

Slides for Augur Steps

Fasta, Snakemake, VCF

Fasta Workflow

- ▶ Highlights all steps
- ▶ Long version (showed abbreviated version at lab meeting - can send too)

▶ Filter

>SWE-2016

AGCAGTAGGTTAGCAATAAC

▶ Align

>ESP-2014

GTAGCAGTAGGTTAGCAGT

▶ Tree

>USA-2016

AGTGAATGGTAGTAGTAATAAT

▶ Refine

>CAN-2015

GCAGTGAATACTAGTGGT

▶ Ancestral

>USA-1960

TAGCAGTGAATGGCAGT

▶ Translate

▶ Clades

▶ Traits

▶ Export

▶ Filter

>SWE-2016 AGCAGTAGGTTAGCAATAAC

▶ Align

>ESP-2014 GTAGCAGTAGGTTAGCAGT

▶ Tree

>USA-2016 AGTGAATGGTAGTAGTAATAAT

▶ Refine

>CAN-2015 GCAGTGAATACTAGTGGT

~~>USA-1960 TAGCAGTGAATGGCAGT~~

▶ Ancestral

▶ Translate

▶ Clades

▶ Traits

▶ Export

- ▶ Filter
- ▶ **Align**
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

>EVD68Ref	GTAGCAGTGAATGGCAGTAGTAATAAT
>SWE-2016	AGCAGTAGGTAGCAATAAC
>ESP-2014	GTAGCAGTAGGTAGCAGT
>USA-2016	AGTGAATGGTAGTAGTAATAAT
>CAN-2015	GCAGTGAATACTAGTGGT

- ▶ Filter
- ▶ **Align**
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

```
>EVD68Ref GTAGCAGTGAATGGCAGTAGTAATAAT
>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN
```

▶ Filter

▶ Align

▶ **Tree**

▶ Refine

▶ Ancestral

▶ Translate

▶ Clades

▶ Traits

▶ Export

>SWE-2016

NNAGCAGTA---GGTAGCAATAACNNN

>ESP-2014

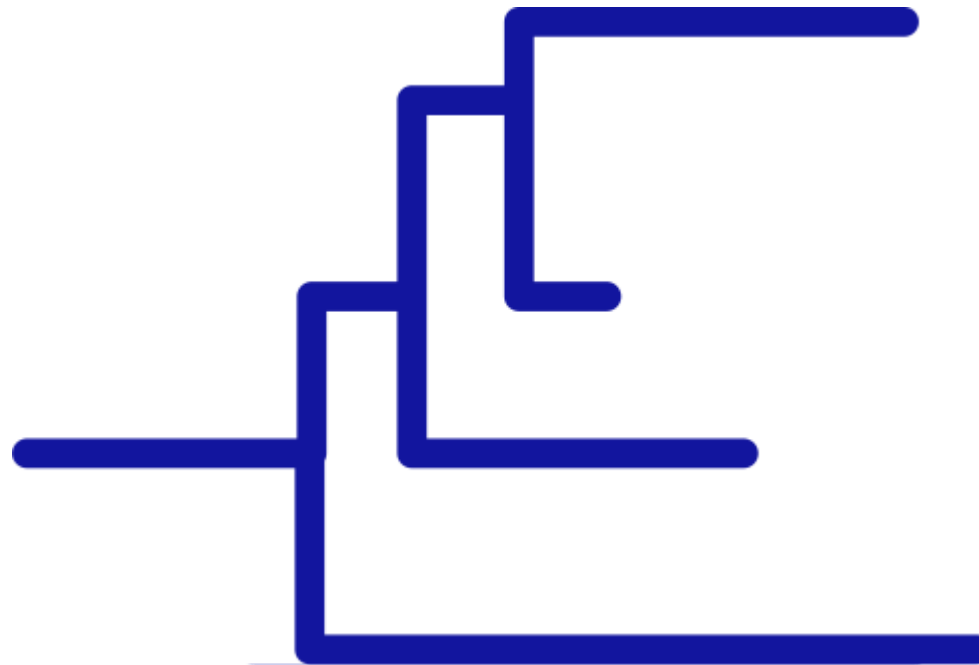
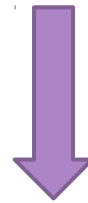
GTAGCAGTA---GGTAGCAGTNNNNNN

