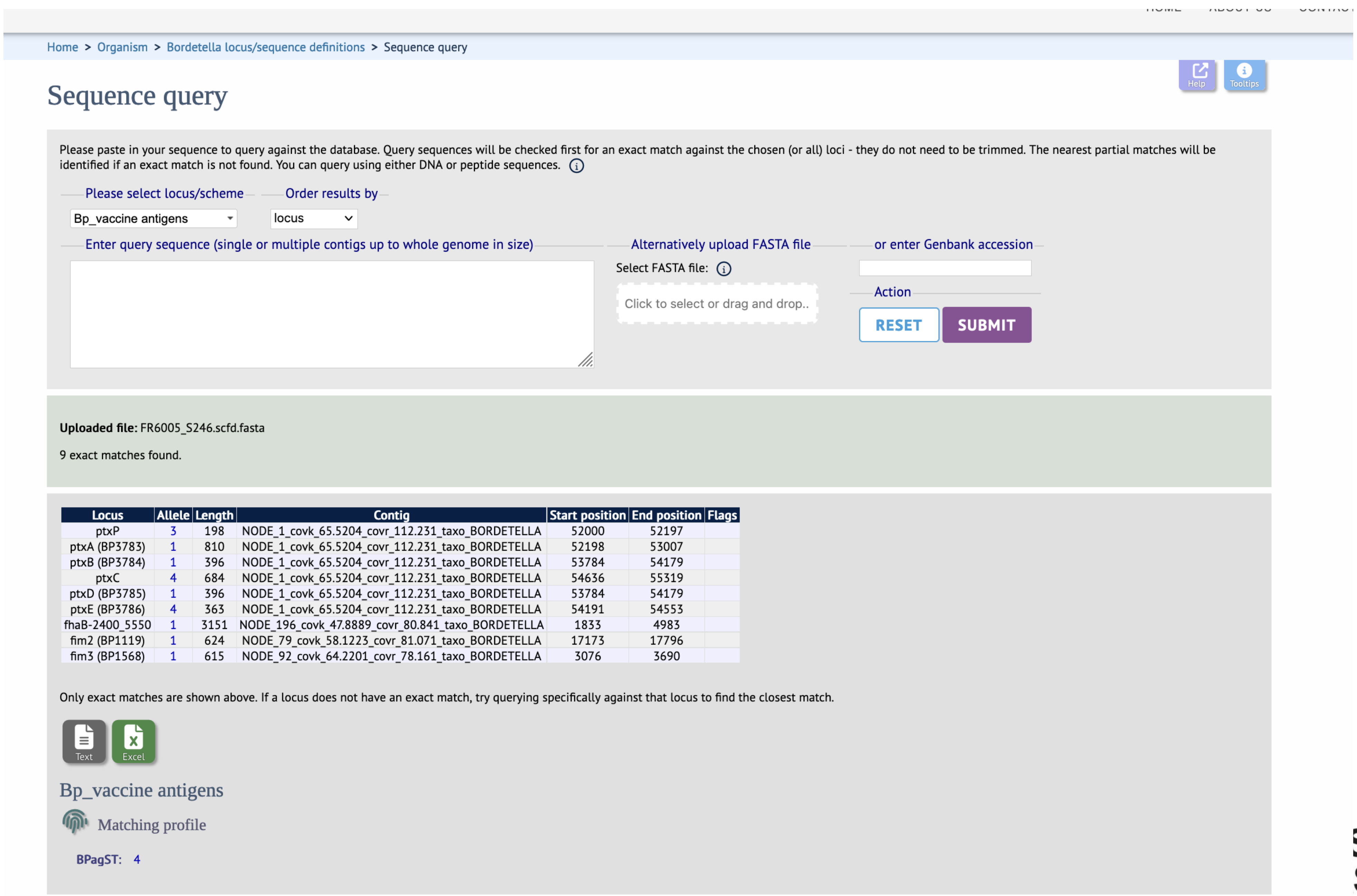
This can include partial matches to find related profiles.

In a batch

Look up multiple allelic profiles together.

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



Sequence query

Please paste in your sequence to query against the database. Query sequences will be checked first for an exact match against the chosen (or all) loci - they do not need to be trimmed. The nearest partial matches will be identified if an exact match is not found. You can query using either DNA or peptide sequences.

Please select locus/scheme: Bp_vaccine antigens

Order results by: locus

Enter query sequence (single or multiple contigs up to whole genome in size)

Alternatively upload FASTA file

Select FASTA file: Click to select or drag and drop..

or enter Genbank accession

Uploaded file: FR6005_S246.scfd.fasta

9 exact matches found.

Locus	Allele	Length	Contig	Start position	End position	Flags
ptxP	3	198	NODE_1_covk_65.5204_covr_112.231_taxo_BORDETELLA	52000	52197	
ptxA (BP3783)	1	810	NODE_1_covk_65.5204_covr_112.231_taxo_BORDETELLA	52198	53007	
ptxB (BP3784)	1	396	NODE_1_covk_65.5204_covr_112.231_taxo_BORDETELLA	53784	54179	
ptxC	4	684	NODE_1_covk_65.5204_covr_112.231_taxo_BORDETELLA	54636	55319	
ptxD (BP3785)	1	396	NODE_1_covk_65.5204_covr_112.231_taxo_BORDETELLA	53784	54179	
ptxE (BP3786)	4	363	NODE_1_covk_65.5204_covr_112.231_taxo_BORDETELLA	54191	54553	
fhaB-2400_5550	1	3151	NODE_196_covk_47.8889_covr_80.841_taxo_BORDETELLA	1833	4983	
fim2 (BP1119)	1	624	NODE_79_covk_58.1223_covr_81.071_taxo_BORDETELLA	17173	17796	
fim3 (BP1568)	1	615	NODE_92_covk_64.2201_covr_78.161_taxo_BORDETELLA	3076	3690	

Only exact matches are shown above. If a locus does not have an exact match, try querying specifically against that locus to find the closest match.

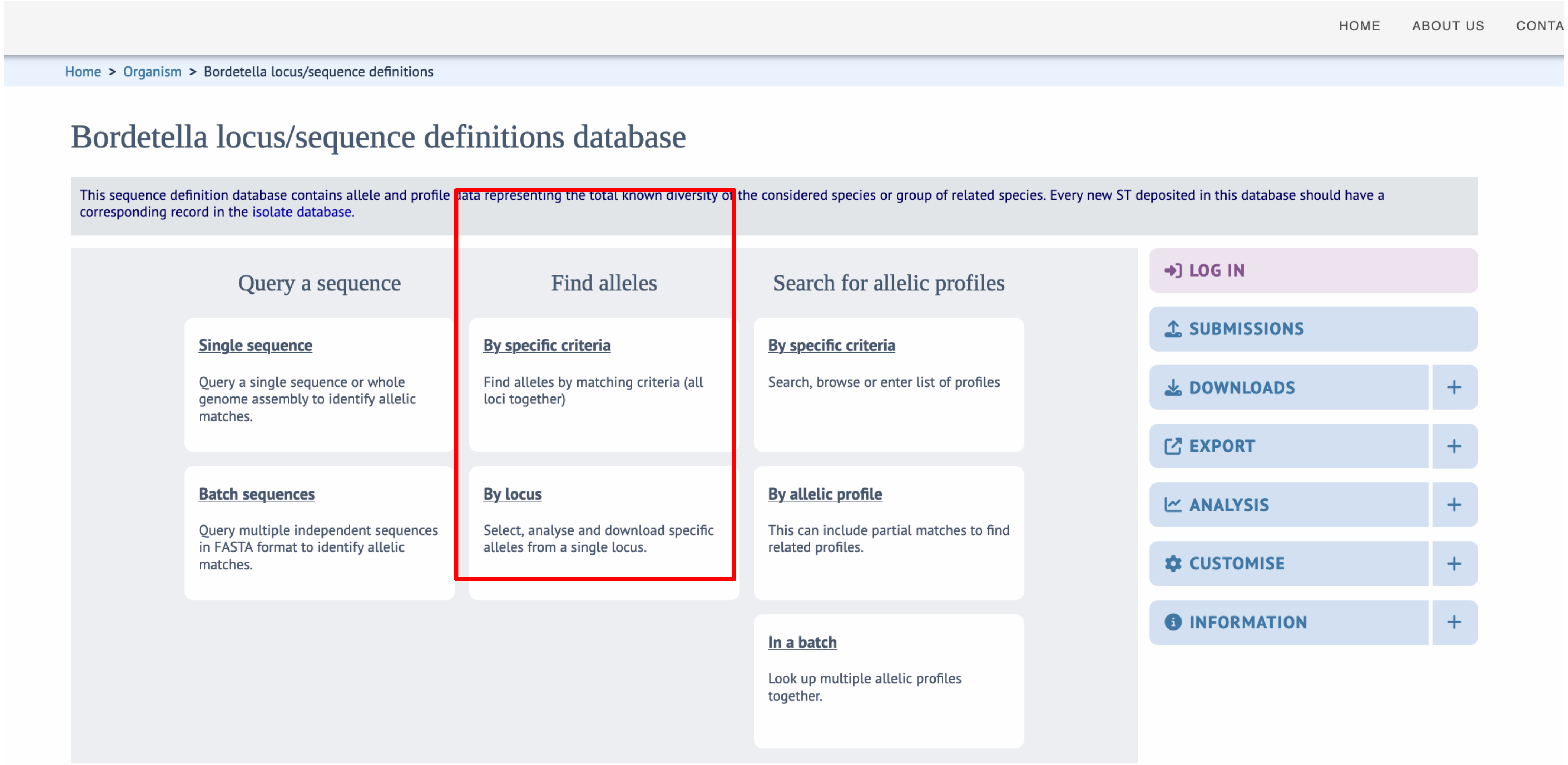
Bp_vaccine antigens

Matching profile

BPagST: 4

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



Bordetella locus/sequence definitions database

This sequence definition database contains allele and profile data representing the total known diversity of the considered species or group of related species. Every new ST deposited in this database should have a corresponding record in the isolate database.

Query a sequence

Single sequence

Query a single sequence or whole genome assembly to identify allelic matches.

Batch sequences

Query multiple independent sequences in FASTA format to identify allelic matches.

Find alleles

By specific criteria

Find alleles by matching criteria (all loci together)

By locus

Select, analyse and download specific alleles from a single locus.

