

# Development of CRISPR diagnostics for *Phytophthora*: Computational tools and approaches to mine whole genomes sequences

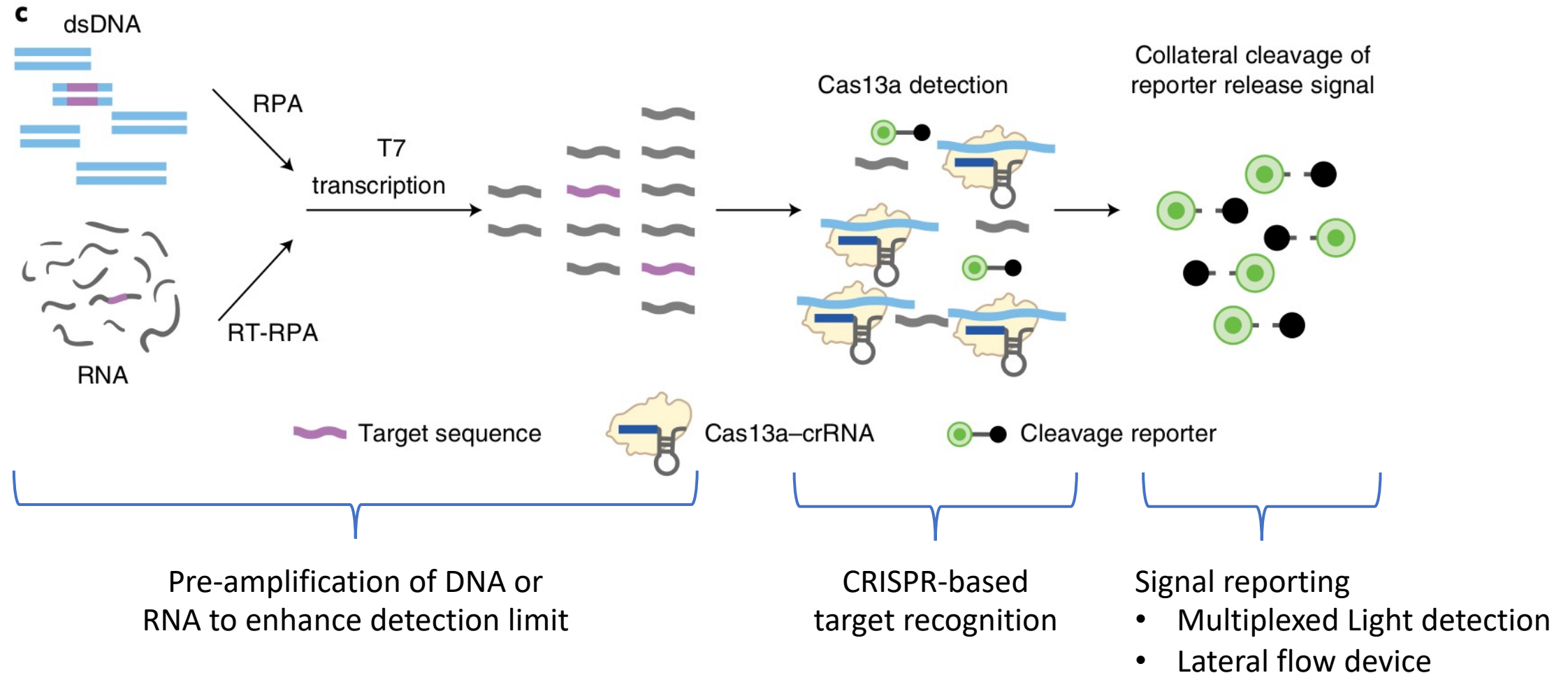
Camilo Parada-Rojas, Caroline Press, Zach Foster,  
Andrew Tupper, Riley Buchanan, Valerie Fieland,  
Nick Cauldron, Jeff Chang, and Nik Grünwald

Horticultural Crops Disease and Pest Management Research Unit,  
USDA Agricultural Research Service  
Corvallis, OR, USA