>USA-2016

NNNNNAGTGAATGGTAGTAGTAATAAT

>CAN-2015

NNNGCAGTGAATAGTAGTGGTNNNNNN



IQ-TREE

(RAxML, FastTree)

▶ Filter

▶ Align

▶ Tree

▶ Refine

▶ Ancestral

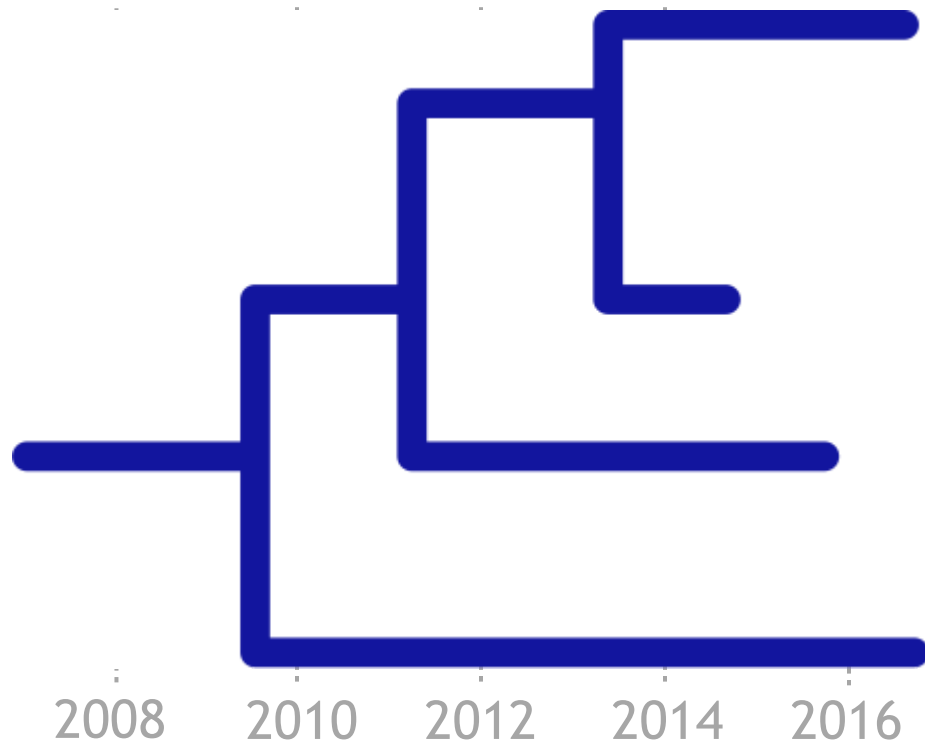
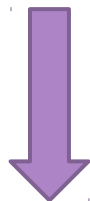
▶ Translate

