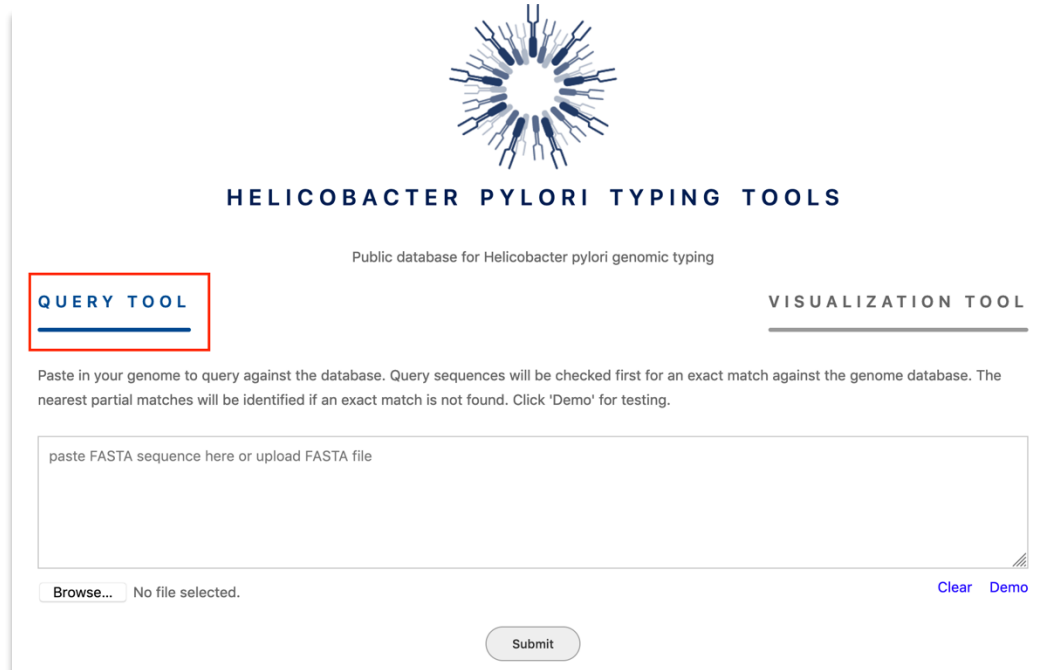


## User instruction

*Helicobacter pylori* typing tool (HPTT) is a genomic typing method based on whole genome sequencing (WGS) data, which can facilitate genetic population analysis and epidemiological origin tracing. In our *H. pylori* typing website (<https://db.cngb.org/HPTT/>), two modules are available, including both genetic typing and geographic visualizing of the global *H. pylori* isolates. Here is the website instruction for users.

### Homepage

You can see the brief introduction of our HPTT on the homepage of our website. Two modules were displayed at this page: the genomic typing tool of *H. pylori* (“[QUERY TOOL](#)”) and the phylogeographic visualization tool of *H. pylori* (“[VISUALIZATION TOOL](#)”).



**HELICOBACTER PYLORI TYPING TOOLS**

Public database for *Helicobacter pylori* genomic typing

**QUERY TOOL** **VISUALIZATION TOOL**

Paste in your genome to query against the database. Query sequences will be checked first for an exact match against the genome database. The nearest partial matches will be identified if an exact match is not found. Click 'Demo' for testing.

paste FASTA sequence here or upload FASTA file

No file selected. [Clear](#) [Demo](#)

### - Data compatibility

In order to achieve the best typing results, it is highly recommended to upload sequencing data generated by the MPS (Massive Parallel Sequencing) method. The common MPS platform includes MGI DNBSEQ® platform and Illumina® platform. General workflow for the MPS including sample pre-treatment, library preparation, sequencing and data analysis. Companies who doing MPS would offer the solutions for all steps, including reagent, instrument, etc.

### - The genomic typing tool of *H. pylori*

At the homepage of website, a tested demo genome is provided. Please see the screenshot below. Click the "Demo" button at the bottom right corner of the input box. It may take few seconds since the genome file is relatively large, and there might be a slight freeze during the loading. If the genome is entered incorrectly, click the "Clear" button just next to the "Demo" to clear the input box. For your own sequenced genome, you can either copy and paste in the input box or upload it by click the "Browse" button. After entering/uploading the genome, click the "Submit" button to query against the database.