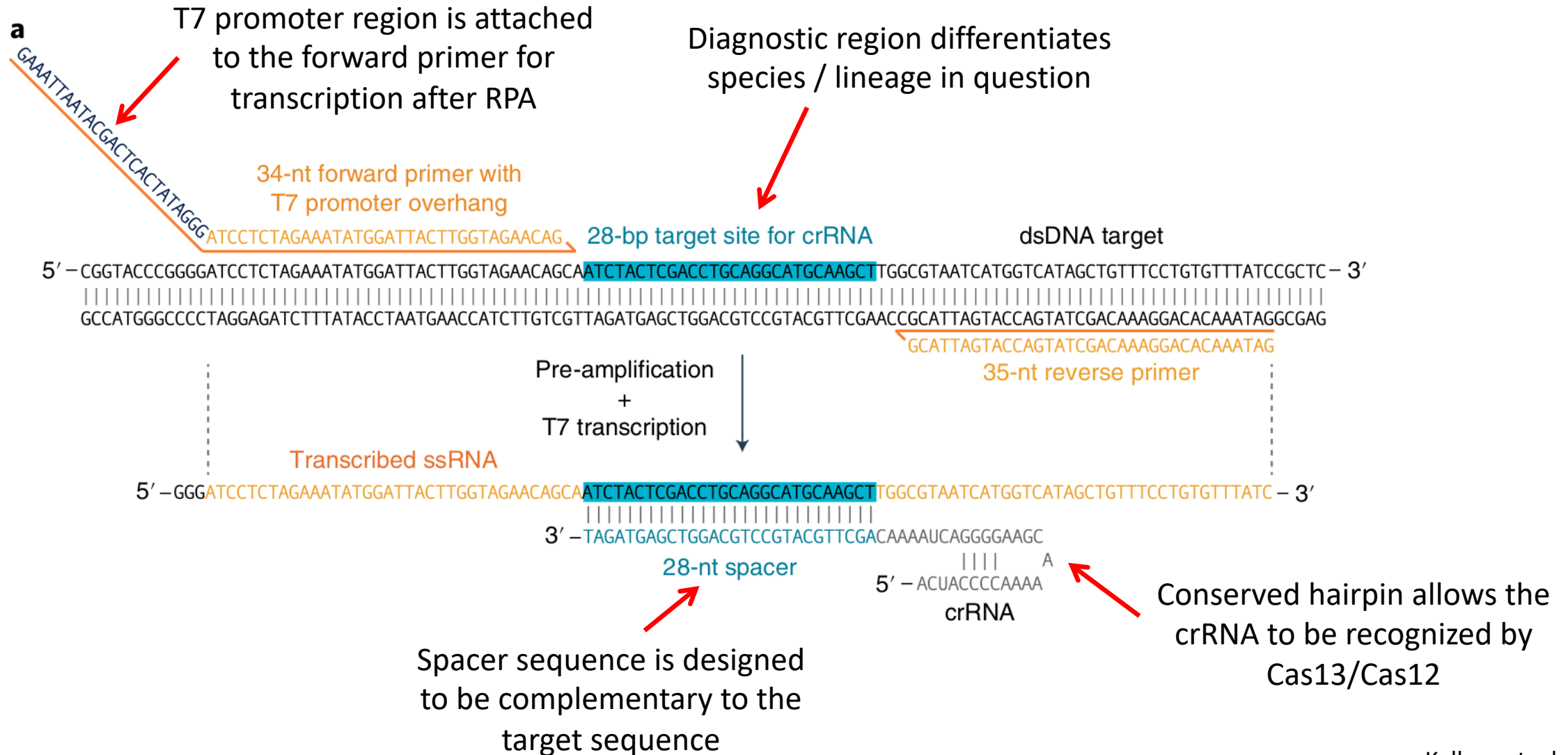




# Basics of SHERLOCK (Kellner et al. 2019)



# CRISPR-Cas Dx pipeline



# **krisp** facilitates development of CRISPR-Dx guides and primers

# krisp facilitates development of CRISPR-Dx spacers and primers

krisp\_vcf

krisp\_fasta

# *krisp*: A tool to find diagnostic markers for SHERLOCK

## Simple to install (Available upon request):

- Github
- PyPI (Python – pip repository) -> pip install krisp
- Command line tools automatically installed
- Free and open source license (MIT)