▶ Clades

▶ Traits

▶ Export

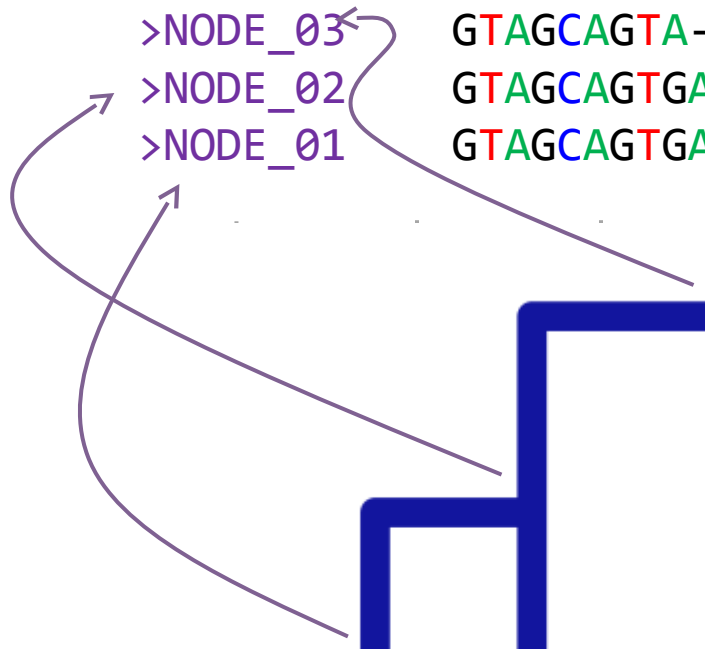
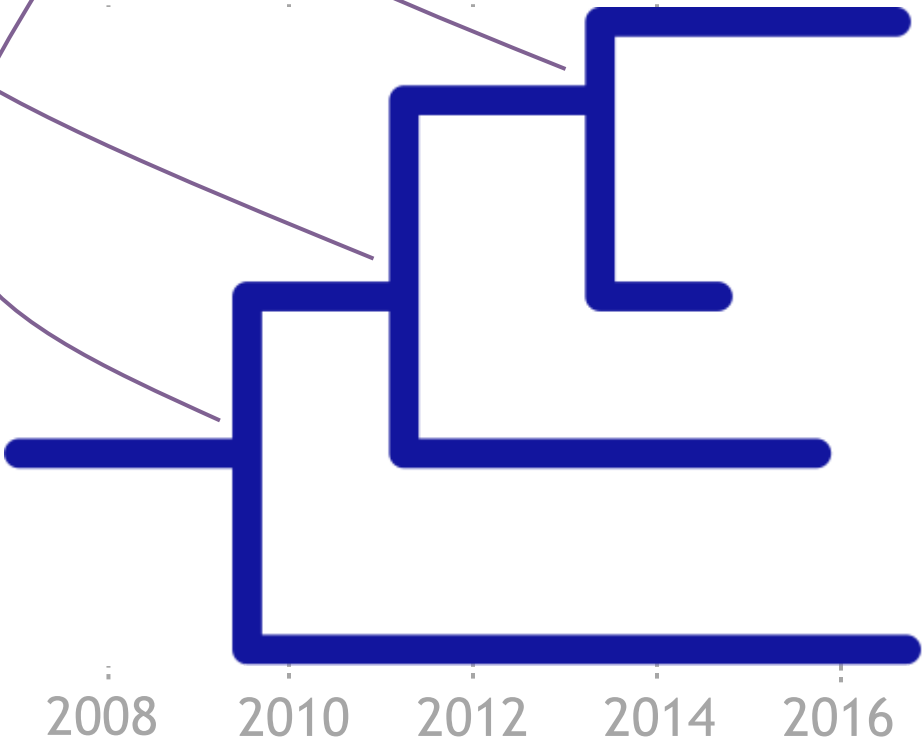
>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN



- ▶ Filter
- ▶ Align
- ▶ Tree
- ▶ Refine
- ▶ **Ancestral**
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

```

>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN
>NODE_03  GTAGCAGTA---GGTAGCAGTAACAAT
>NODE_02  GTAGCAGTGAATGGTAGTAGTAACAAT
>NODE_01  GTAGCAGTGAATGGTAGTAGTAATAAT
  