Search for allelic profiles

By specific criteria

Search, browse or enter list of profiles

By allelic profile

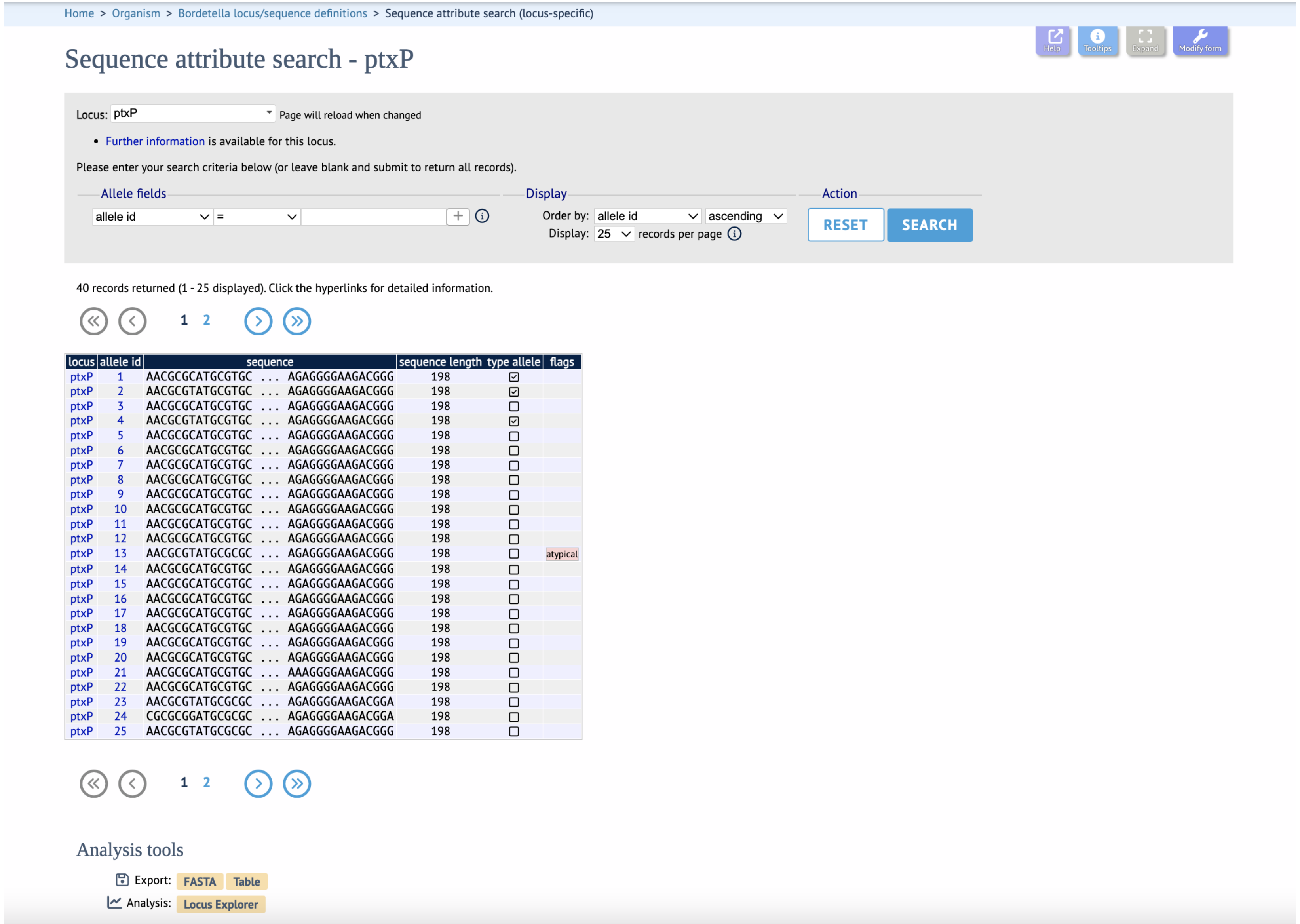
This can include partial matches to find related profiles.

In a batch

Look up multiple allelic profiles together.

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



Home > Organism > Bordetella locus/sequence definitions > Sequence attribute search (locus-specific)

Sequence attribute search - ptxP

Locus: ptxP Page will reload when changed

- Further information is available for this locus.

Please enter your search criteria below (or leave blank and submit to return all records).

Allele fields: allele id =

Display: Order by: allele id ascending

Display: 25 records per page

Action

40 records returned (1 - 25 displayed). Click the hyperlinks for detailed information.

1 2

locus	allele id	sequence	sequence length	type allele	flags
ptxP	1	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☑	
ptxP	2	AACGCGTATGCGTGC ... AGAGGGGAAGACGGG	198	☑	
ptxP	3	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	4	AACGCGTATGCGTGC ... AGAGGGGAAGACGGG	198	☑	
ptxP	5	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	6	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	7	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	8	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	9	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	10	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	11	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	12	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	13	AACGCGTATGCGCGC ... AGAGGGGAAGACGGG	198	☐	atypical
ptxP	14	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	15	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	16	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	17	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	18	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	19	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	20	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	21	AACGCGCATGCGTGC ... AAAGGGGAAGACGGG	198	☐	
ptxP	22	AACGCGCATGCGTGC ... AGAGGGGAAGACGGG	198	☐	
ptxP	23	AACGCGTATGCGCGC ... AGAGGGGAAGACGGA	198	☐	
ptxP	24	CGCGCGGATGCGCGC ... AGAGGGGAAGACGGA	198	☐	
ptxP	25	AACGCGTATGCGCGC ... AGAGGGGAAGACGGG	198	☐	

1 2

Analysis tools

Export: FASTA Table

Analysis: Locus Explorer

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases

Locus Explorer

Polymorphic site analysis

ptxP

The colour codes represent the percentage of alleles that have a particular nucleotide at each position. Click anywhere within the sequence to drill down to allele and profile information. The width of the display can be altered by going to the options page - change this if the display goes off the page.

2 alleles included in analysis. 1 polymorphic site found.

Key: 0 - 10 | >10 - 20 | >20 - 30 | >30 - 40 | >40 - 50 | >50 - 60 | >60 - 70 | >70 - 80 | >80 - 90 | >90 - 100

```
  1        10        20        30        40        50        60        70        80        90        100
  AACGCGCATGCGTGCAGATTCGTCGTACAAAACCCTCGATTCTTCCGTACATCCCGCTACTGCAATCCAACACGGCATGAACGCTCCTTCGGCGCAAAGT
101        110       120       130       140       150       160       170       180       190
  CGCGCGATAGTACCGGTCACCGTCCGGACCGTGCTGACCCCCCTGCCATGGTGTGATCCGTAAAATAGGCACCATCAAAACGCAGAGGGGAAGACGGG
          G
```

Nucleotide frequencies

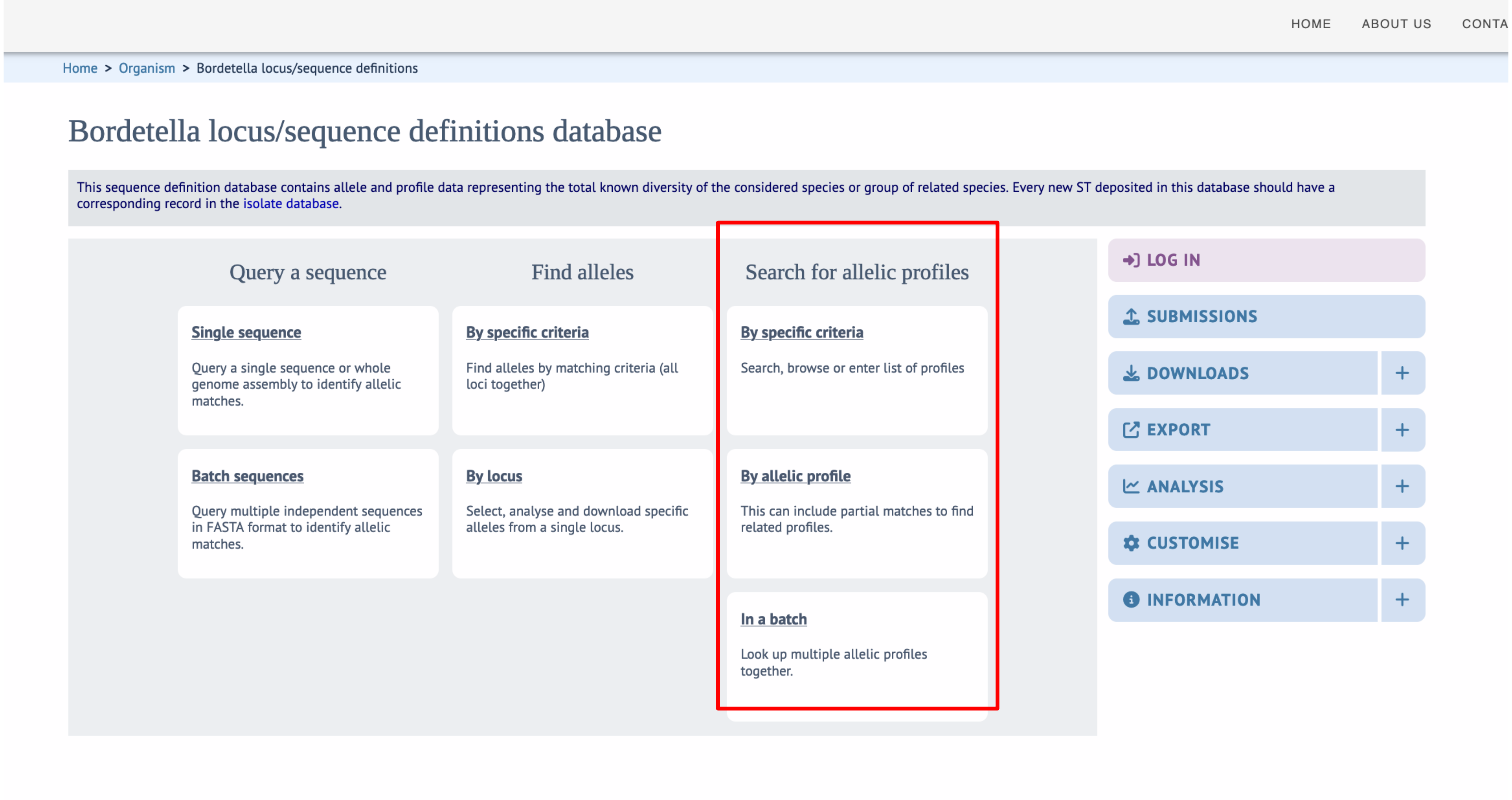
Position	Nucleotide									
	A	C	G	T	-	%A	%C	%G	%T	%-
109	1	0	1	0	0	50.00		50.00		

Download: Tab-delimited text format

22/09/2024

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



Bordetella locus/sequence definitions database

This sequence definition database contains allele and profile data representing the total known diversity of the considered species or group of related species. Every new ST deposited in this database should have a corresponding record in the isolate database.

Query a sequence

Single sequence

Query a single sequence or whole genome assembly to identify allelic matches.

Batch sequences

Query multiple independent sequences in FASTA format to identify allelic matches.

Find alleles

By specific criteria

Find alleles by matching criteria (all loci together)

By locus

Select, analyse and download specific alleles from a single locus.

Search for allelic profiles

By specific criteria

Search, browse or enter list of profiles

By allelic profile

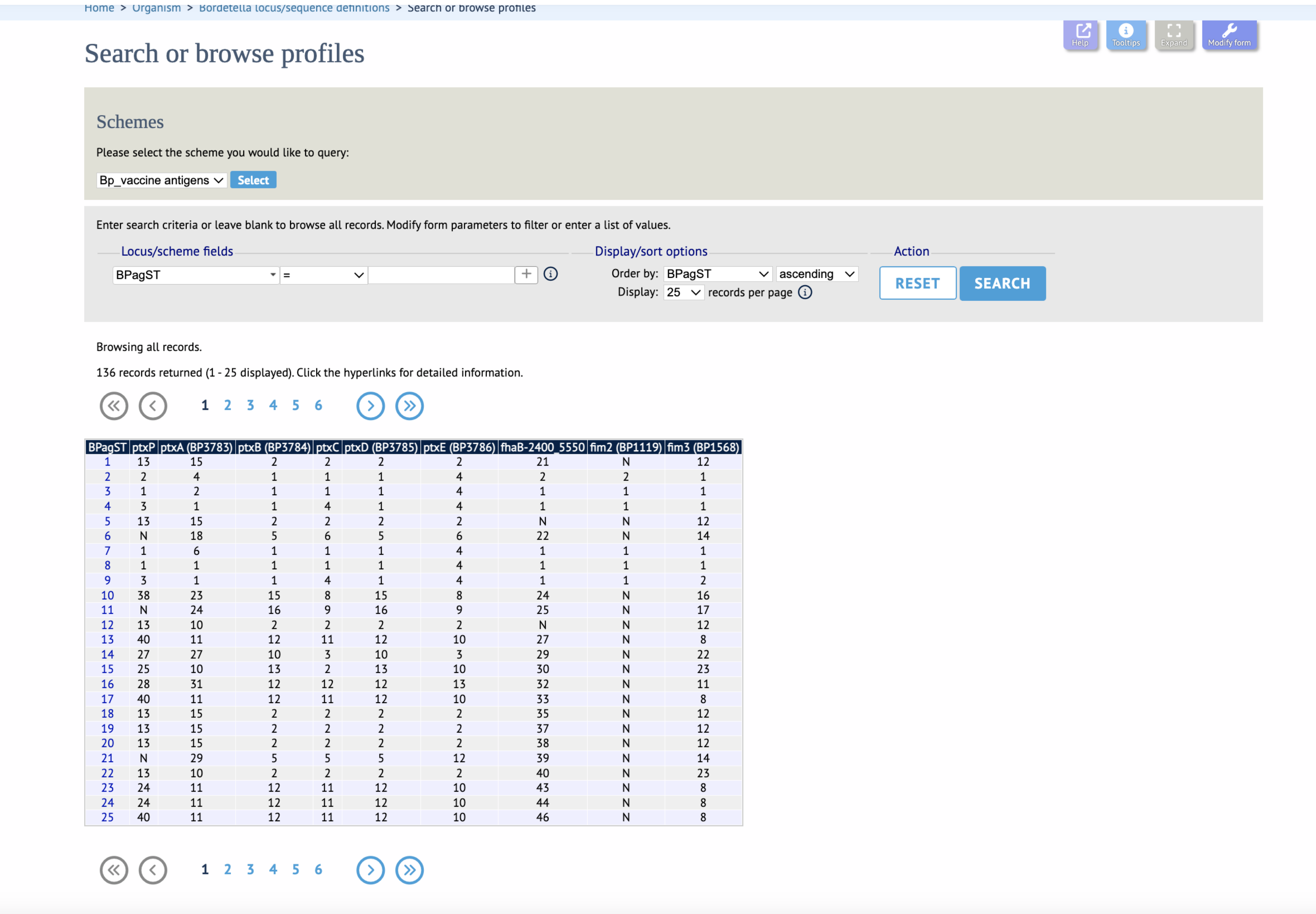
This can include partial matches to find related profiles.