## Simple to use:

- Specify FASTA/VCF files and run: `$ krisp lineage1.fasta lineage2.fasta ...`
- Optimal primers pairs suggested for each potential marker found
- Results are clear: No complex algorithms, weighting function, or magic

## Fast:

- Whole genomes: ~10-15 minutes (5 lineages @ ~60 Mbp / genome using 4 cores)
- Population SNP data: ~2 hours (~350 isolates on a laptop)
- Written in parallel: compute time decreases with number of cores available

New Results

[Follow this preprint](#)

**Krisp: A python package for designing CRISPR and amplification-based diagnostic assays from whole genome data**

[Zachary S. L. Foster](#), [Andrew S. Tupper](#), [Caroline M. Press](#), [Niklaus J. Grünwald](#)

doi: <https://doi.org/10.1101/2023.11.16.567433>

This article is a preprint and has not been certified by peer review [what does this mean?].



Andrew Tupper  
Zach Foster

# Examples

```
krisp_fasta cerevisiae/*.fna.gz --outgroup kudriavzevii/*.fna.gz --
conserved 50 --diagnostic 10 --out_align align.txt --out_csv out.csv --
primer3
```

```
TGCAAGTTAATTGGAACGGAAGCACC...TTGTCAACTTGAAC...AGATGAAATCTTACCTTCTTGACCCTT : SC1;SC2;SC3;SC4;SC5;SC6
TGCAAGTTAATTGGAACGGAAGCACC...TTGTCAACTTGGAC...AGATGAAATCTTACCTTCTTGACCCTT : SK1;SK2;SK3;SK4;SK5;SK6
┌──────────Forward──────────┐ ... {-----#--#} ... ┌──────────Reverse──────────┐
```

Primer statistics:

Direction	Penalty	Sequence	Tm	Gc Percent	Self Any Th	Self End Th	...
Forward	7.22384	TGCAAGTTAATTGGAACGGAAGCAC	62.72384	44.0	0.0	0.0	...
Reverse	6.91895	AGGGTCAAGAAGGTAAGATTTTCATC	58.58105	40.0	0.0	0.0	...

Pair statistics:

Penalty	Compl Any Th	Compl End Th	Product Size	Product Tm	Product Tm Oligo Tm Diff	...
14.14278	0.0	0.0	109	79.88148	21.30042	...

```
krisp_vcf metadata.csv reference.fasta --vcf variants.vcf --
groups NA1 NA2 EU1 --out_align alignments.txt --gc_clamp 2
```

```
## Phyram_PR-102_s0001:209731-209856 is diagnostic for NA2
```

```
Reference : gtcccgtcaccgtatatatgtactaaacgca...gtggtagcatactgacgacgagaagt C agt...ctggtaggacagttaaatgtaccGagag
EU1 : ..... C9 .....R....
NA1 : ..... C6 .....
NA2 : .....<T11>.....
oligos : gtcccgtcaccgtatatatgtactaaacg ... ggtagcatactgacgacgagaagt T ag ... ggtaggacagttaaatgtaccgagag
┌──────────Left primer──────────┐ ... ┌──────────crRNA──────────┐ ... ┌──────────Right primer──────────┐
```

Primer statistics:

Direction	Penalty	Sequence	Tm	Gc Percent	Self Any Th	Self End Th	Hairpin Th	...
Forward	3.42431	gtcccgtcaccgtatatatgtactaaacg	62.92431	44.82759	7.87208	10.60241	0.0	...
Reverse	3.60529	ctctcggtacatttaactgtcctaacc	61.10529	44.44444	1.18023	3.98189	0.0	...

Pair statistics:

Penalty	Compl Any Th	Compl End Th	Product Size	Product Tm	Product Tm Oligo Tm Diff	T Opt A	...
7.0296	2.27423	0.0	126	83.22136	22.11607	61.68654	...

New Results

[Follow this preprint](#)

**Krisp: A python package for designing CRISPR and amplification-based diagnostic assays from whole genome data**

© Zachary S.L. Foster, © Andrew S. Tupper, © Caroline M. Press, © Niklaus J. Grünwald

doi: <https://doi.org/10.1101/2023.11.16.567433>

This article is a preprint and has not been certified by peer review [what does this mean?].





