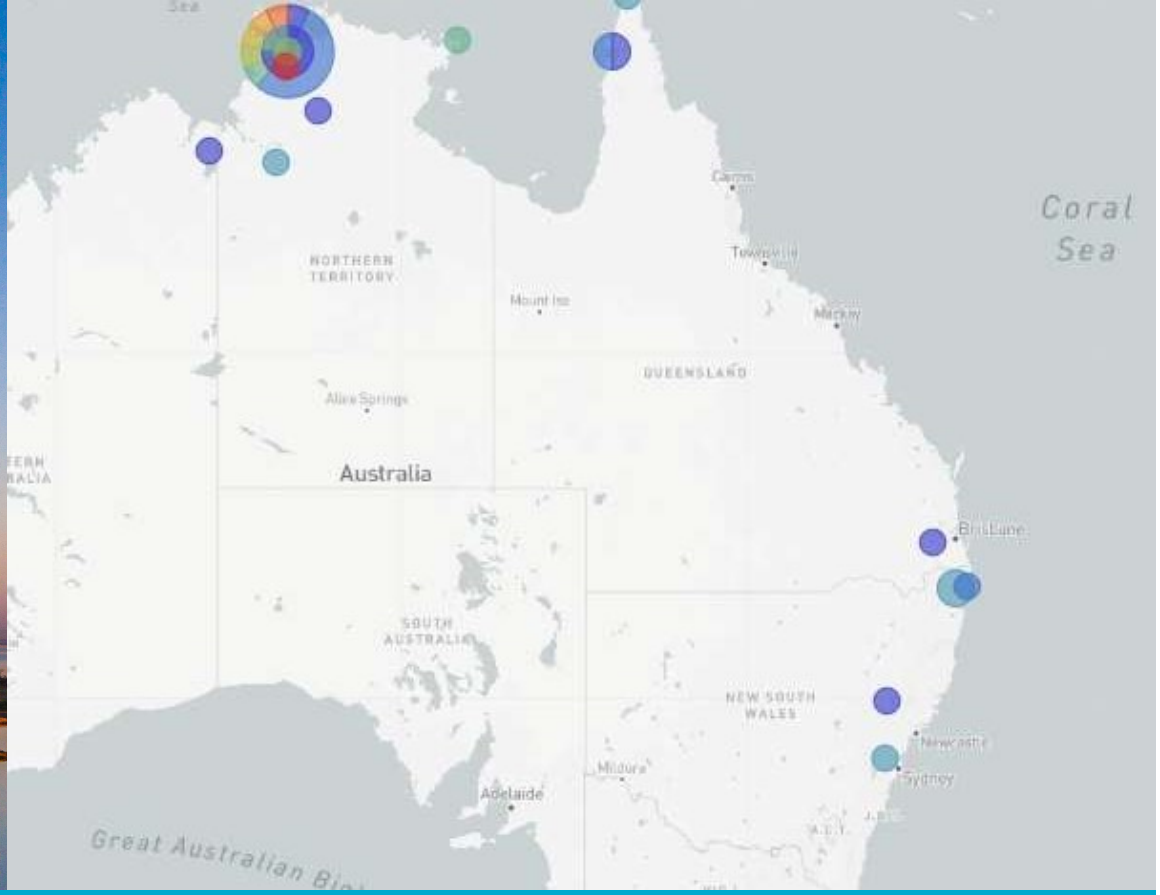


ACDP update on Whole Genome Sequencing and Nextstrain

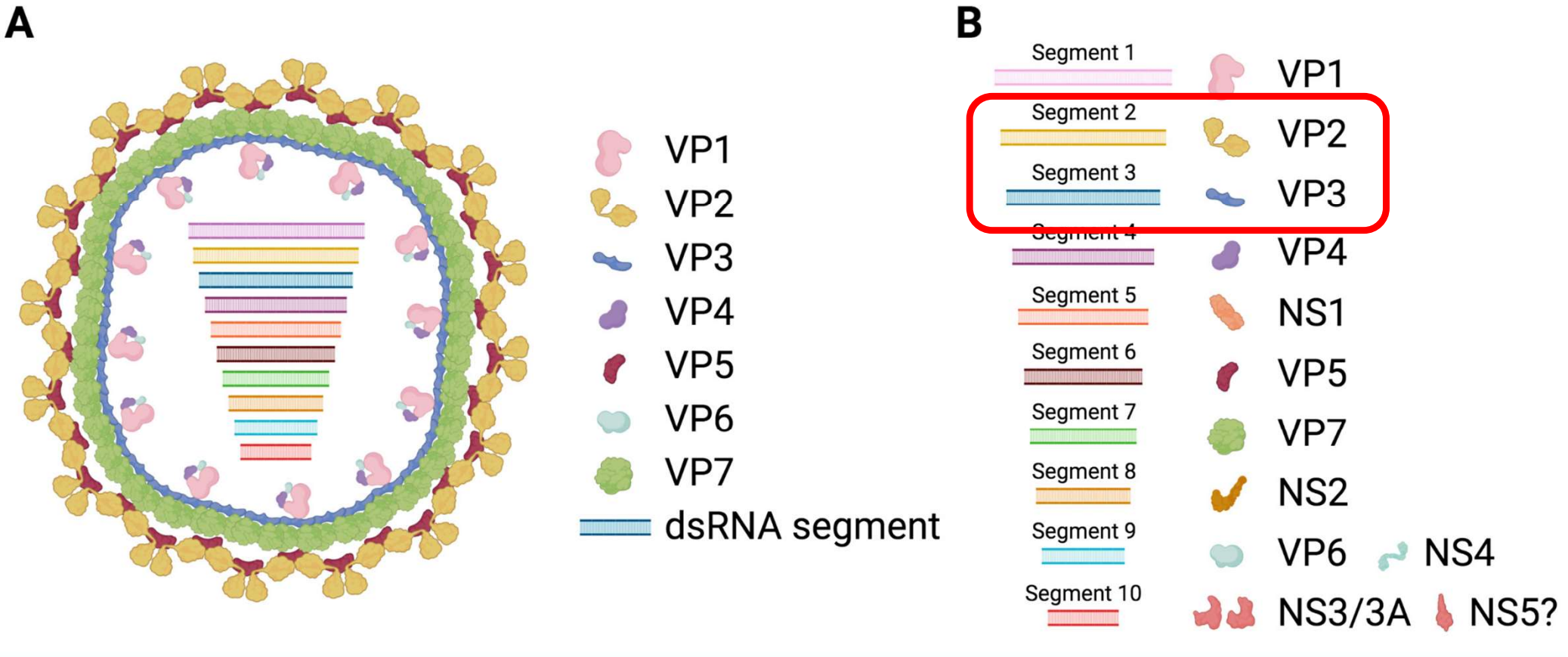
CSIRO Australian Centre for Disease Preparedness

Tristan Reid, Matt Neave & John White

NAMP Technical Committee –Annual Meeting August 24, 2023

AUSTRALIAN ANIMAL HEALTH LABORATORY
www.csiro.au

Bluetongue Virus



Rojas, J.M.; Martín, V.; Sevilla, N. Vaccination as a Strategy to Prevent Bluetongue Virus Vertical Transmission. *Pathogens* **2021**, *10*, 1528. https://doi.org/10.3390/pathogens10111528