In a batch

Look up multiple allelic profiles together.

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



Search or browse profiles

Schemes

Please select the scheme you would like to query: Bp_vaccine antigens

Enter search criteria or leave blank to browse all records. Modify form parameters to filter or enter a list of values.

BPagST =

Order by: BPagST ascending

Display: 25 records per page

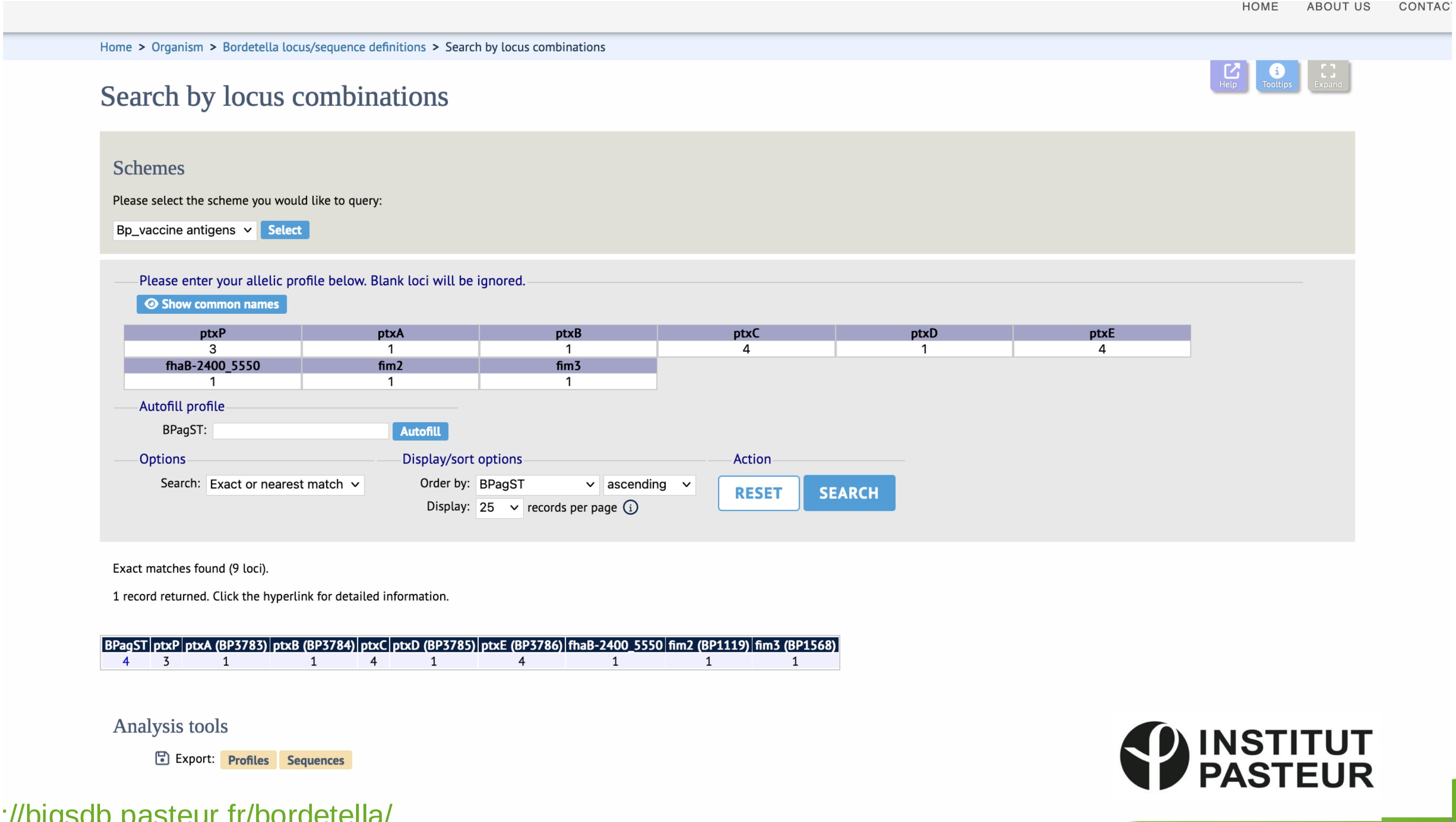
Browsing all records.

136 records returned (1 - 25 displayed). Click the hyperlinks for detailed information.

BPagST	ptxP	ptxA (BP3783)	ptxB (BP3784)	ptxC	ptxD (BP3785)	ptxE (BP3786)	fhaB-2400_5550	fim2 (BP1119)	fim3 (BP1568)
1	13	15	2	2	2	2	21	N	12
2	2	4	1	1	1	4	2	2	1
3	1	2	1	1	1	4	1	1	1
4	3	1	1	4	1	4	1	1	1
5	13	15	2	2	2	2	N	N	12
6	N	18	5	6	5	6	22	N	14
7	1	6	1	1	1	4	1	1	1
8	1	1	1	1	1	4	1	1	1
9	3	1	1	4	1	4	1	1	2
10	38	23	15	8	15	8	24	N	16
11	N	24	16	9	16	9	25	N	17
12	13	10	2	2	2	2	N	N	12
13	40	11	12	11	12	10	27	N	8
14	27	27	10	3	10	3	29	N	22
15	25	10	13	2	13	10	30	N	23
16	28	31	12	12	12	13	32	N	11
17	40	11	12	11	12	10	33	N	8
18	13	15	2	2	2	2	35	N	12
19	13	15	2	2	2	2	37	N	12
20	13	15	2	2	2	2	38	N	12
21	N	29	5	5	5	12	39	N	14
22	13	10	2	2	2	2	40	N	23
23	24	11	12	11	12	10	43	N	8
24	24	11	12	11	12	10	44	N	8
25	40	11	12	11	12	10	46	N	8

https://bigsdb.pasteur.fr/bordetella/

Short Practical Demos on BIGSdb for exploration of Bordetella Allele and profiles databases



Home > Organism > Bordetella locus/sequence definitions > Search by locus combinations

Search by locus combinations

Schemes

Please select the scheme you would like to query:

Bp_vaccine antigens

Please enter your allelic profile below. Blank loci will be ignored.

ptxP	ptxA	ptxB	ptxC	ptxD	ptxE
3	1	1	4	1	4
fhaB-2400_5550	**fim2**	**fim3**			
1	1	1			

Autofill profile

BPagST:

Options

Search: Exact or nearest match

Display/sort options

Order by: BPagST ascending

Display: 25 records per page

Action

Exact matches found (9 loci).

1 record returned. Click the hyperlink for detailed information.

BPagST	ptxP	ptxA (BP3783)	ptxB (BP3784)	ptxC	ptxD (BP3785)	ptxE (BP3786)	fhaB-2400_5550	fim2 (BP1119)	fim3 (BP1568)
4	3	1	1	4	1	4	1	1	1

Analysis tools

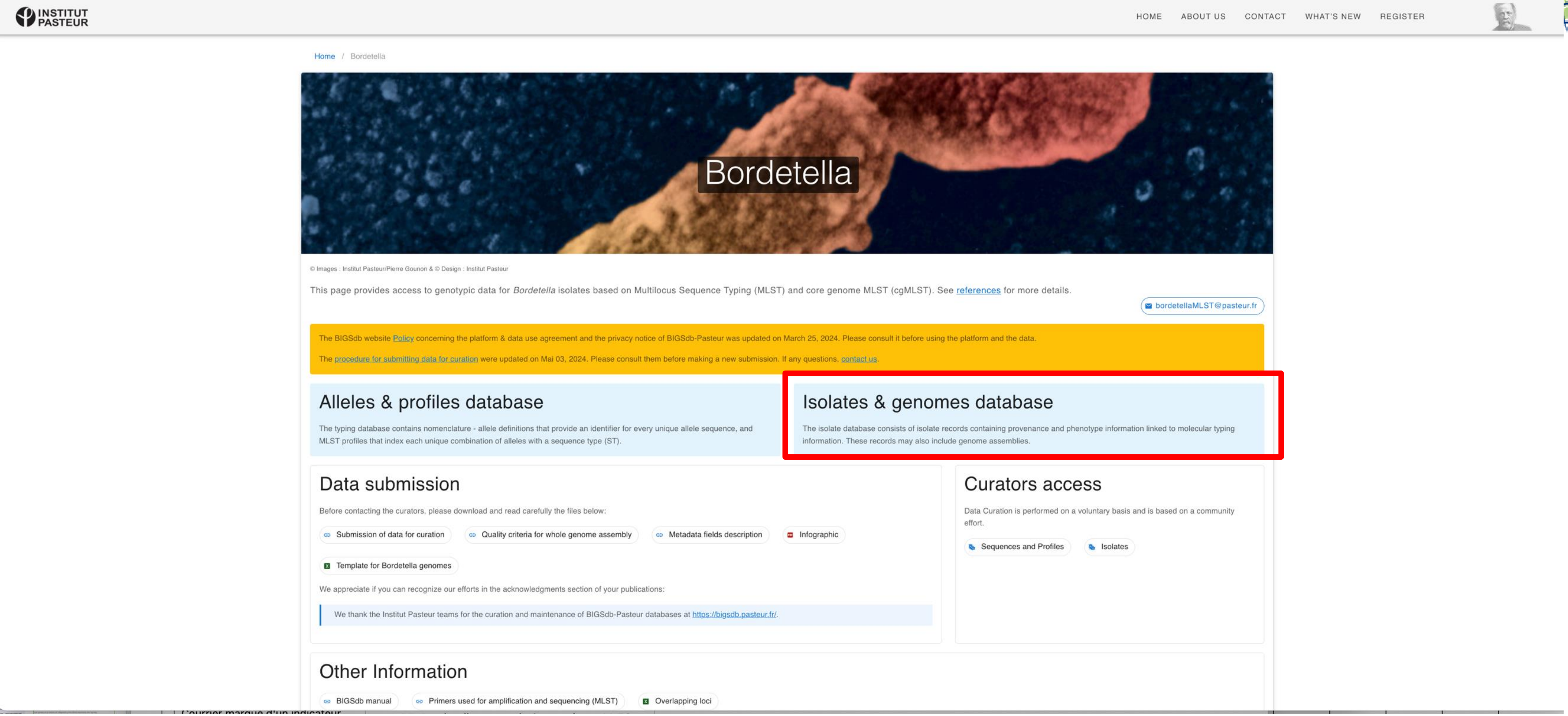
Export:

https://bigsdb.pasteur.fr/bordetella/

BIGSdb Bordetella : overview of Isolates database, genotyping schemes and public projects

https://bigsdb.pasteur.fr/bordetella/

Short Demos on Allele and profiles database



© Images : Institut Pasteur/Pierre Gounon & © Design : Institut Pasteur

This page provides access to genotypic data for *Bordetella* isolates based on Multilocus Sequence Typing (MLST) and core genome MLST (cgMLST). See references for more details.

bordetellaMLST@pasteur.fr

The BIGSdb website Policy concerning the platform & data use agreement and the privacy notice of BIGSdb-Pasteur was updated on March 25, 2024. Please consult it before using the platform and the data.

The procedure for submitting data for curation were updated on Mai 03, 2024. Please consult them before making a new submission. If any questions, contact us.