```

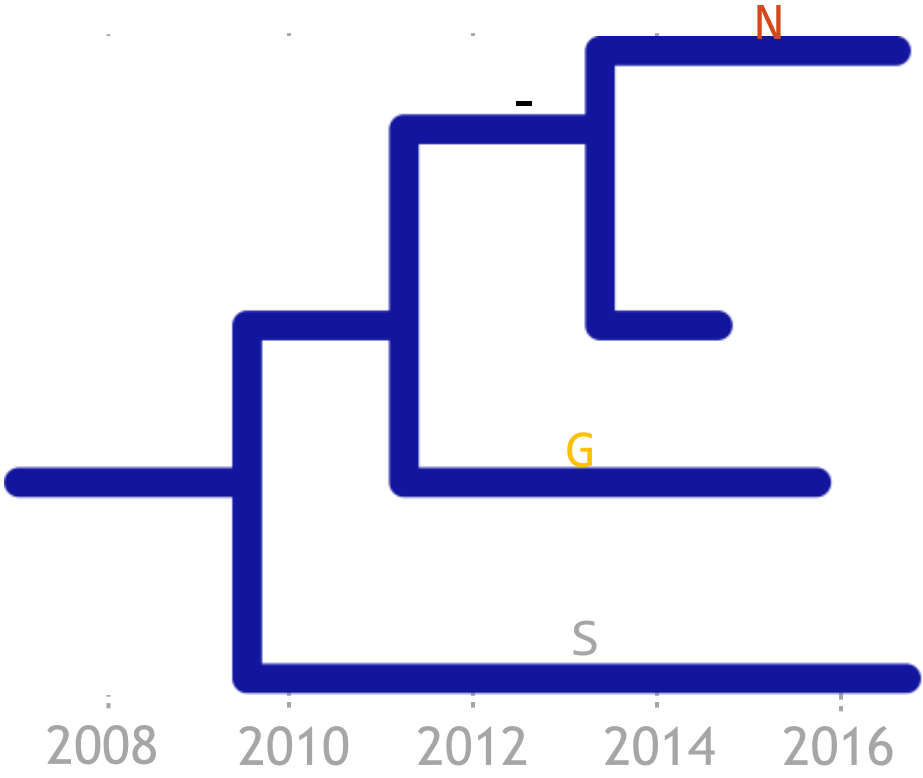


- ▶ Filter
- ▶ Align
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

```

>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN XAV-GSNNX
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN VAV-GSSXX
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT XXVNGSSNN
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN XAVNSSGXX
>NODE_03 GTAGCAGTA---GGTAGCAGTAACAAT VAV-GSSNN
>NODE_02 GTAGCAGTGAATGGTAGTAGTAACAAT VAVNGSSNN
>NODE_01 GTAGCAGTGAATGGTAGTAGTAATAAT VAVNGSSNN

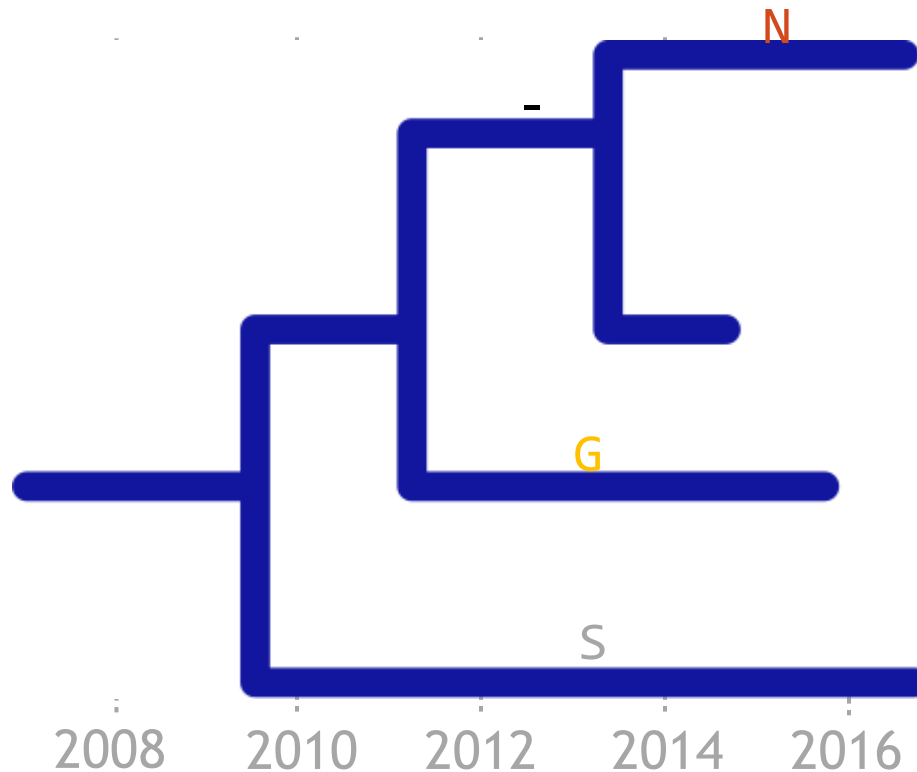
```



- ▶ Filter
- ▶ Align
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

```

>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN XAV-GSNNX
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN VAV-GSSXX
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT XXVNGSSNN
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN XAVNSSGXX
>NODE_03  GTAGCAGTA---GGTAGCAGTAACAAT VAV-GSSNN
>NODE_02  GTAGCAGTGAATGGTAGTAGTAACAAT VAVNGSSNN
>NODE_01  GTAGCAGTGAATGGTAGTAGTAATAAT VAVNGSSNN
  