An update on Whole Genome Sequencing and Nextstrain

- ACDP has undertaken a project to develop a novel method of characterisation of the genetic relationships between BTV isolates, at the whole genome level.

- Aimed to overcome the limitations of the current BTV characterisation methods
  - Current reliance on 2 segments (VP2, VP3) alone

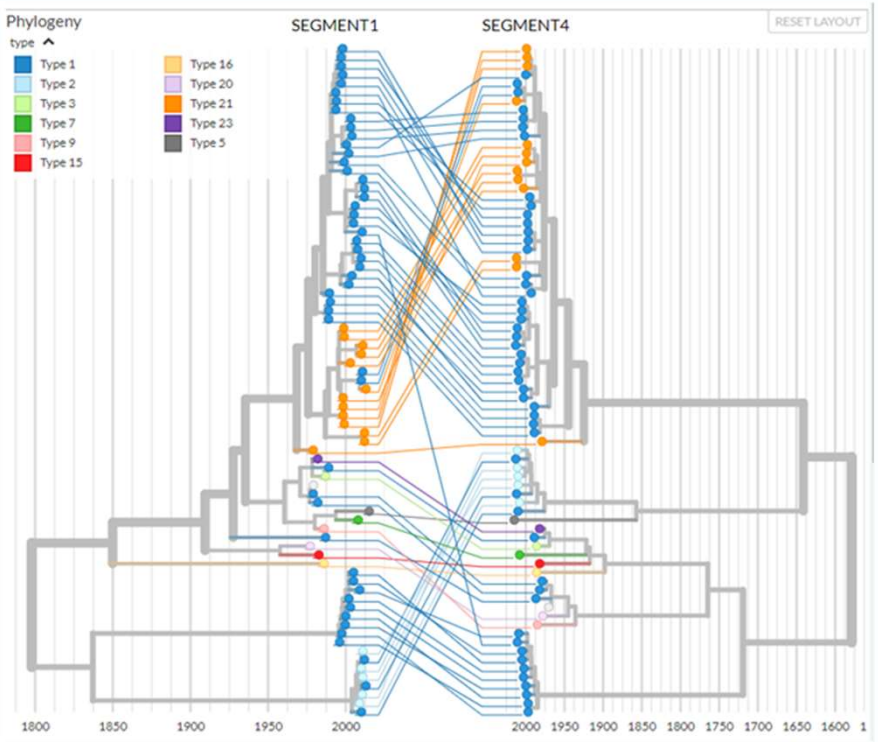
  → Poorly sensitive for the detection of potentially significant changes in other segments, recombination events, etc