Alleles & profiles database

The typing database contains nomenclature - allele definitions that provide an identifier for every unique allele sequence, and MLST profiles that index each unique combination of alleles with a sequence type (ST).

Isolates & genomes database

The isolate database consists of isolate records containing provenance and phenotype information linked to molecular typing information. These records may also include genome assemblies.

Data submission

Before contacting the curators, please download and read carefully the files below:

We appreciate if you can recognize our efforts in the acknowledgments section of your publications:

> We thank the Institut Pasteur teams for the curation and maintenance of BIGSdb-Pasteur databases at https://bigsdb.pasteur.fr/.

Curators access

Data Curation is performed on a voluntary basis and is based on a community effort.

Other Information

https://bigsdb.pasteur.fr/bordetella/

BIGSdb Bordetella : overview of Isolates database, genotyping schemes and public projects

BigsDB public library currently includes
 -**3,299 *Bordetella* isolates** and their provenance data,
- 2,799 genomes (mainly from *B. pertussis*, *B. parapertussis*, *B. bronchiseptica*)

Bordetella cgMLST database

This database contains data for a collection of isolates that represent the total known diversity of Bordetella pertussis. For every allelic profile in the profiles/sequence definition database there is at least one corresponding isolate deposited here. Isolates submitted to this database are (in general) those that represent novel allelic profiles and consequently it should be noted that the database does not represent an unbiased population sample.

Restricted view: Note that you are currently restricted to viewing or downloading data that was submitted on or prior to 2024-10-07. Please log in to access the full dataset.

Record versions: current; Record creation: all time

Continent
Top 5 values

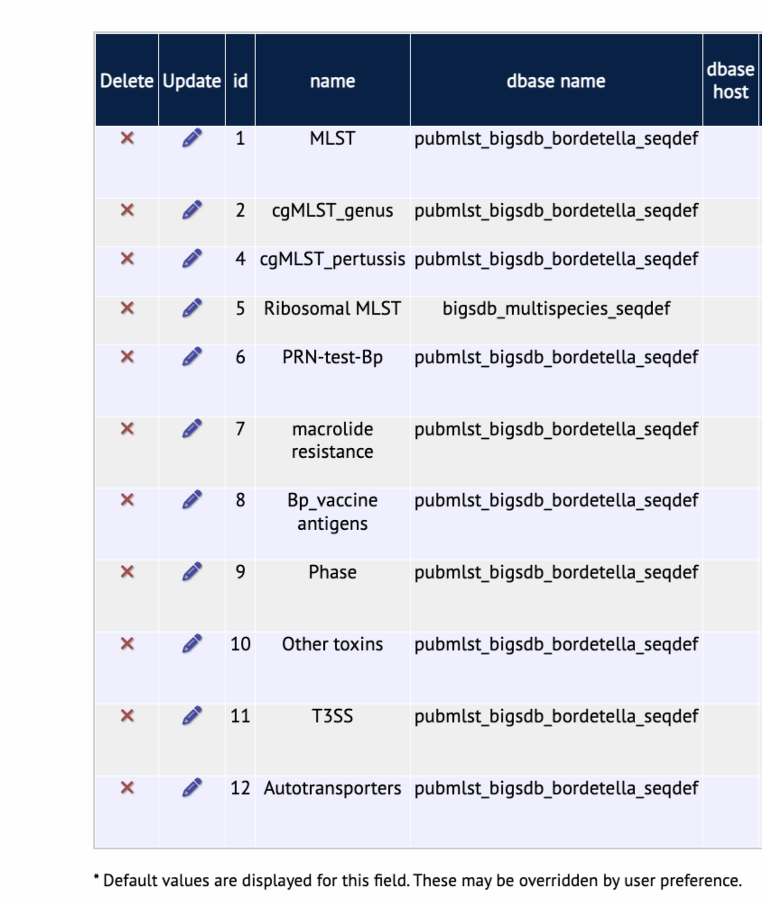
Value	Frequency
Europe	1,203
North America	738
Asia	168
Oceania	154
Africa	49

22/09/2024

https://bigsdb.pasteur.fr/bordetella/

-from different continents
-mostly collected after the year 2000

BigsDB Database for Bordetellae spp. Genotyping schemes



Delete	Update	id	name	dbase name	dbase host
✕	✎	1	MLST	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	2	cgMLST_genus	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	4	cgMLST_pertussis	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	5	Ribosomal MLST	bigsdb_multispecies_seqdef	
✕	✎	6	PRN-test-Bp	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	7	macrolide resistance	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	8	Bp_vaccine antigens	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	9	Phase	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	10	Other toxins	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	11	T3SS	pubmlst_bigsdb_bordetella_seqdef	
✕	✎	12	Autotransporters	pubmlst_bigsdb_bordetella_seqdef	

* Default values are displayed for this field. These may be overridden by user preference.

MLST (*adk, fumC, glyA, icd, pepA, pgm, tyrB)*

cgMLST genus (1415 core gene loci) : for all Bordetella spp.
cgMLST_pertussis (2038 core gene loci) : only for B. pertussis

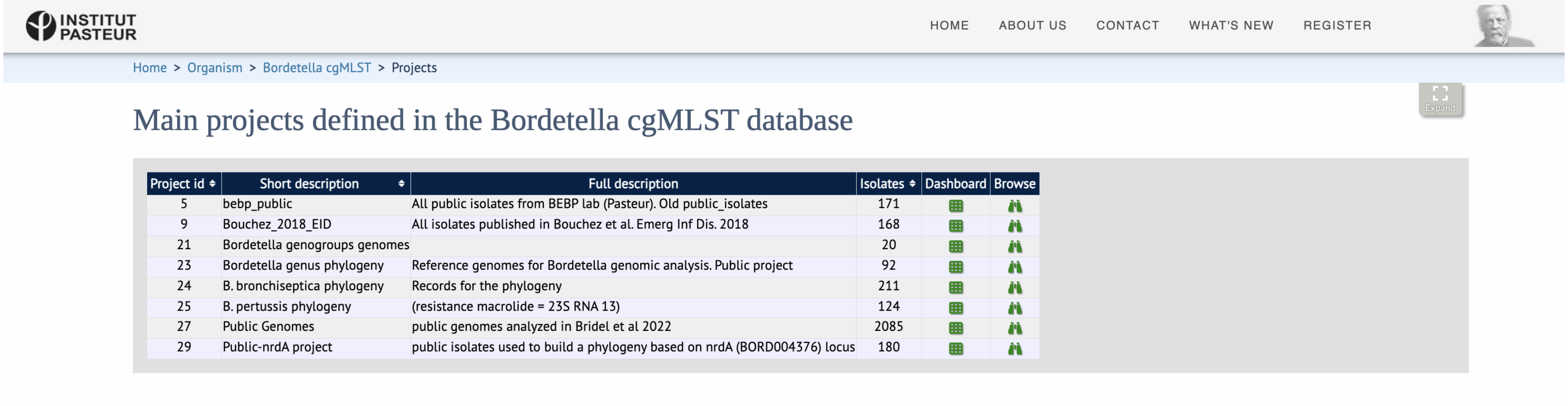
Ribosomal MLST :Ribosomal Multilocus Sequence Typing (rMLST) is an approach that indexes variation of the 53 genes encoding the bacterial ribosome protein subunits (*rps* genes) as a means of integrating microbial taxonomy and typing.
rMLST is described in Jolley *et al.* 2012 *Microbiology* **158:**1005-15

Macrolide resistance (23S_rRNA, fhaB loci)
Bp_vaccine antigens (ptxP, ptxA to E, fim2, fim3, FhaB-2500-5550)
Phase (bipA, bvgA, bvgS)
Other toxins (cyaA, dnt)
T3SS (bopB, bopD, bopN, bsp22, bteA)
Autotransporters (bapC, brkA, prn, tcfA, vag8)

https://bigsdb.pasteur.fr

BigsDB Database for Bordetellae spp.

Public projects



Home > Organism > Bordetella cgMLST > Projects

Main projects defined in the Bordetella cgMLST database

Project id	Short description	Full description	Isolates	Dashboard	Browse
5	bebp_public	All public isolates from BEBP lab (Pasteur). Old public_isolates	171	▦	🔍
9	Bouchez_2018_EID	All isolates published in Bouchez et al. Emerg Inf Dis. 2018	168	▦	🔍
21	Bordetella genogroups genomes		20	▦	🔍
23	Bordetella genus phylogeny	Reference genomes for Bordetella genomic analysis. Public project	92	▦	🔍
24	B. bronchiseptica phylogeny	Records for the phylogeny	211	▦	🔍
25	B. pertussis phylogeny	(resistance macrolide = 23S RNA 13)	124	▦	🔍
27	Public Genomes	public genomes analyzed in Bridel et al 2022	2085	▦	🔍
29	Public-nrdA project	public isolates used to build a phylogeny based on nrdA (BORD004376) locus	180	▦	🔍

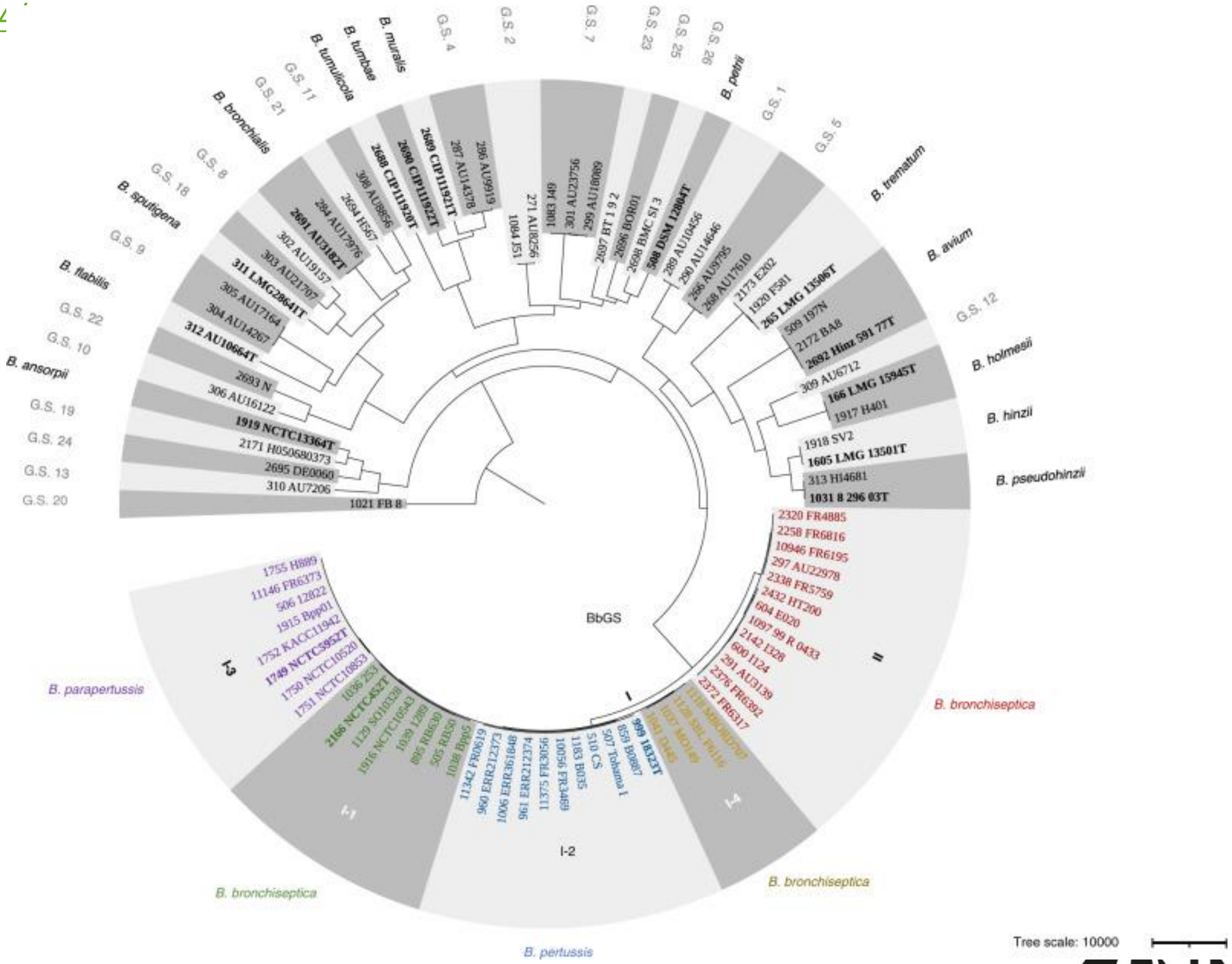
https://bigsdb.pasteur.fr/bordetella/

A comprehensive resource for Bordetella genomic epidemiology and biodiversity studies