# *krisp*: Candidate regions displayed as alignments

>Alignment 0:

```

aaaaaaagaaaaagaaaccagaacgcgacaCCTGATTGCCAAGCAGCGAAATGGAGCAGATACTGCAGCGAGGCTTGGTGTCTGAGGACCcacgtttgcccctcgacgatctccatagt <- PR-102
aaaaaaagaaaaagaaaccagaacgcgacaCTTGATTGCCAAGCAGCGAAATGGAGCAGATACTGCAGCGAGACTTGGTGTCTGAGGACCcacgtttgcccctcgacgatctccatagt <- PR-15-019
aaaaaaagaaaaagaaaccagaacgcgacaCCTGATTGCCAAGCAGCGAAATGGAGCAGATACTGCAGCGAGACTTGGTGTCTGAGGACCcacgtttgcccctcgacgatctccatagt <- PR-18-108, PR-18-126, PR-18-069
      {-*-----*-----}
  
```

>Alignment 1:

```

aaaaaacaatgtccatcgttcgcggtagttCGGGATTGGTTGCATATCAATTTCTGTTGCATCCACACTGGCAAGCATAACCGTGACCCAccctccgttgtggtcatttcaaagaatgt <- PR-102
aaaaaacaatgtccatcgttcgcggtagttCGGGATTGGTTGCATATCAATTTCTGTTGCATCCACACTGGCAAGCATAATCGTGATCCAccctccgttgtggtcatttcaaagaatgt <- PR-15-019
aaaaaacaatgtccatcgttcgcggtagttCGGGATTGGTTGCATATCAATTTCTGTTGCATCCACACTGGCAAGCATAATCGTGACCCAccctccgttgtggtcatttcaaagaatgt <- PR-18-108, PR-18-126, PR-18-069
      {-----*-----*----}
  
```

>Alignment 2:

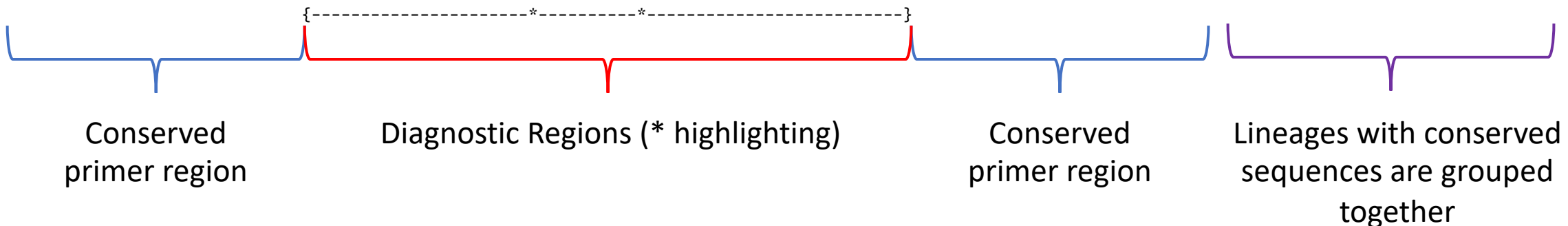
```

aaaaaaagaaaaagaaaccagaacgcgacacCTGATTGCCAAGCAGCGAAATGGAGCAGATACTGCAGCGAGGCTTGGTGTCTGAGGACCCcacgtttgcccctcgacgatctccatagtg <- PR-102
aaaaaaagaaaaagaaaccagaacgcgacacTTGATTGCCAAGCAGCGAAATGGAGCAGATACTGCAGCGAGACTTGGTGTCTGAGGACCCcacgtttgcccctcgacgatctccatagtg <- PR-15-019
aaaaaaagaaaaagaaaccagaacgcgacacCTGATTGCCAAGCAGCGAAATGGAGCAGATACTGCAGCGAGACTTGGTGTCTGAGGACCCcacgtttgcccctcgacgatctccatagtg <- PR-18-108, PR-18-126, PR-18-069
      {*-----*-----}
  