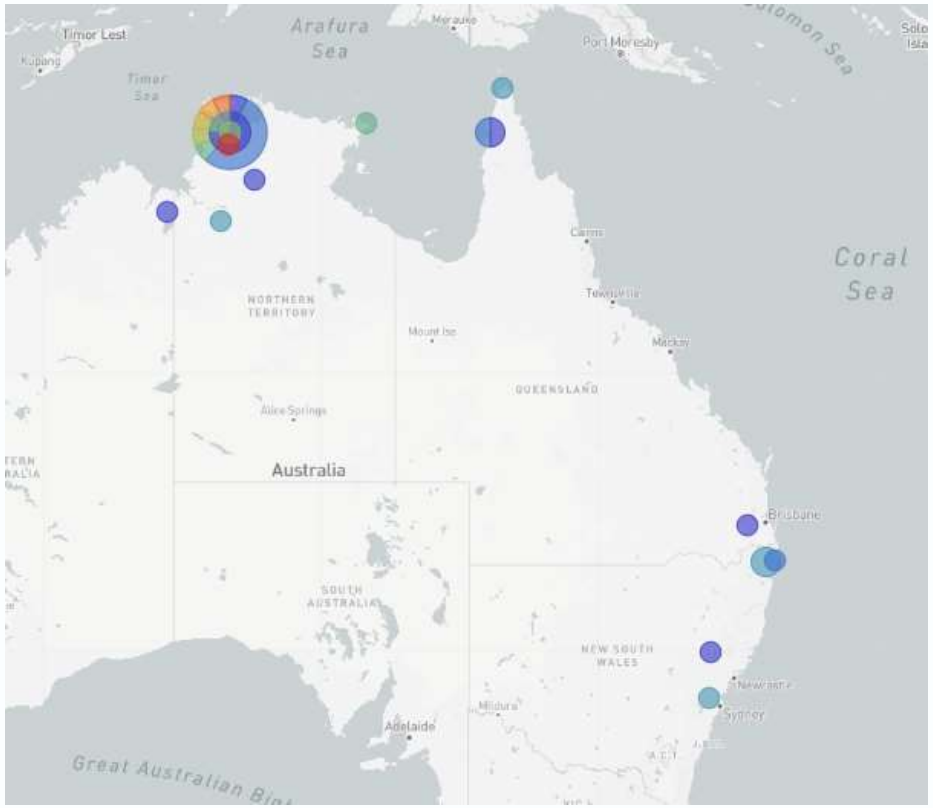
- Whole genome analysis complicated by the to segmented nature of the virus

  BUT has the potential to greatly improve our understanding of the molecular epidemiology of BTV, E.g.

  - Improved understanding of phylogenetic relationships between isolates to understand changes in virus transmission
  - Improved ability to detect:
    - recombination events
    - incursion & spread of exotic gene segments

Nextstrain


Phylogeny
type
Type 1
Type 2
Type 3
Type 7
Type 9
Type 15
Type 16
Type 20
Type 21
Type 23
Type 5
SEGMENT1
SEGMENT4
RESET LAYOUT
1800 1850 1900 1950 2000
2000 1950 1900 1850 1800 1750 1700 1650 1600

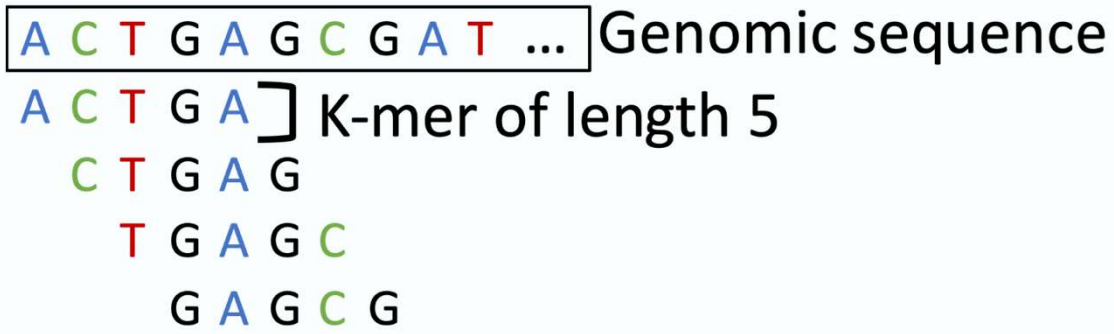

Arafura Sea
Timor Leste
Kupang
Timor Sea
Merauke
Port Moresby
Coral Sea
NORTHERN TERRITORY
Mount Isa
Cairns
Townsville
Mackay
QUEENSLAND
Alice Springs
Australia
Brisbane
SOUTH AUSTRALIA
NEW SOUTH WALES
Newcastle
Sydney
Mildura
Adelaide
Great Australian Bight