Bridel et al., 2022, Nat. Com. PMID: 35778384

https://pubmed.ncbi.nlm.nih.gov/3577838

Bordetella genus Phylogeny
Project id=23. ***B. genus***



92 Bordetella genomes

1415 core gene sequences
cgMLST_genus scheme

https://itol.embl.de/shared/1l7Fw0AvKOoCF.

https://bigsdb.pasteur.fr

BIGSdb Bordetella : overview of Isolates database, genotyping schemes and public projects

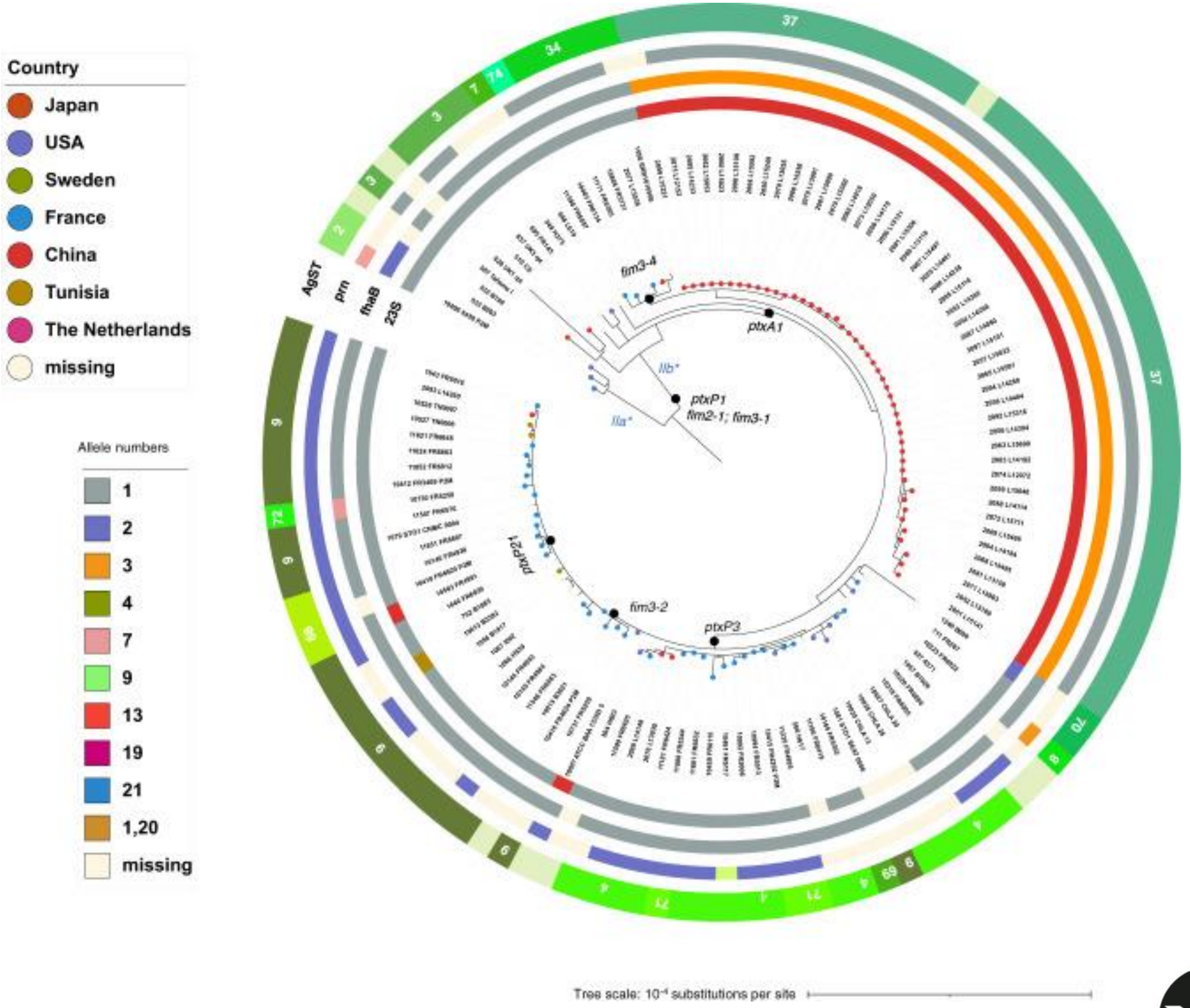
A comprehensive resource for Bordetella genomic epidemiology and biodiversity studies
Bridel et al., 2022, Nat. Com. PMID: 35778384
https://pubmed.ncbi.nlm.nih.gov/35778384/

Bordetella genus Phylogeny
Project id=25. *B. pertussis (macrolide resistance)*

124 genomes

2038 core genome loci
cgMLST_pertussis scheme

https://itol.embl.de/shared/1l7Fw0AvKOoCF

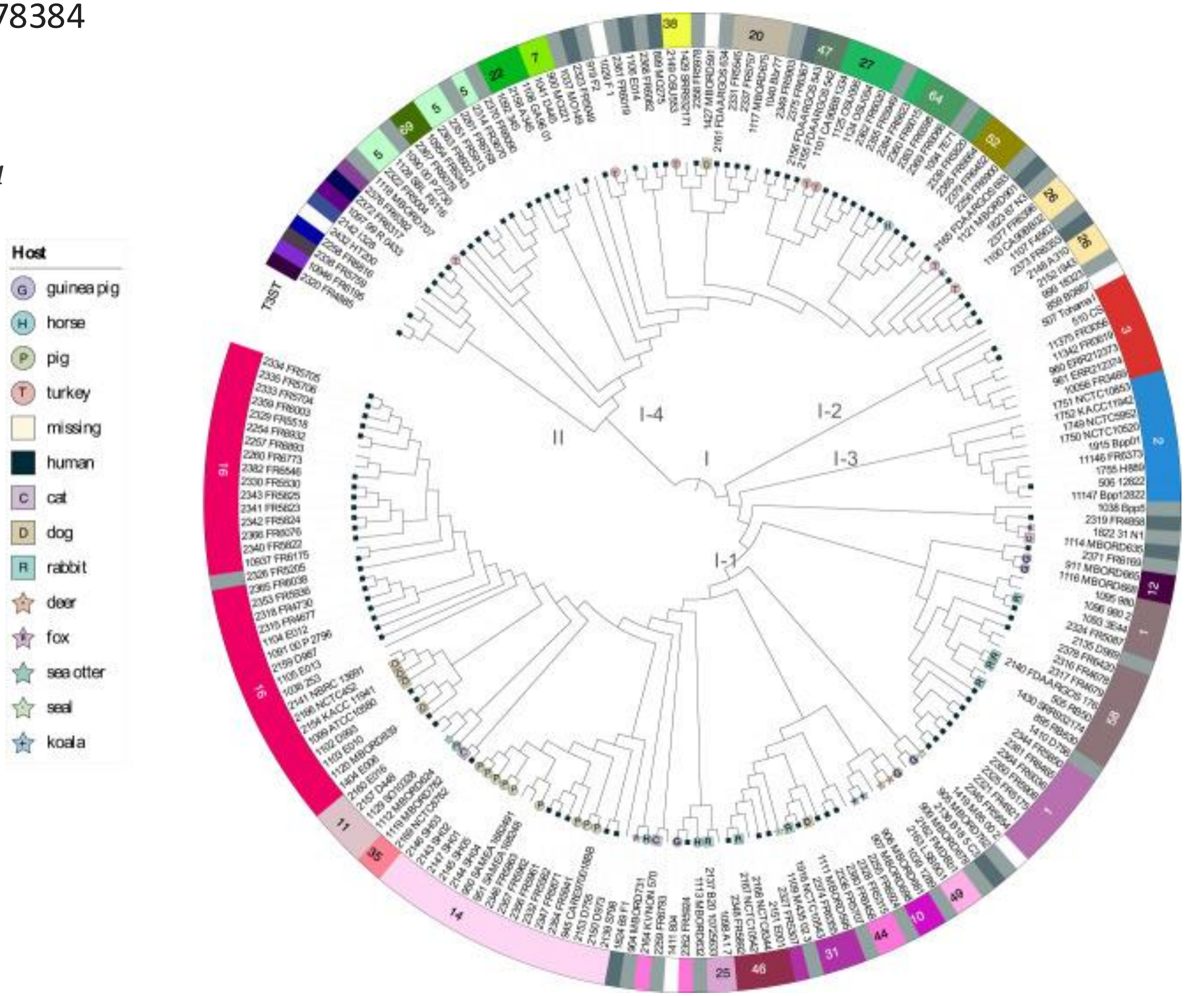


https://bigsdb.pasteur.fr/bordetella/

BIGSdb Bordetella : overview of Isolates database, genotyping schemes and public projects

A comprehensive resource for Bordetella genomic epidemiology and biodiversity studies
Bridel et al., 2022, Nat. Com. PMID: 35778384
https://pubmed.ncbi.nlm.nih.gov/35778384/

Cladogram of the *Bordetella bronchiseptica*



186 *B. bronchiseptica* genomes, and representatives of the phylogenetic diversity of *B. pertussis* (9 genomes) and *B. parapertussis* (9 genomes)

1415 core gene loci (*cgMLST_genus* scheme)

https://itol.embl.de/shared/1l7Fw0AvKOoCF

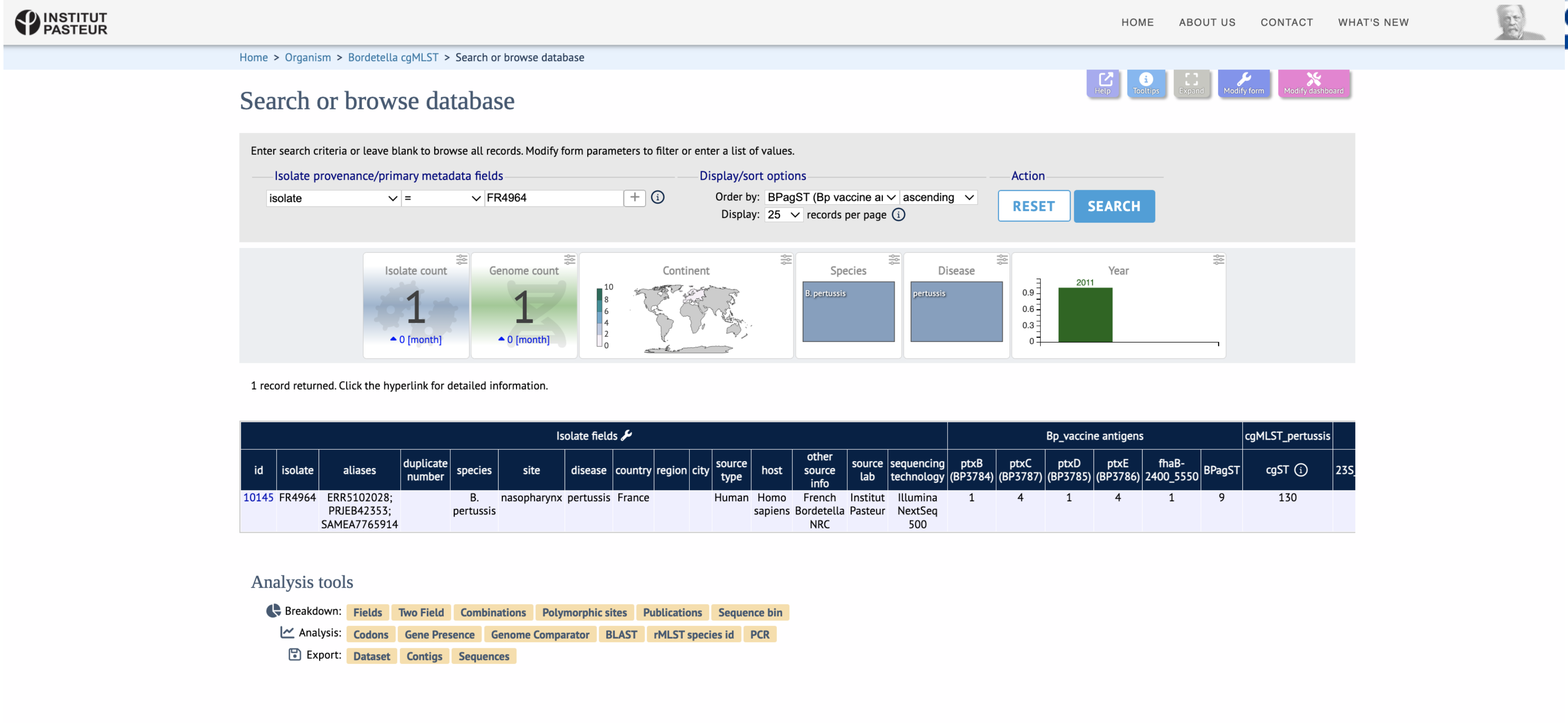
https://bigsdb.pasteur.fr/bordetella/

https://bigsdb.pasteur.fr/cgi-bin/bigsdb/bigsdb.pl?db=pubmlst_bordetella_isolates&page=logout

Analysis on BIGSdb Bordetella Isolates database (search, genome comparator, GrapeTree)

https://bigsdb.pasteur.fr/bordetella/

Analysis on BIGSdb Bordetella isolates database (genome comparator, GrapeTree…)



Home > Organism > Bordetella cgMLST > Search or browse database

Search or browse database

Enter search criteria or leave blank to browse all records. Modify form parameters to filter or enter a list of values.