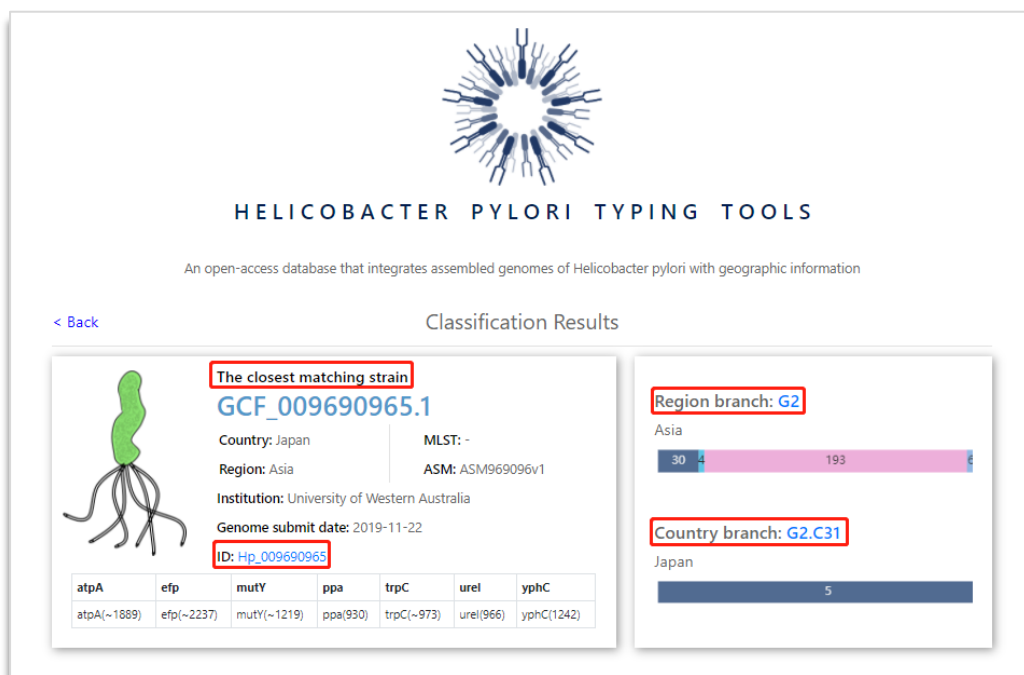
```
>NZ_BEGT01000001.1 Helicobacter pylori strain K21, whole genome shotgun sequence
ACACATAAGCTAAAAAGCCAGCGTTTAAAAAAGAATGATTTTAGAGAGGTGAGCGGATGGGGTGGGTAAAAAGTTTTATTGTTTTGATGCATGGAATTTCC
TAATTTAAAAATAACTAAAGATTAACATGTTTTTAATTTAATATTCATTAAGTGGTGGTTTTGTGGCTATTTCAATTTAATTTTCCATTAACCCAATAACT
AAAATCTTATTTTTATGATAAACTACCCAATTATAATGTCAAATCTTAACTAAAGAAAAACCATGAAAAAAACCCTTTACTAACTCTTTTTCTCGTTTTGGC
TCCACGCTGAAAGGAATGGATTTTTATTAGGTTAAATTTTCAGAAAGGAAAGCTATATTAAGGACAAGGCAGCATCGGCGAAAAAGCTTCAGCAGAAAAACG
CTTTAAATCAAGCGATCAATAACGCAAAAAATTCATTATTCGCCGAACAAAACAAAAAGCCATAAGAGACGCGCAAAACGCTTTAAATGAAGTGAAAGATTTC
AACA AAAATCGCTAACCGATTTCGAGGAAATGGTGGATCGGGCGGCTTTTTAATGAACCTCAGCTTTGGGTATAAATATTTTTGGGTAAAAAAGGATTATA
```

Browse... No file selected. Clear Demo Submit

If your uploaded file is neither *H. pylori* nor sequenced by the second-generation sequencing technology, it might be reported as an error. You will need to check whether your genome fit our criteria.



You will get the typing results if it is successfully typed. Please see the image below. Firstly, the strain with the highest identity ("the closest matching strain") that is most similar to our database will be displayed, including region, country, MLST type, data release time, sequenced research institution, etc. Secondly, if the most similar bacterial species has a typing patterns in our database, it will be displayed on the right. If the corresponding types not completely refer to one continent/country of patterns in our database, the proportion of each continent/country will be displayed in different colours in the bar.



For the region and country groups in our database, each strain can be listed to the corresponding groups. In addition, you can easily access to the visualization tool and NCBI information page by clicking the link (letters in blue).