An update on Whole Genome Sequencing and Nextstrain

- BTV Reassortment is not completely random – some segments have a greater tendency to reassort together

- Developed a Machine Learning model (Support Vector Machine, SVM) that assigns isolates into clusters based on
  - 1-3-4-9 (core)
  - 5+8+10 (non-structural)
  - 2-6-7 (capsid)

- Dataset? WGS from 880 BTV Isolates Worldwide
  - 276 Eastern Hemisphere
  - 125 from Australia



An update on Whole Genome Sequencing and Nextstrain

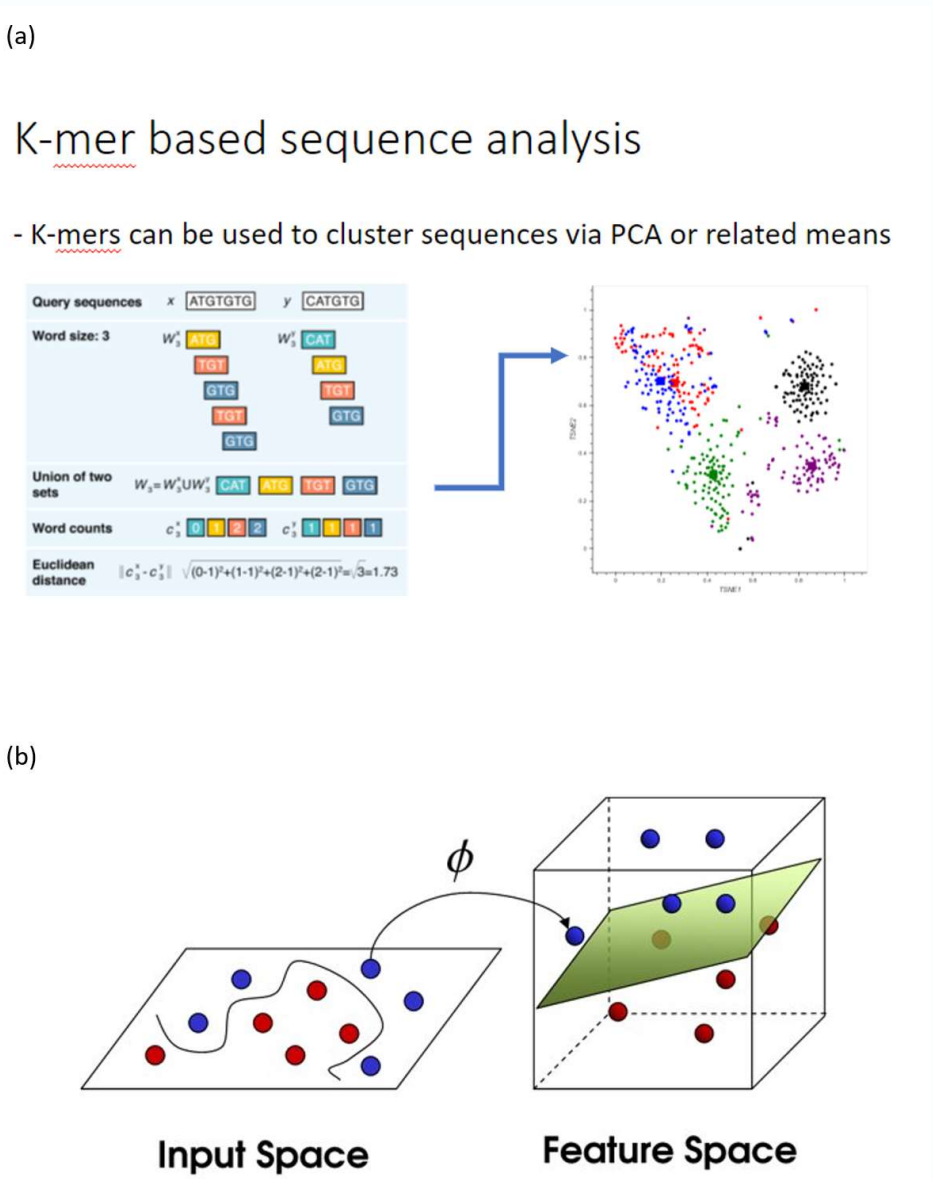


Figure 1. Schematic representations of the sequence data input approach for BTV-ML model analysis. Overlapping K-mer oligonucleotides for each concatemer provided individual feature points enabling the creation of a genetic signature which could be converted to a specific point in space **(a)**.