Isolate provenance/primary metadata fields: isolate = FR4964

Display/sort options:
Order by: BPagST (Bp vaccine ar ascending
Display: 25 records per page

Action

1 record returned. Click the hyperlink for detailed information.

Isolate fields															Bp_vaccine antigens						cgMLST_pertussis	
id	isolate	aliases	duplicate number	species	site	disease	country	region	city	source type	host	other source info	source lab	sequencing technology	ptxB (BP3784)	ptxC (BP3787)	ptxD (BP3785)	ptxE (BP3786)	fhaB-2400_5550	BPagST	cgST	23S
10145	FR4964	ERR5102028; PRJEB42353; SAMEA7765914		B. pertussis	nasopharynx	pertussis	France			Human	Homo sapiens	French Bordetella NRC	Institut Pasteur	Illumina NextSeq 500	1	4	1	4	1	9	130	

Analysis tools

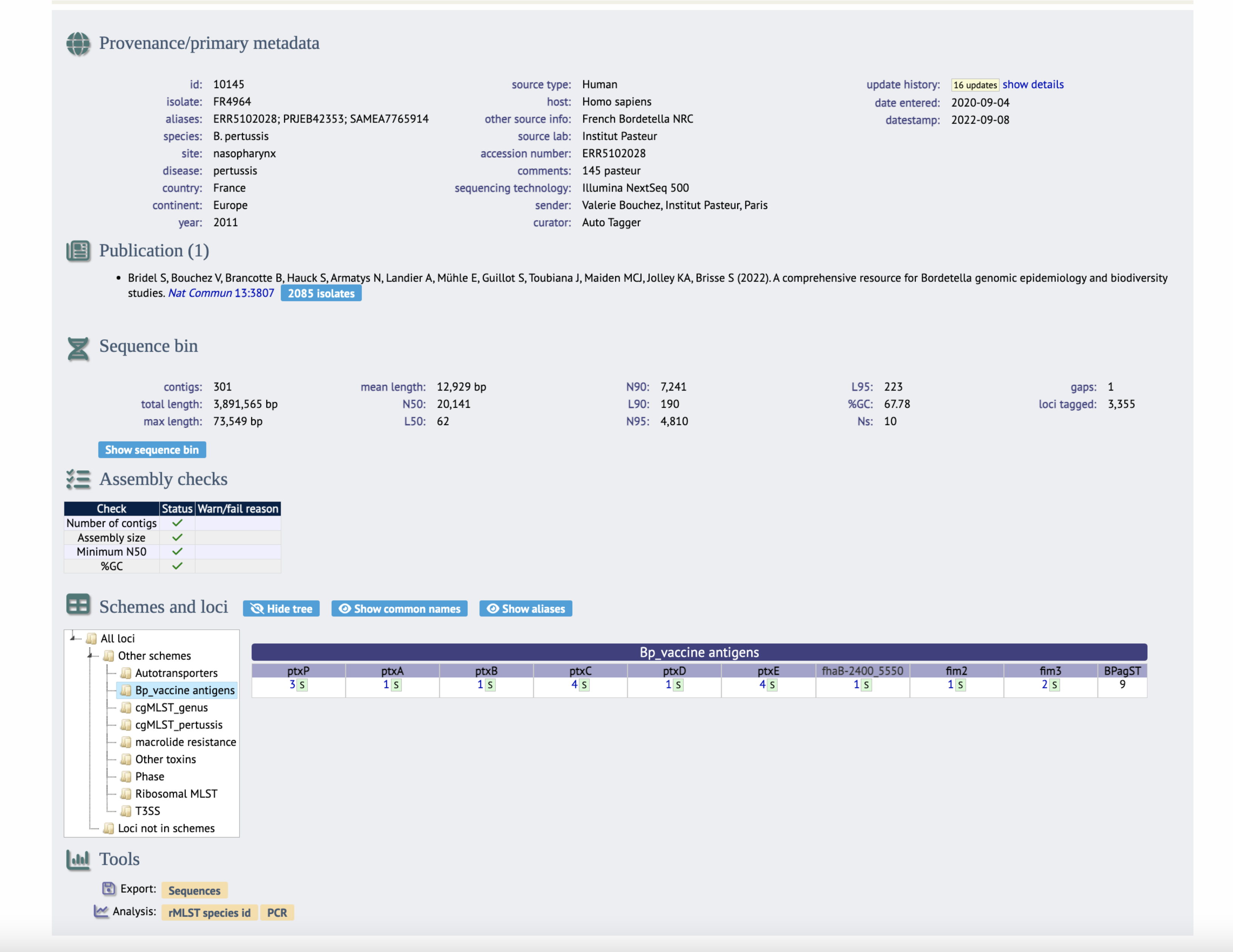
Breakdown: Fields | Two Field | Combinations | Polymorphic sites | Publications | Sequence bin

Analysis: Codons | Gene Presence | Genome Comparator | BLAST | rMLST species id | PCR

Export: Dataset | Contigs | Sequences

https://bigsdb.pasteur.fr/bordetella/

Analysis on BIGSdb Bordetella isolates database (genome comparator, GrapeTree…)



Provenance/primary metadata

id: 10145
isolate: FR4964
aliases: ERR5102028; PRJEB42353; SAMEA7765914
species: B. pertussis
site: nasopharynx
disease: pertussis
country: France
continent: Europe
year: 2011

source type: Human
host: Homo sapiens
other source info: French Bordetella NRC
source lab: Institut Pasteur
accession number: ERR5102028
comments: 145 pasteur
sequencing technology: Illumina NextSeq 500
sender: Valerie Bouchez, Institut Pasteur, Paris
curator: Auto Tagger

update history: 16 updates show details
date entered: 2020-09-04
datestamp: 2022-09-08

Publication (1)

- Bridel S, Bouchez V, Brancotte B, Hauck S, Armatys N, Landier A, Mühle E, Guillot S, Toubiana J, Maiden MCJ, Jolley KA, Brisse S (2022). A comprehensive resource for Bordetella genomic epidemiology and biodiversity studies. *Nat Commun* 13:3807 2085 isolates

Sequence bin

contigs: 301
total length: 3,891,565 bp
max length: 73,549 bp
mean length: 12,929 bp
N50: 20,141
L50: 62
N90: 7,241
L90: 190
N95: 4,810
L95: 223
%GC: 67.78
Ns: 10
gaps: 1
loci tagged: 3,355

Assembly checks

Check	Status	Warn/fail reason
Number of contigs	✓	
Assembly size	✓	
Minimum N50	✓	
%GC	✓	

Schemes and loci

Bp_vaccine antigens

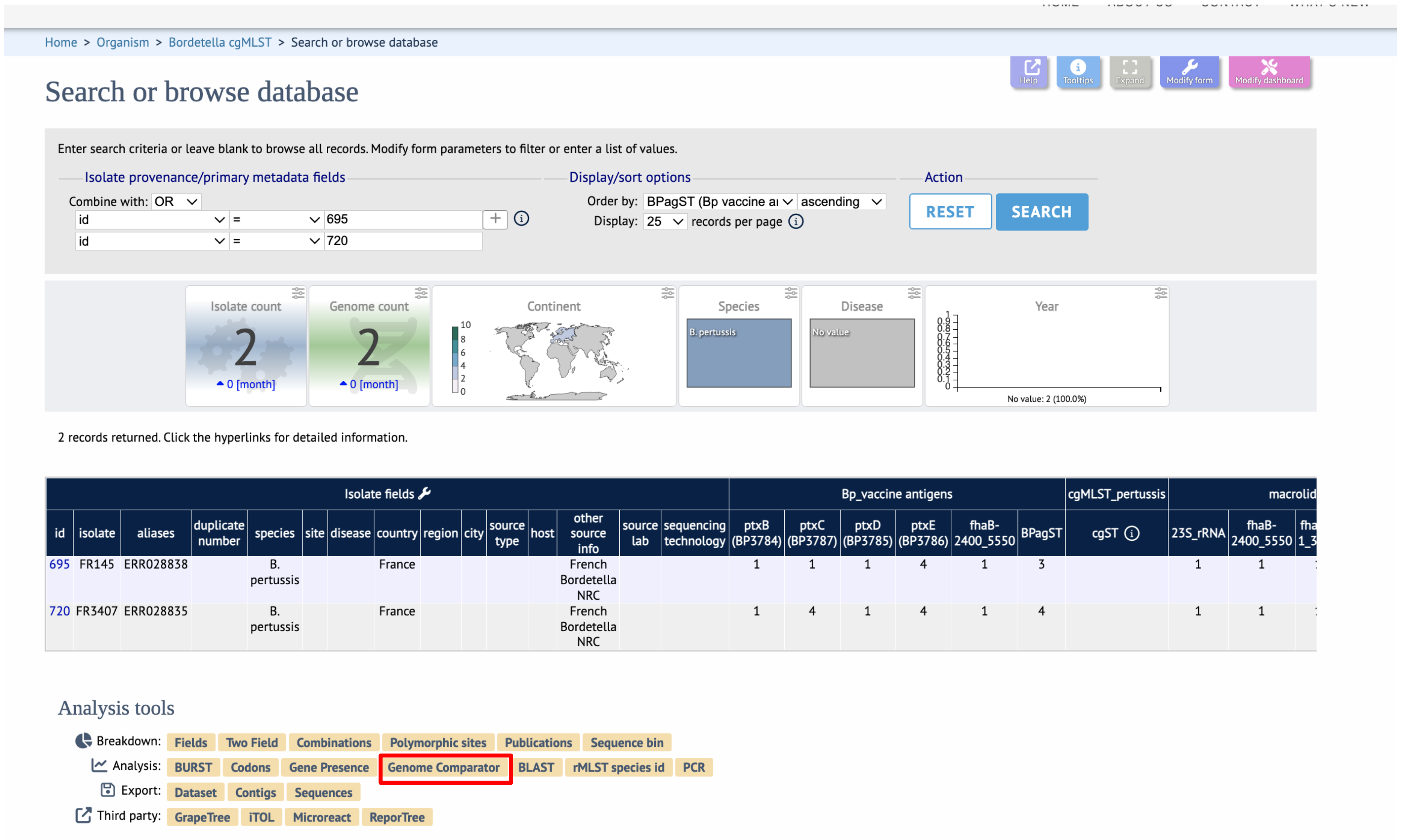
ptxP	ptxA	ptxB	ptxC	ptxD	ptxE	fhaB-2400_5550	fim2	fim3	BPagST
3 S	1 S	1 S	4 S	1 S	4 S	1 S	1 S	2 S	9

Tools

Export: Sequences
Analysis: rMLST species id PCR

https://bigsdb.pasteur.fr/bordetella/

Analysis on BIGSdb Bordetella isolates database (genome comparator, GrapeTree…)



Home > Organism > Bordetella cgMLST > Search or browse database

Search or browse database

Enter search criteria or leave blank to browse all records. Modify form parameters to filter or enter a list of values.