```

>Alignment 3:

```

aaaaaaagagcacactgtgattggtcaatccACTTCAAGTTGATGTGGATCGACTATACGTGACGGTAGAGTAGTTGCCAAAGGTAGCGCGccatgagcgcgaatcgctgtcgctgtcgcg <- PR-102
aaaaaaagagcacactgtgattggtcaatccACTTCAAGTTGATGTGGATCGATTATACGTGACTGTAGAGTAGTTGCCAAAGGTAGCGCGccatgagcgcgaatcgctgtcgctgtcgcg <- PR-15-019
aaaaaaagagcacactgtgattggtcaatccACTTCAAGTTGATGTGGATCGACTATACGTGACTGTAGAGTAGTTGCCAAAGGTAGCGCGccatgagcgcgaatcgctgtcgctgtcgcg <- PR-18-108, PR-18-126, PR-18-069
      {-----*-----*-----}
  
```



*Phytophthora ramorum*  
Sudden Oak Death

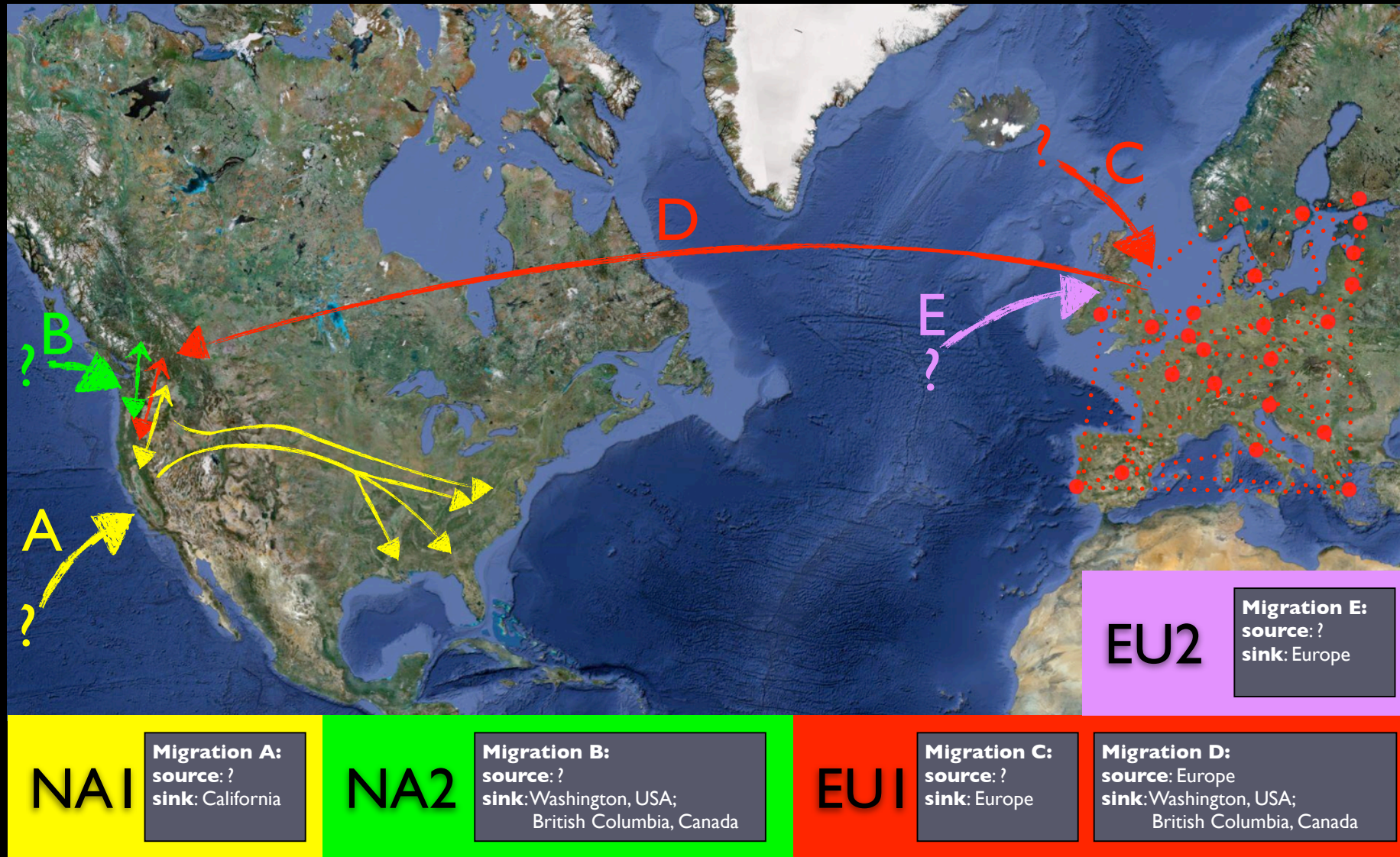


Coast Live Oak, Marin County CA, 2000

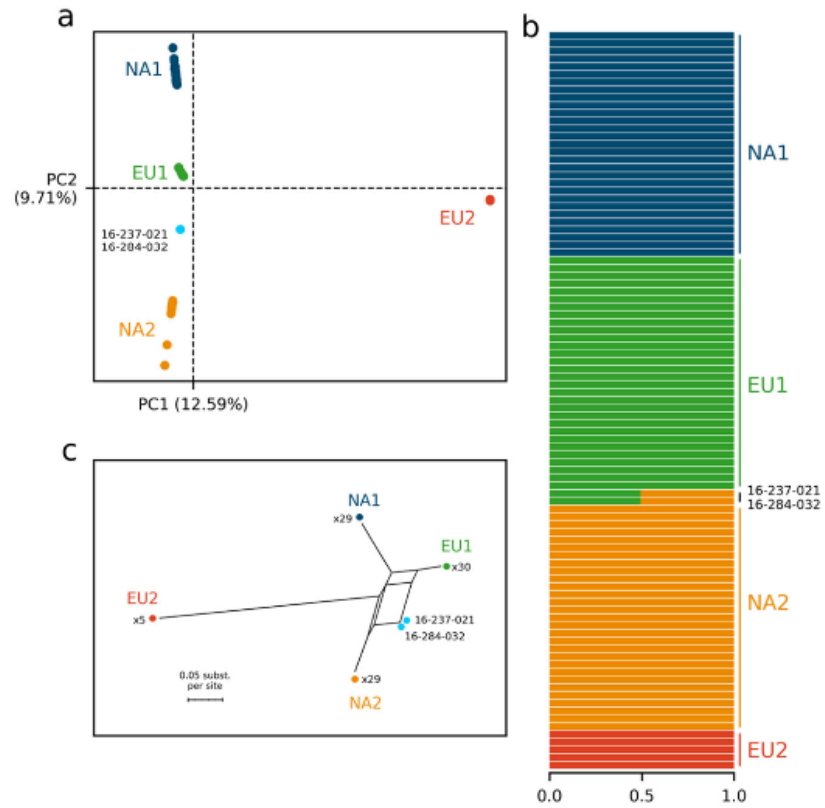


Tan oak, Marin County CA, 2000

# 5 migrations



# Detection of first sexual variants



## communications biology

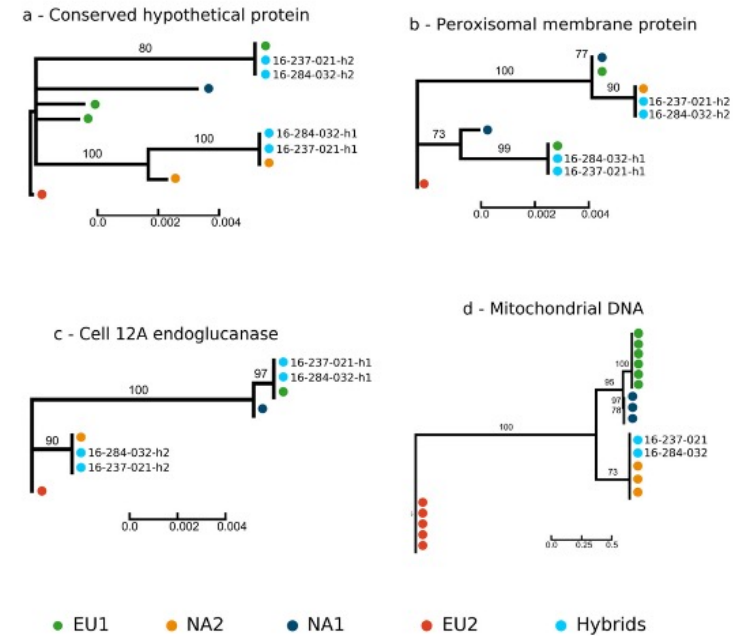
ARTICLE

[Check for updates](#)