The multidimensional nature of the signature placement enables clear partitioning between points that appear close to each other in two-dimensional space **(b).**

Ouput

- The BTV-ML Model allows assignment of a WG characterisation of BTV based on assessment of 3 groups of genes

- Relationships between isolates in a cluster can be visualised and examined in Nextstrain to determine secondary clade designations

- A seamless workflow pipeline enables us to rapidly assign serotype, topotype, clade and sub-clade designations to inputted test isolate WGS data.

Detection of significant changes in transmission patterns?

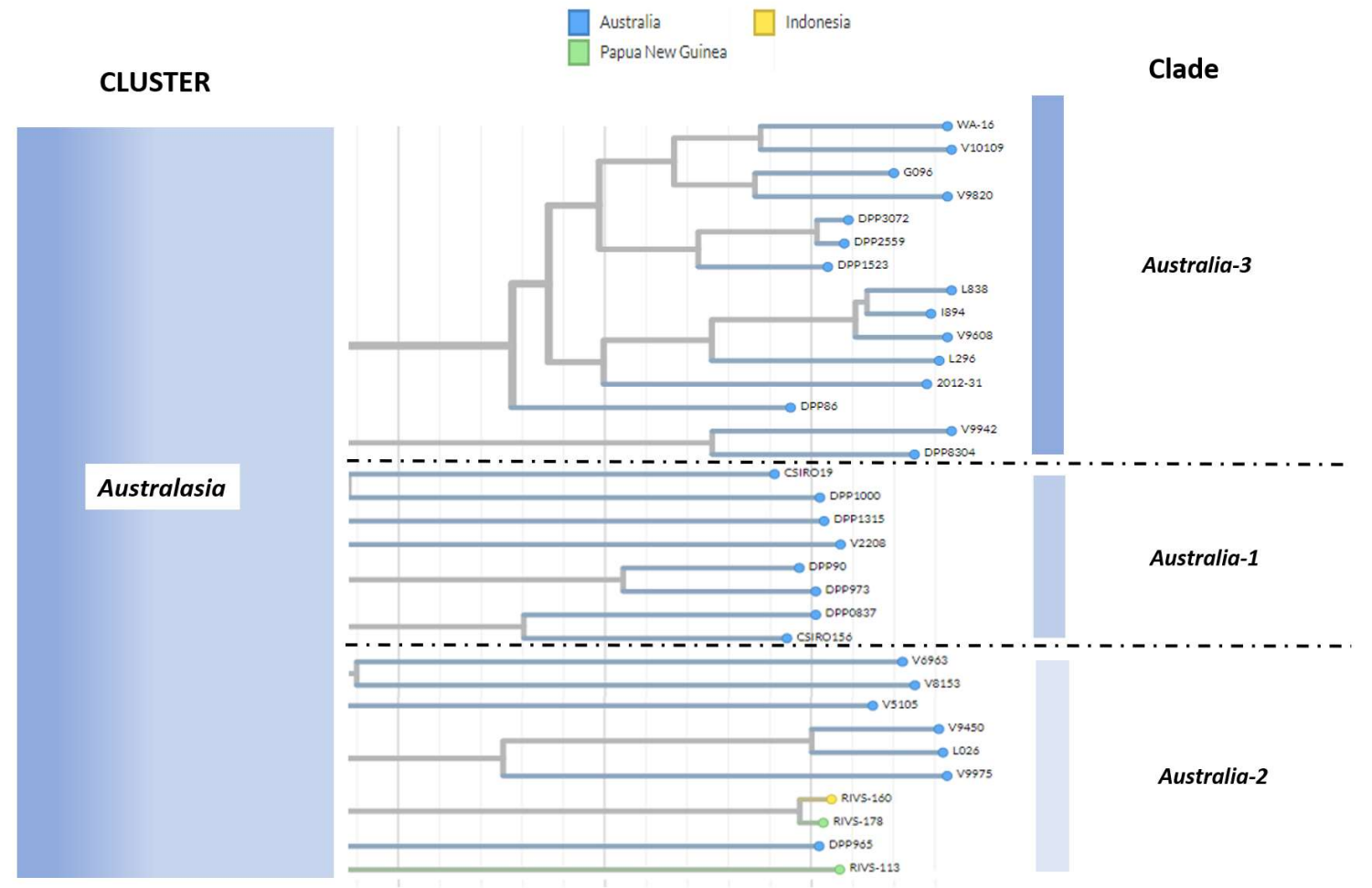
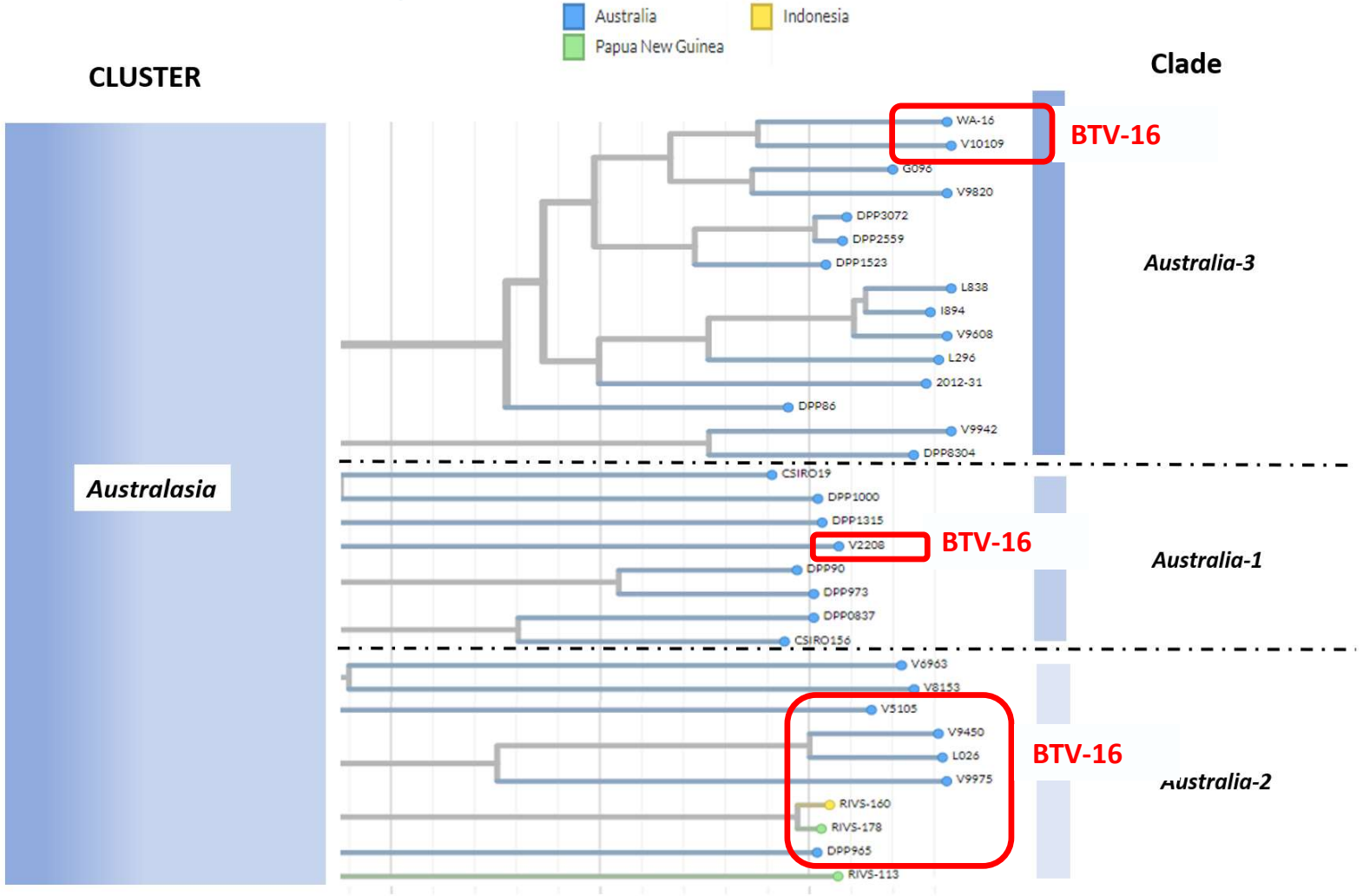