G2 Region Group List - [233](#) strains

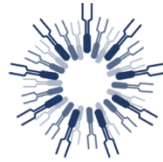
Group	Genome_id	Tip_id	Region	Country	Submit_date	MLST	Link(NCBI)
G2	GCF_000262655.1	<a href="#">Hp_000262655</a>	Asia	China	2012-05-14	3039	<a href="#">more</a>
G2	GCF_000287735.1	<a href="#">Hp_000287735</a>	Asia	China	2013-02-23	2164	<a href="#">more</a>
G2	GCF_000287755.1	<a href="#">Hp_000287755</a>	Asia	China	2013-02-23	2128	<a href="#">more</a>

G2.C31 Country Group List - [5](#) strains

Group	Genome_id	Tip_id	Region	Country	Submit_date	MLST	Link(NCBI)
G2.C31	GCF_002357315.1	<a href="#">Hp_002357315</a>	Asia	Japan	2017-01-17	-	<a href="#">more</a>
G2.C31	GCF_002357415.1	<a href="#">Hp_002357415</a>	Asia	Japan	2017-01-17	-	<a href="#">more</a>
G2.C31	GCF_002357715.1	<a href="#">Hp_002357715</a>	Asia	Japan	2017-01-17	-	<a href="#">more</a>
G2.C31	GCF_002357775.1	<a href="#">Hp_002357775</a>	Asia	Japan	2017-01-17	-	<a href="#">more</a>
G2.C31	GCF_009690965.1	<a href="#">Hp_009690965</a>	Asia	Japan	2019-11-22	-	<a href="#">more</a>

### - The visualization tool of *H. pylori*

You can easily access to the visualization tool at the homepage by clicking the “[VISUALIZATION TOOL](#)” button. All the genomes in our database can be interacted and displayed in both phylogenetic relationships and geographic locations.



## HELICOBACTER PYLORI TYPING TOOLS

Public database for Helicobacter pylori genomic typing

**QUERY TOOL**

**VISUALIZATION TOOL**

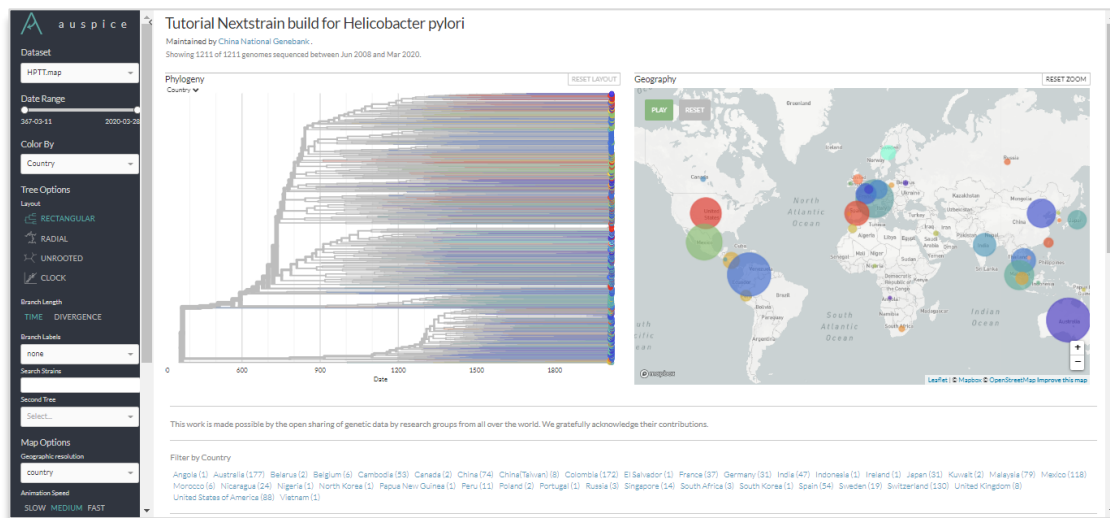
Paste in your genome to query against the database. Query sequences will be checked first for an exact match against the genome database. The nearest partial matches will be identified if an exact match is not found. Click 'Demo' for testing.

paste FASTA sequence here or upload FASTA file

Browse... No file selected.

Clear Demo

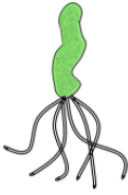
Submit



### - The linking functions

When the genome matched the data in our database, it is easily to see the geographic location by clicking [ID](#).