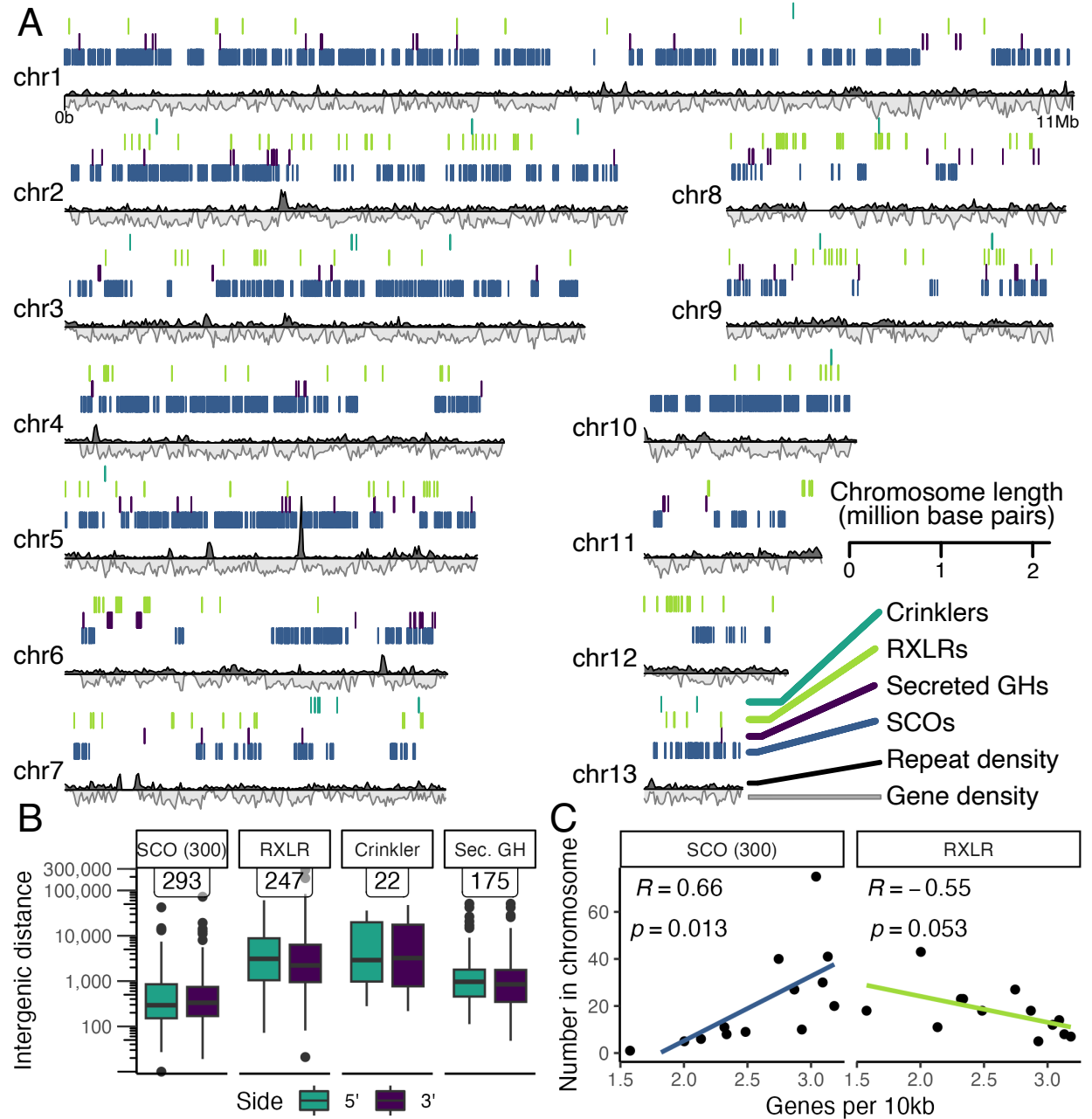
<https://doi.org/10.1038/s42003-022-03394-w> OPEN