Figure 2. Nextstrain phylogenetic analysis of the Core (Seg 1, 3, 4, 9) concatemer for isolates placed in the Australasian cluster revealed three clades. Clades Australia 1 and 2 contained genotypes normally restricted to the north of Australia whilst the Australia 3 clade described a single genotype found throughout Australia.

Detection of significant changes in transmission patterns?



If we were to see Clade 1 or 2 viruses emerge in Eastern Australia, we would know there has been a significant change in BTV transmission.

(Even if is 'just another BTV-16')

Figure 2. Nextstrain phylogenetic analysis of the Core (Seg 1, 3, 4, 9) concatemer for isolates placed in the Australasian cluster revealed three clades. Clades Australia 1 and 2 contained genotypes normally restricted to the north of Australia whilst the Australia 3 clade described a single genotype found throughout Australia.

Recognising new incursions

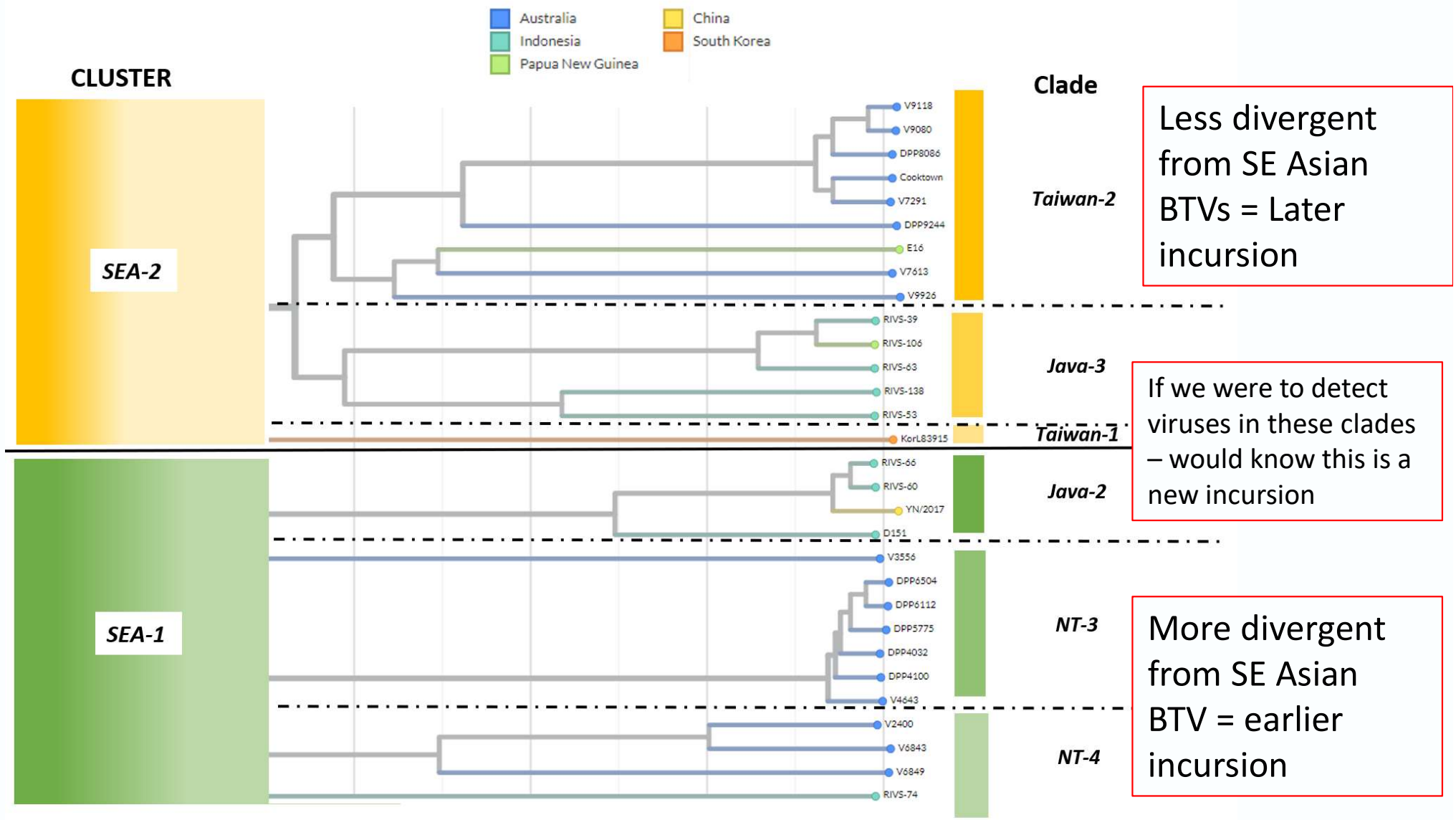


Figure 3. Nextstrain phylogenetic analysis of the Core (Seg 1, 3, 4, 9) concatemer for isolates placed in the Southeast Asia 1 and 2 clusters revealed three clades within each cluster. The cluster SEA-1 contained genotypes entering and remaining restricted to northern Australia in more recent times (from early 1990's). A second later incursion of unique genotypes is evident in the isolates contained in the SEA-2 cluster.