< Back Classification Results



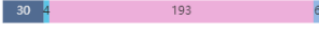
**The closest matching strain**  
**GCF\_009690965.1**

Country: Japan MLST: -  
 Region: Asia ASM: ASM969096v1  
 Institution: University of Western Australia  
 Genome submit date: 2019-11-22  
**ID: Hp\_009690965** →

atpA	efp	mutY	ppa	trpC	urel	yphC
atpA(~1889)	efp(~2237)	mutY(~1219)	ppa(930)	trpC(~973)	urel(966)	yphC(1242)


**Region branch: G2**

Asia



**Country branch: G2.C31**

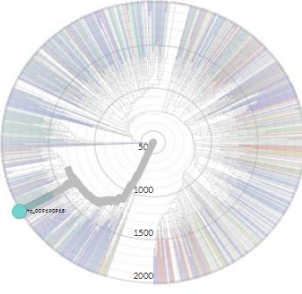
Japan



Tutorial Nextstrain build for *Helicobacter pylori*  
 Maintained by China National Genebank  
 Showing a single strain [Hp\\_009690965](#)

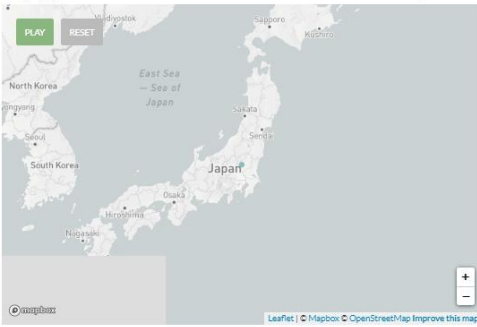
Phylogeny

Country ▾



Geography

PLAY RESET

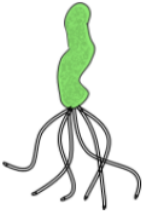


RESET ZOOM

This work is made possible by the open sharing of genetic data by research groups from all over the world. We gratefully acknowledge their contributions.

When the genome matched the data in our database, it is easily to see the phylogenetic relationship and geographic location by clicking [the group ID \(the letters in blue\)](#).

< Back Classification Results




**The closest matching strain**  
**GCF\_009690965.1**

Country: Japan MLST: -  
 Region: Asia ASM: ASM969096v1  
 Institution: University of Western Australia  
 Genome submit date: 2019-11-22  
 ID: [Hp\\_009690965](#)

atpA	efp	mutY	ppa	trpC	urel	yphC
atpA(~1889)	efp(~2237)	mutY(~1219)	ppa(930)	trpC(~973)	urel(966)	yphC(1242)


**Region branch: G2** ←

Asia



**Country branch: G2.C31** ←

Japan



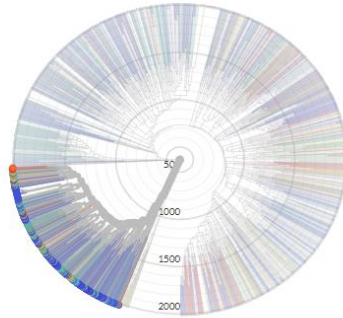
## Tutorial Nextstrain build for *Helicobacter pylori*

Maintained by China National Genebank.

Showing 233 of 1211 genomes sequenced between May 2012 and Dec 2019. Filtered to x | G2 (233).

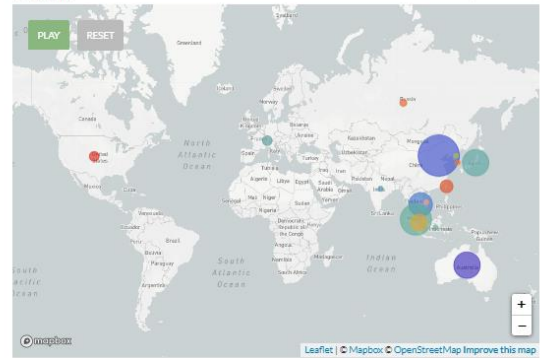
Phylogeny  
Country ▼

RESET LAYOUT



Geography

RESET ZOOM



This work is made possible by the open sharing of genetic data by research groups from all over the world. We gratefully acknowledge their contributions.

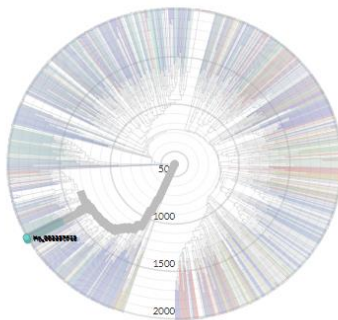
## Tutorial Nextstrain build for *Helicobacter pylori*

Maintained by China National Genebank.

Showing 5 of 1211 genomes sequenced between Jan 2017 and Nov 2019. Filtered to x | G2.C31 (5).

Phylogeny  
Country ▼

RESET LAYOUT



Geography

RESET ZOOM



This work is made possible by the open sharing of genetic data by research groups from all over the world. We gratefully acknowledge their contributions.