```



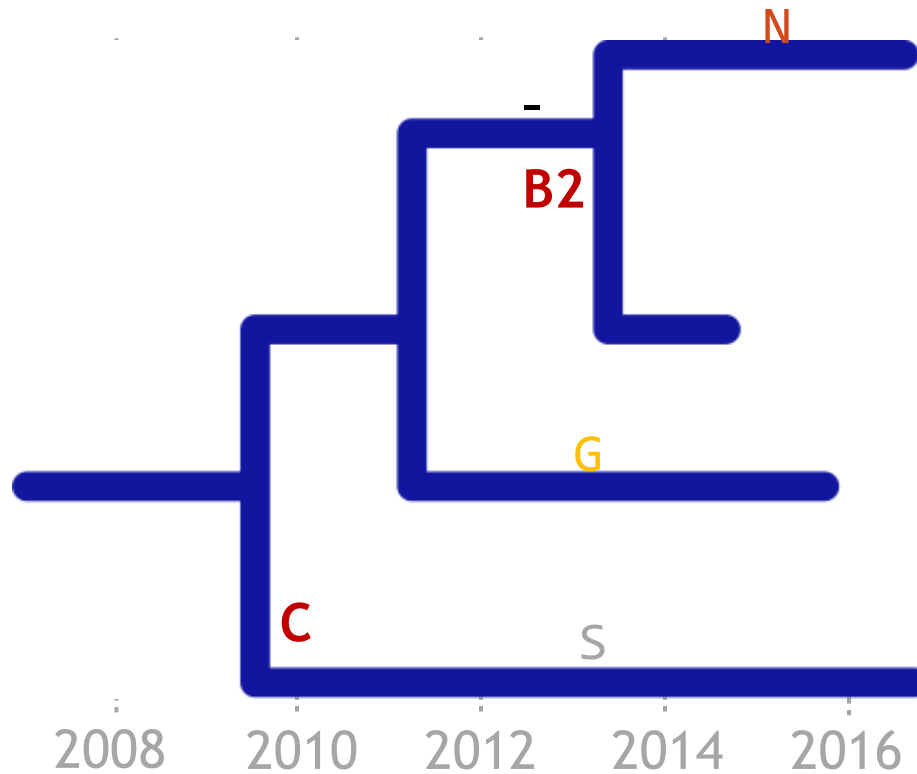
clade	gene	site	alt
B2	nuc	18	C
C	nuc	19	G
A1	VP1	144	S

- ▶ Filter
- ▶ Align
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

```

>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN XAV-GSNNX
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN VAV-GSSXX
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT XXVNGSSNN
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN XAVNSSGXX
>NODE_03  GTAGCAGTA---GGTAGCAGTAACAAT VAV-GSSNN
>NODE_02  GTAGCAGTGAATGGTAGTAGTAACAAT VAVNGSSNN
>NODE_01  GTAGCAGTGAATGGTAGTAGTAATAAT VAVNGSSNN