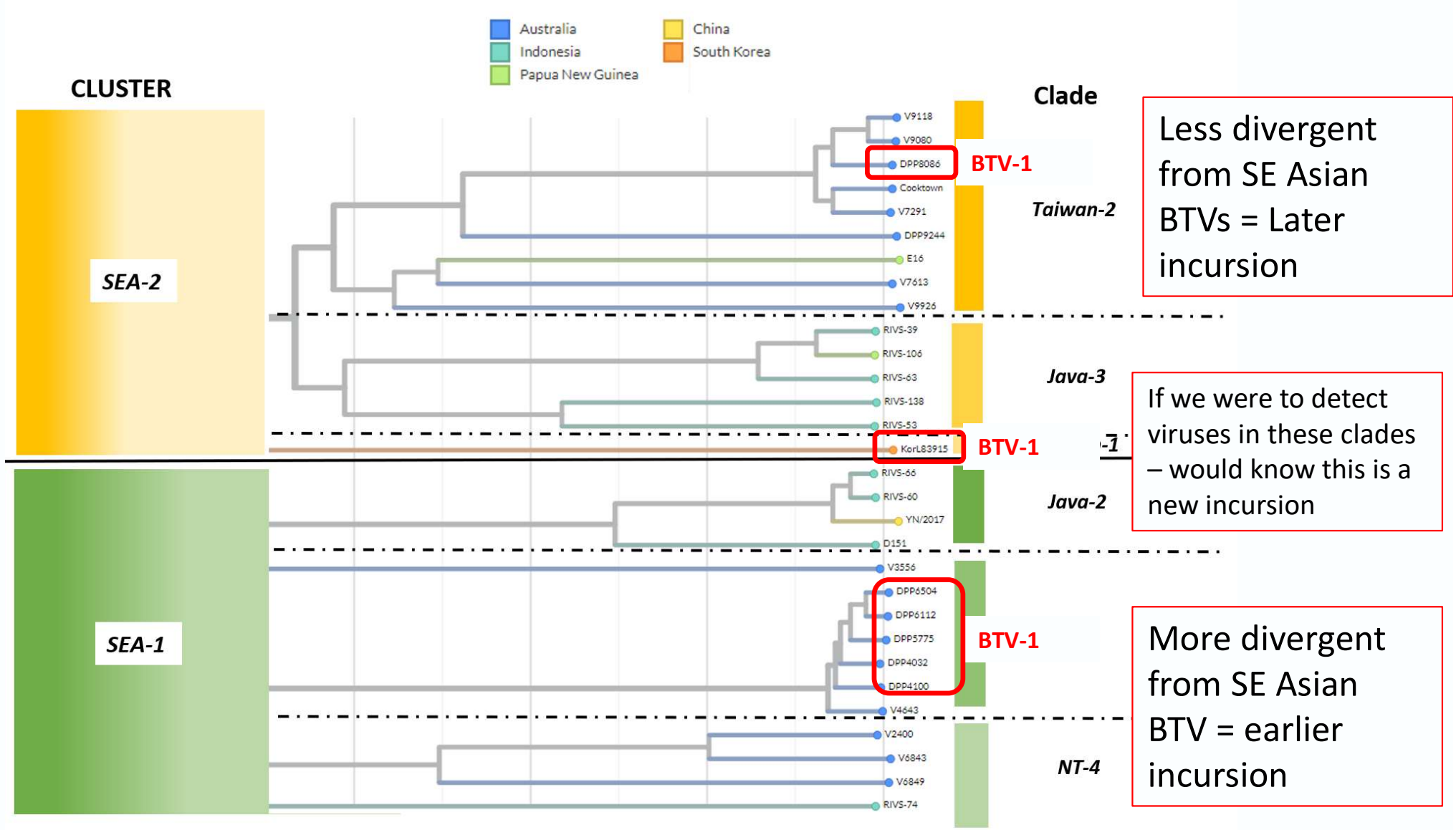


Figure 3. Nextstrain phylogenetic analysis of the Core (Seg 1, 3, 4, 9) concatemer for isolates placed in the Southeast Asia 1 and 2 clusters revealed three clades within each cluster. The cluster SEA-1 contained genotypes entering and remaining restricted to northern Australia in more recent times (from early 1990's). A second later incursion of unique genotypes is evident in the isolates contained in the SEA-2 cluster.

An update on Whole Genome Sequencing and Nextstrain

- WGS limited by requirement for high titre isolates – currently not applicable to field samples
- % isolate recovery is low, isolation is slow and resource intensive

- ACDP is continuing to investigate the use of myBaits target capture kits for probe-enriched next-generation sequencing of BTV samples

→ACDP has had good success in generating WGS data from blood samples with pan-BTV TQM Cts of <28

How can NAMP help?

- The database and analysis is only as strong as the baseline data
  - Inclusion of as many WGS as possible in the dataset will improve accuracy and value of the tool

- Greatly aided with recent success with WGS of low Ct value blood samples

- Please continue to submit isolates and low Ct samples for WGS ☺

Future?

- Details of methods, proposed nomenclature to appear in a future manuscript
  → Will be circulated to NAMP for comment
- Subject to peer review, aim to report on WGS characterisation of isolates received in the future
- Annual/biannual reports to NAMP highlighting potentially significant events?
- NAMP stakeholder login to Nextstrain
- Further research potential? E.g. identification of virulence factors

An update on Whole Genome Sequencing and Nextstrain

ACDP would like to thank all NAMP submitters who have assisted in providing isolates and PCR positive samples for use in these activities.