Isolate provenance/primary metadata fields

Combine with: OR

id = 695

id = 720

Display/sort options

Order by: BPagST (Bp vaccine ai ascending

Display: 25 records per page

2 records returned. Click the hyperlinks for detailed information.

Isolate fields															Bp_vaccine antigens						cgMLST_pertussis	macrolid		
id	isolate	aliases	duplicate number	species	site	disease	country	region	city	source type	host	other source info	source lab	sequencing technology	ptxB (BP3784)	ptxC (BP3787)	ptxD (BP3785)	ptxE (BP3786)	fhaB-2400_5550	BPagST	cgST	23S_rRNA	fhaB-2400_5550	fha 1_3
695	FR145	ERR028838		B. pertussis			France					French Bordetella NRC			1	1	1	4	1	3		1	1	
720	FR3407	ERR028835		B. pertussis			France					French Bordetella NRC			1	4	1	4	1	4		1	1	

Analysis tools

Breakdown: Fields, Two Field, Combinations, Polymorphic sites, Publications, Sequence bin

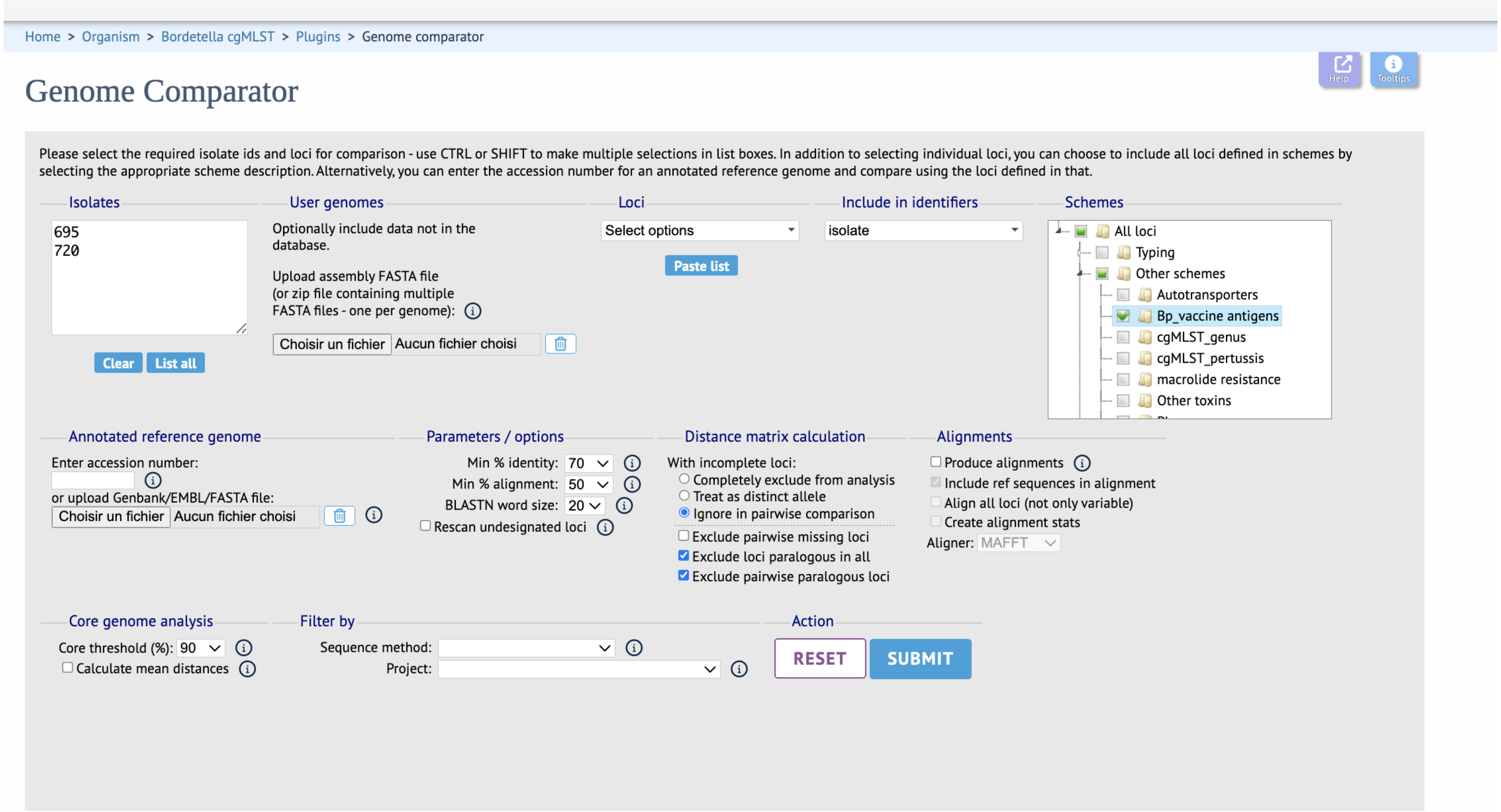
Analysis: BURST, Codons, Gene Presence, Genome Comparator, BLAST, rMLST species id, PCR

Export: Dataset, Contigs, Sequences

Third party: GrapeTree, iTOL, Microreact, ReporTree

https://bigsdb.pasteur.fr/bordetella/

Analysis on BIGSdb Bordetella isolates database (genome comparator, GrapeTree…)



Genome Comparator

Please select the required isolate ids and loci for comparison - use CTRL or SHIFT to make multiple selections in list boxes. In addition to selecting individual loci, you can choose to include all loci defined in schemes by selecting the appropriate scheme description. Alternatively, you can enter the accession number for an annotated reference genome and compare using the loci defined in that.

Isolates

695
720

User genomes

Optionally include data not in the database.

Upload assembly FASTA file (or zip file containing multiple FASTA files - one per genome):

Choisir un fichier Aucun fichier choisi

Loci

Select options

Include in identifiers

isolate

Schemes

- ☑ All loci
  - ☐ Typing
  - ☑ Other schemes
    - ☐ Autotransporters
    - ☑ Bp_vaccine antigens
    - ☐ cgMLST_genus
    - ☐ cgMLST_pertussis
    - ☐ macrolide resistance
    - ☐ Other toxins

Annotated reference genome

Enter accession number:

or upload Genbank/EMBL/FASTA file:

Choisir un fichier Aucun fichier choisi

Parameters / options

Min % identity: 70

Min % alignment: 50

BLASTN word size: 20

☐ Rescan undesignated loci

Distance matrix calculation

With incomplete loci:

- ○ Completely exclude from analysis
- ○ Treat as distinct allele
- ● Ignore in pairwise comparison

☐ Exclude pairwise missing loci

☑ Exclude loci paralogous in all

☑ Exclude pairwise paralogous loci

Alignments

☐ Produce alignments

☑ Include ref sequences in alignment

☐ Align all loci (not only variable)

☐ Create alignment stats

Aligner: MAFFT

Core genome analysis

Core threshold (%): 90

☐ Calculate mean distances

Filter by

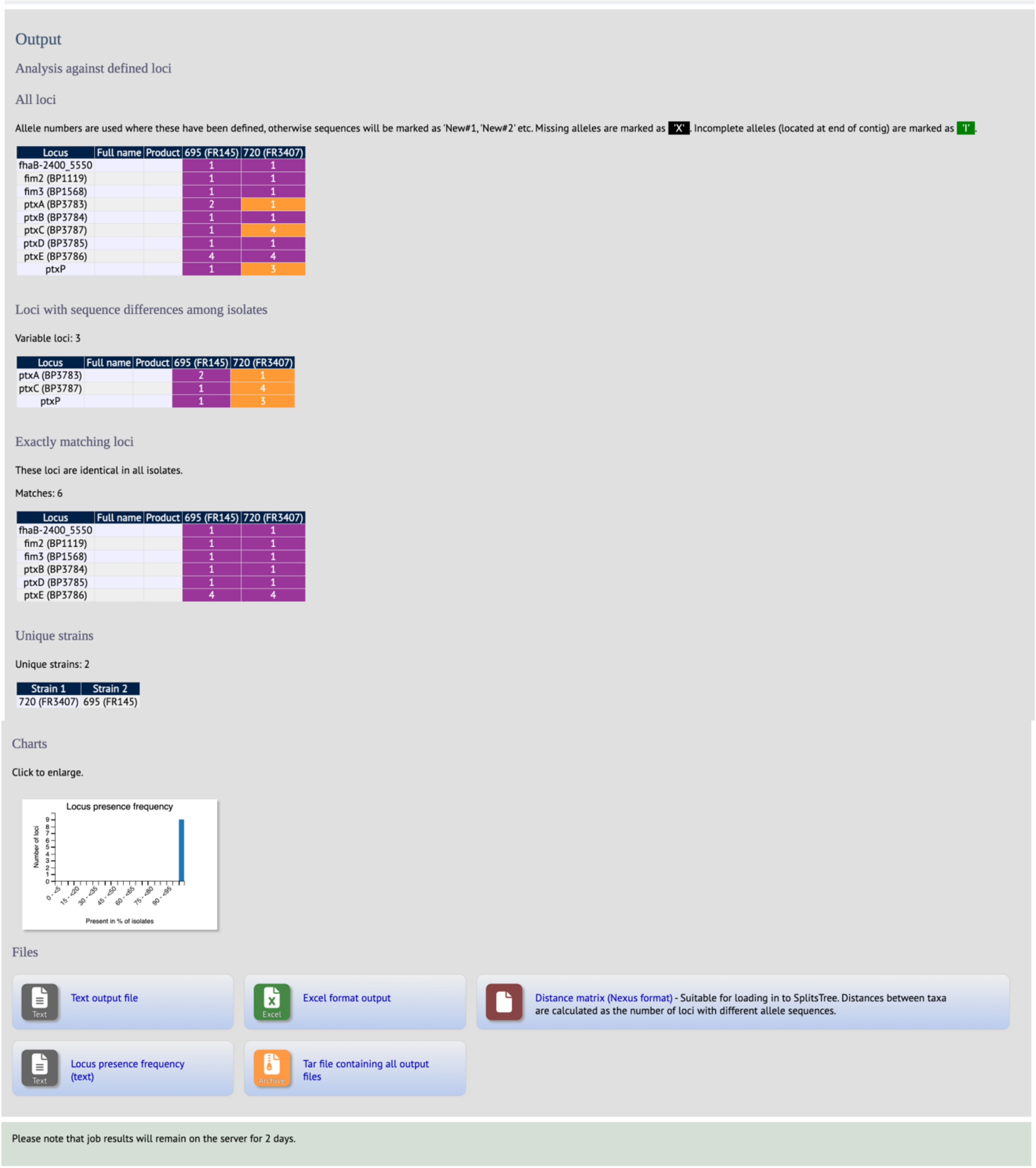
Sequence method:

Project:

Action

https://bigsdb.pasteur.fr/bordetella/

Analysis on BIGSdb Bordetella isolates database (genome comparator, GrapeTree…)



Output

Analysis against defined loci

All loci

Allele numbers are used where these have been defined, otherwise sequences will be marked as 'New#1, 'New#2' etc. Missing alleles are marked as 'X'. Incomplete alleles (located at end of contig) are marked as 'I'.

Locus	Full name	Product	695 (FR145)	720 (FR3407)
fhaB-2400_5550			1	1
fim2 (BP1119)			1	1
fim3 (BP1568)			1	1
ptxA (BP3783)			2	1
ptxB (BP3784)			1	1
ptxC (BP3787)			1	4
ptxD (BP3785)			1	1
ptxE (BP3786)			4	4
ptxP			1	3

Loci with sequence differences among isolates

Variable loci: 3

Locus	Full name	Product	695 (FR145)	720 (FR3407)
ptxA (BP3783)			2	1
ptxC (BP3787)			1	4
ptxP			1	3

Exactly matching loci

These loci are identical in all isolates.