```



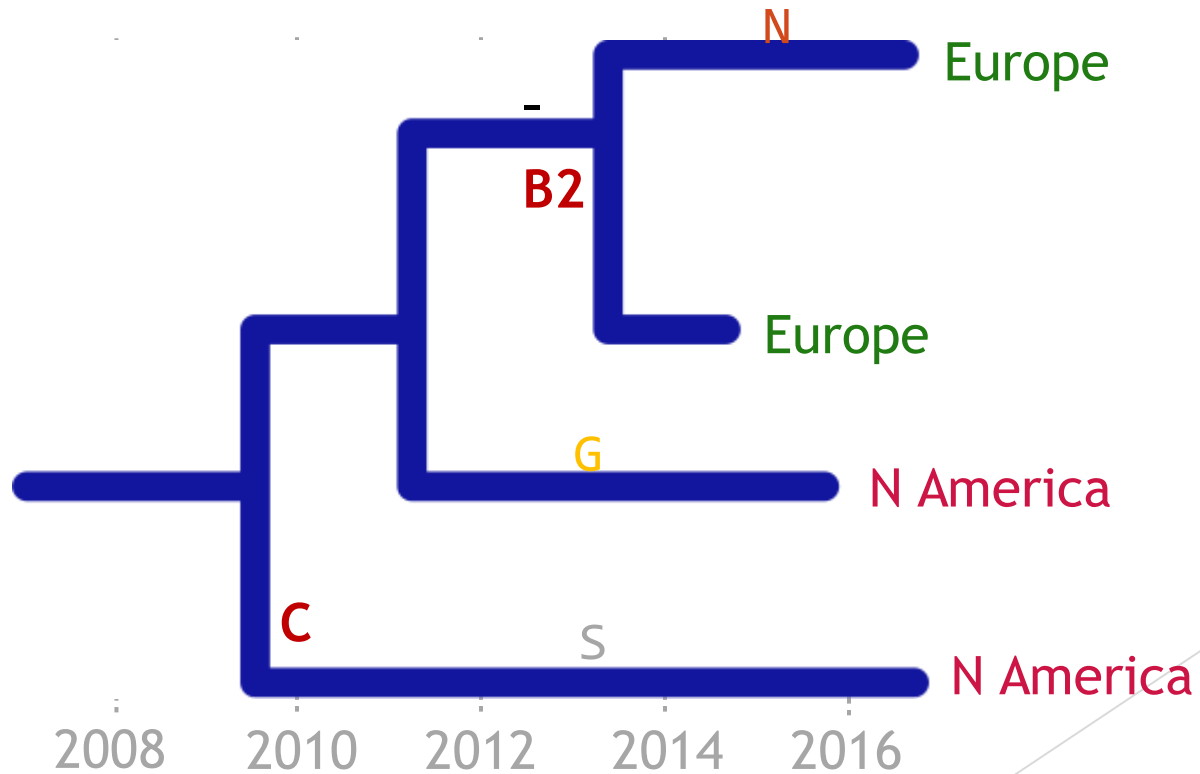
clade	gene	site	alt
B2	nuc	18	C
C	nuc	19	G
A1	VP1	144	S

- ▶ Filter
- ▶ Align
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

```

>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN XAV-GSNNX
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN VAV-GSSXX
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT XXVNGSSNN
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN XAVNSSGXX
>NODE_03 GTAGCAGTA---GGTAGCAGTAACAAT VAV-GSSNN
>NODE_02 GTAGCAGTGAATGGTAGTAGTAACAAT VAVNGSSNN
>NODE_01 GTAGCAGTGAATGGTAGTAGTAATAAT VAVNGSSNN

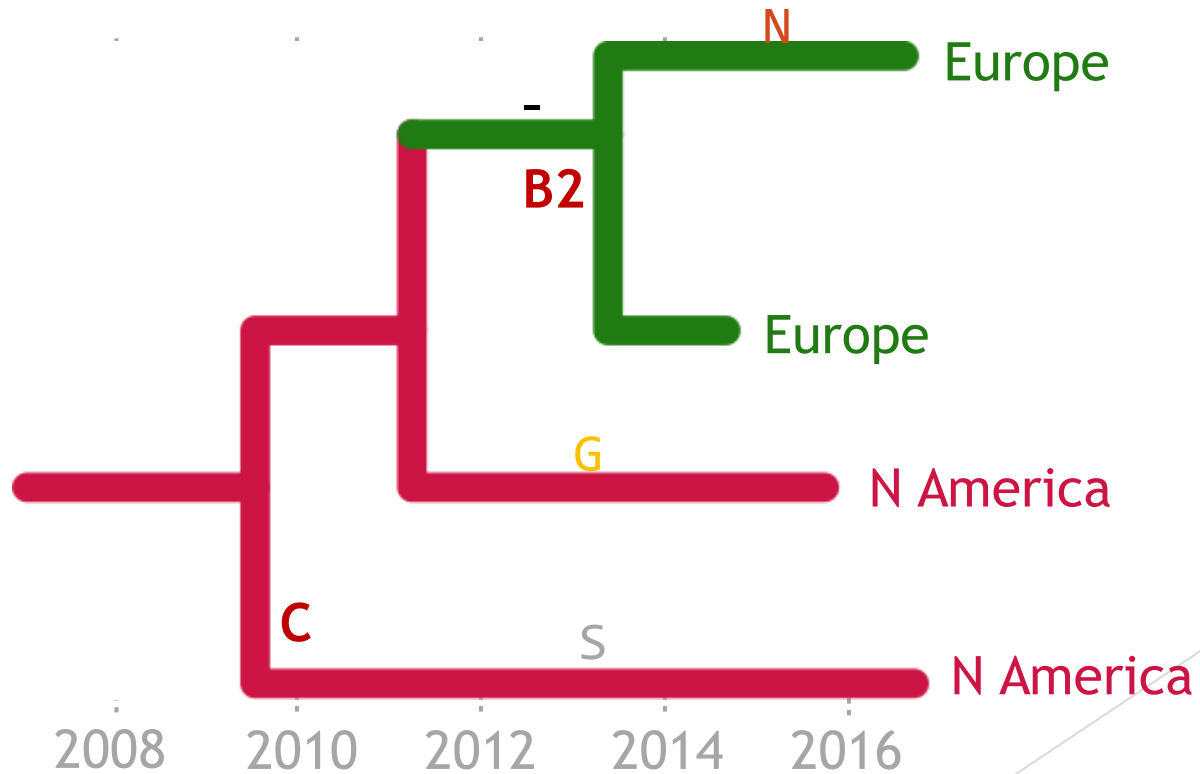
```



- ▶ Filter
- ▶ Align
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

```

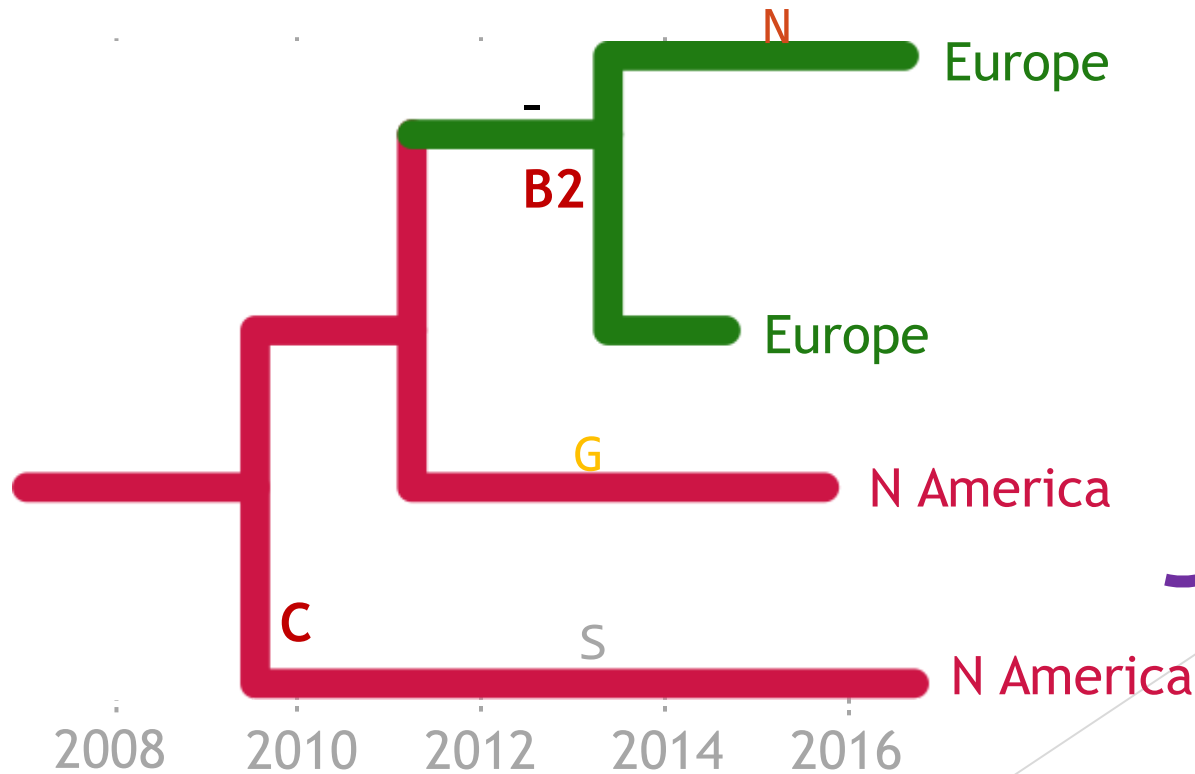
>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN XAV-GSNNX
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN VAV-GSSXX
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT XXVNGSSNN
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN XAVNSSGXX
>NODE_03 GTAGCAGTA---GGTAGCAGTAACAAT VAV-GSSNN
>NODE_02 GTAGCAGTGAATGGTAGTAGTAACAAT VAVNGSSNN
>NODE_01 GTAGCAGTGAATGGTAGTAGTAATAAT VAVNGSSNN
  