Matches: 6

Locus	Full name	Product	695 (FR145)	720 (FR3407)
fhaB-2400_5550			1	1
fim2 (BP1119)			1	1
fim3 (BP1568)			1	1
ptxB (BP3784)			1	1
ptxD (BP3785)			1	1
ptxE (BP3786)			4	4

Unique strains

Unique strains: 2

Strain 1	Strain 2
720 (FR3407)	695 (FR145)

Charts

Click to enlarge.

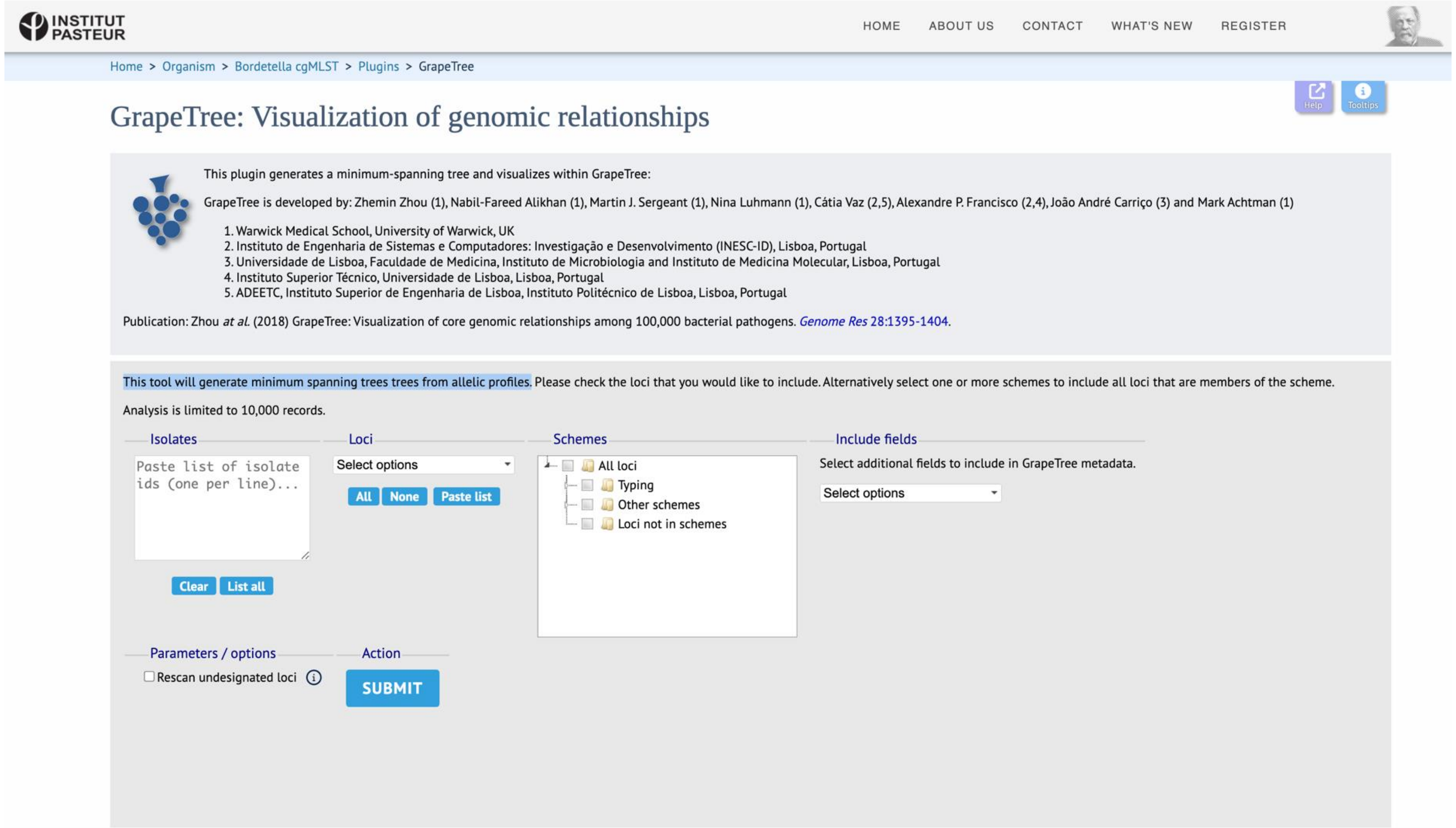
Files

- Text output file
- Excel format output
- Distance matrix (Nexus format) - Suitable for loading in to SplitsTree. Distances between taxa are calculated as the number of loci with different allele sequences.
- Locus presence frequency (text)
- Tar file containing all output files

Please note that job results will remain on the server for 2 days.

https://bigsdb.pasteur.fr/bordetella/

GrapeTree : generate minimum spanning trees trees from allelic profiles.
Analysis available on BIGSdb



GrapeTree: Visualization of genomic relationships

This plugin generates a minimum-spanning tree and visualizes within GrapeTree:

GrapeTree is developed by: Zhemin Zhou (1), Nabil-Fareed Alikhan (1), Martin J. Sergeant (1), Nina Luhmann (1), Cátia Vaz (2,5), Alexandre P. Francisco (2,4), João André Carriço (3) and Mark Achtman (1)

1. Warwick Medical School, University of Warwick, UK
2. Instituto de Engenharia de Sistemas e Computadores: Investigação e Desenvolvimento (INESC-ID), Lisboa, Portugal
3. Universidade de Lisboa, Faculdade de Medicina, Instituto de Microbiologia and Instituto de Medicina Molecular, Lisboa, Portugal
4. Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal
5. ADEETC, Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, Lisboa, Portugal

Publication: Zhou *at al.* (2018) GrapeTree: Visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Res* 28:1395-1404.

This tool will generate minimum spanning trees trees from allelic profiles. Please check the loci that you would like to include. Alternatively select one or more schemes to include all loci that are members of the scheme.

Analysis is limited to 10,000 records.

https://bigsdb.pasteur.fr/bordetella/

GrapeTree : generate minimum spanning trees trees from allelic profiles.
Analysis available on BIGSdb

Eurosurveillance
Europe's journal on infectious disease epidemiology, prevention and control since 1996
Publisher: European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden

Journal Articl

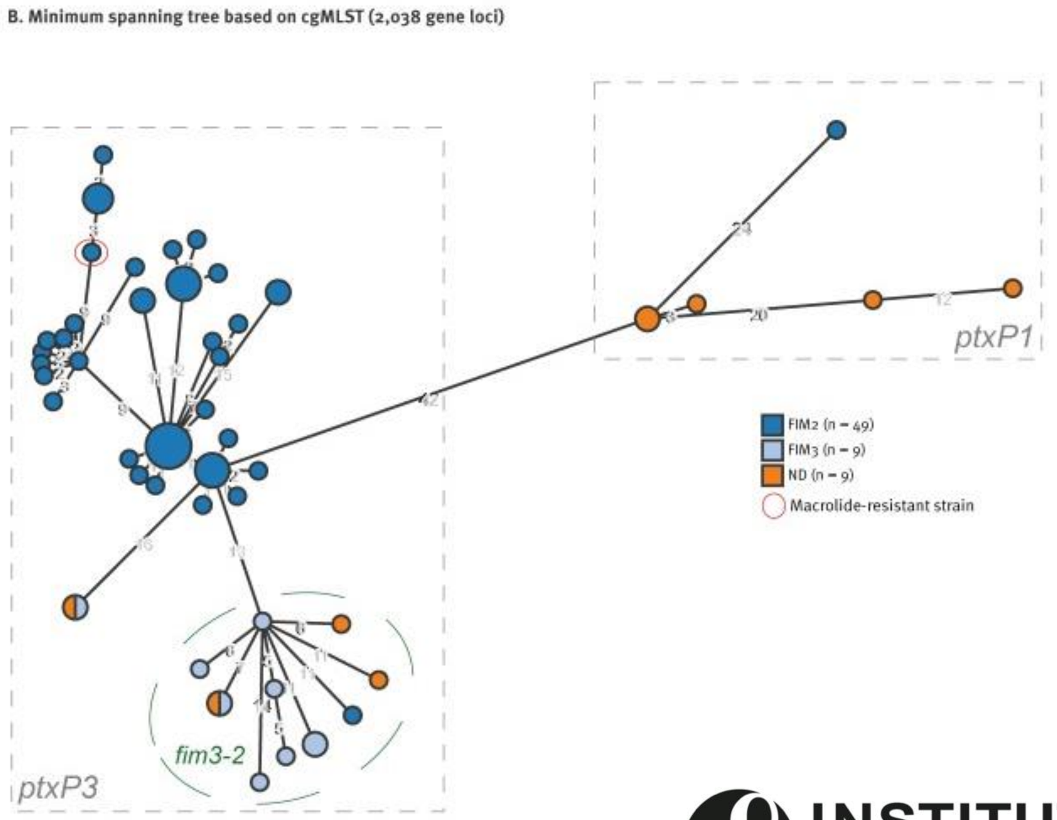
Euro Surveill. 2024 Aug 1; 29(31): 2400459. PMCID: PMC11295439
doi: 10.2807/1560-7917.ES.2024.29.31.2400459 PMID: 39092529

Resurgence of *Bordetella pertussis*, including one macrolide-resistant isolate, France, 2024

Carla Rodrigues,[1,2,*] Valérie Bouchez,[1,2,*] Anaïs Soares,[3] Sabine Trombert-Paolantoni,[4] Fatima Aït El Belghiti,[5] Jérémie F Cohen,[6,7] Nathalie Armatys,[1,2] Annie Landier,[1,2] Thomas Blanchot,[3] Marie Hervo,[3] REMICOQ study group,[8] Julie Toubiana,[1,2,6,**] and Sylvain Brisse[1,2,**]

Minimum spaninning tree based on cgMLST_pertussis (2038 loci)
The numbers on the branches indicate the number of allelic differences among cgST profiles. Each circle represents a cgST type, coloured by FIM serotype. The diameter of circles is related to the number of isolates they comprise

https://bigsdb.pasteur.fr/bordetella/



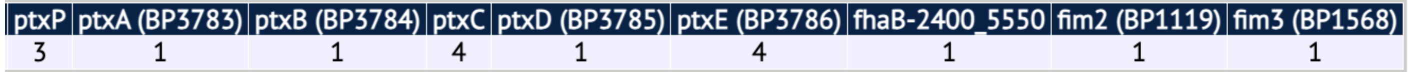
B. Minimum spanning tree based on cgMLST (2,038 gene loci)

Practical session on BIGSdb Bordetella

Part1 :

In BIGSdb Bordetella Alleles & profiles database :

Considering locus *fim3*, compare alleles 1 and 2 (i.e fim3-1 and fim 3-2) using locus explorer;
What Bp-agST correspond to the following profiles :



ptxP	ptxA (BP3783)	ptxB (BP3784)	ptxC	ptxD (BP3785)	ptxE (BP3786)	fhaB-2400_5550	fim2 (BP1119)	fim3 (BP1568)
3	1	1	4	1	4	1	1	1

ptxP	ptxA (BP3783)	ptxB (BP3784)	ptxC	ptxD (BP3785)	ptxE (BP3786)	fhaB-2400_5550	fim2 (BP1119)	fim3 (BP1568)
3	1	1	4	1	4	1	1	2

Part2 :

In BIGSdb Bordetella Isolates database :

Considering all isolates from the public project 25,
Visualize the percentage of isolates per country in a pie chart (Field breakdown option)
Visualize the number of isolates resistant to macrolide per country (Two field breakdown option)
Using genome comparator, compare isolates FR5015 and FR5016, is it the same strain?
Using genome comparator, compare isolates FR5015 and FR4964, is it the same strain?

Part 3 :

In BIGSdb Bordetella Isolates database :

Considering all isolates from the public **project 25,**
Perform a first GrapeTree analysis based on Bp-vaccine-Antigen scheme
Color isolates per country, label BpAg-ST, indicate branch length
Visualize metadata

Perform a second GrapeTree based on cgMLST_pertussis
Color isolates per BpAg-ST, indicate branch length
Export image
Use Microreact to visualize analysis based on Bp-vaccine antigens and explore Microreact functions

https://bigsdb.pasteur.fr/bordetella/

Corrections of pratical exercices on

Novembre 5 th

Acknowledgements

The creation of this training material was commissioned by ECDC to Institut Pasteur with the direct involvement of Valérie Bouchez.