```



- ▶ Filter
- ▶ Align
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ Clades
- ▶ Traits
- ▶ Export

```

>SWE-2016 NNAGCAGTA---GGTAGCAATAACNNN XAV-GSNNX
>ESP-2014 GTAGCAGTA---GGTAGCAGTNNNNNN VAV-GSSXX
>USA-2016 NNNNNAGTGAATGGTAGTAGTAATAAT XXVNGSSNN
>CAN-2015 NNNGCAGTGAATAGTAGTGGTNNNNNN XAVNSSGXX
>NODE_03 GTAGCAGTA---GGTAGCAGTAACAAT VAV-GSSNN
>NODE_02 GTAGCAGTGAATGGTAGTAGTAACAAT VAVNGSSNN
>NODE_01 GTAGCAGTGAATGGTAGTAGTAATAAT VAVNGSSNN
  
```



evd68_tree.json
evd68_meta.json

TB workflow

- ▶ Filter
- ▶ Mask
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ DRM
- ▶ Traits
- ▶ Export

>TBRef	GTAGCAGTGAATGGCAGTAGTAATAATAATGTG
>S-1078	GNAGCAGTAAATGGTAGCAATACTAATAATCTG
>P-2283	GTAGCAGTAAATGGTAGCAGTAATAATAATGTG
>S-1002	GTTGGAGTGAATGGTAGTAGTAATAATAATGTG
>L-0310	GTAGCAGTGAATAGTAGTGGTAATCCTAATGTC

TB workflow

- ▶ Filter
- ▶ Mask
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ DRM
- ▶ Traits
- ▶ Export

>TBRef	GTAGCAGTGAATGGCAGTAGTAATAATAATGTG
>S-1078	GNAGCAGTAAATGGTAGCAATAATAATAATCTG
>P-2283	GTAGCAGTAAATGGTAGCAGTAATAATAATGTG
>S-1002	GTTGGAGTGAATGGTAGTAGTAATAATAATGTG
>L-0310	GTAGCAGTGAATAGTAGTGGTAATAATAATGTC

TB workflow

- ▶ Filter
- ▶ Mask
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ DRM
- ▶ Traits
- ▶ Export

```
>TRef  GTAGCAGTGAATGGCAGTAGTAATAATAATGTG
>S-1078 GNAGCAGTAAATGGTAGCAATAATAATAATCTG
>P-2283 GTAGCAGTAAATGGTAGCAGTAATAATAATGTG
>S-1002 GTTGGAGTGAATGGTAGTAGTAATAATAATGTG
>L-0310 GTAGCAGTGAATAGTAGTGGTAATAATAATGTC
```



Get topology



Adjust branch-lengths

TB workflow

- ▶ Filter
- ▶ Mask
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ DRM
- ▶ Traits
- ▶ Export

>TBRef	GTAGCAGTGAATGGCAGTAGTAATAATAATGTG
>S-1078	GNAGCAGTAAATGGTAGCAATAATAATAATCTG
>P-2283	GTAGCAGTAAATGGTAGCAGTAATAATAATGTG
>S-1002	GTTGGAGTGAATGGTAGTAGTAATAATAATGTG
>L-0310	GTAGCAGTGAATAGTAGTGGTAATAATAATGTC
	G L
	G V
	G V
	S V

TB workflow

- ▶ Filter
- ▶ Mask
- ▶ Tree
- ▶ Refine
- ▶ Ancestral
- ▶ Translate
- ▶ SeqTraits
- ▶ Traits
- ▶ Export

```
>TBRef      GTAGCAGTGAATGGCAGTAGTAATAATAATGTG
>S-1078     GNAGCAGTAAATGGTAGCAATAATAATAATCTG
>P-2283     GTAGCAGTAAATGGTAGCAGTAATAATAATGTG
>S-1002     GTTGGAGTGAATGGTAGTAGTAATAATAATGTG
>L-0310     GTAGCAGTGAATAGTAGTGGTAATAATAATGTC
```

GENE	SITE	ALT	DISPLAY_NAME	FEATURE
rpoB	170	F		Rifampicin
katG	315	T		Isoniazid
nuc	1472358	T	rrs:C513T	Streptomycin

(Here: drug resistance mutations)