## Genomic biosurveillance detects a sexual hybrid in the sudden oak death pathogen

Richard C. Hamelin<sup>1</sup>, Guillaume J. Bildeau<sup>2</sup>, Renate Heinzelmann<sup>1,3</sup>, Kelly Hrykiw<sup>1</sup>, Arnaud Capron<sup>1</sup>, Erika Dort<sup>1</sup>, Angela L. Dale<sup>4</sup>, Emilie Giroux<sup>2</sup>, Stacey Kus<sup>4</sup>, Nick C. Carleson<sup>5</sup>, Niklaus J. Grünwald<sup>5,6</sup> & Nicolas Feau<sup>1</sup>



# *P. ramorum* gene content



# Genomes available: Nextstrain implementation

**Auspice**

Dataset: nextstrainoomy

Date Range: 609-09-13 to 2019-01-05

Color By: Lineage

Filter Data: Type filter query here...

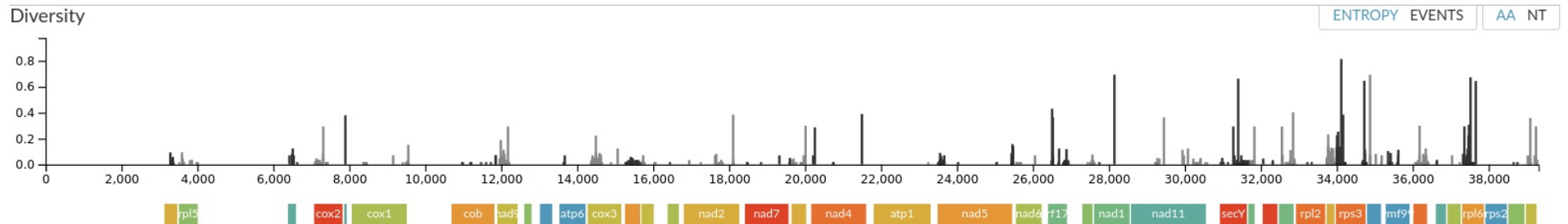
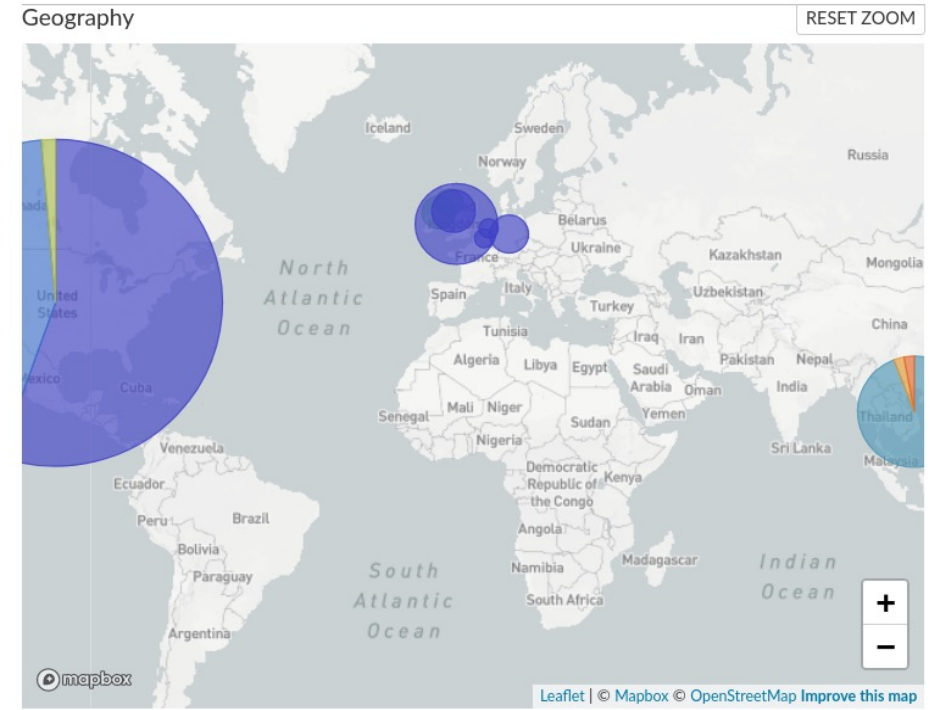
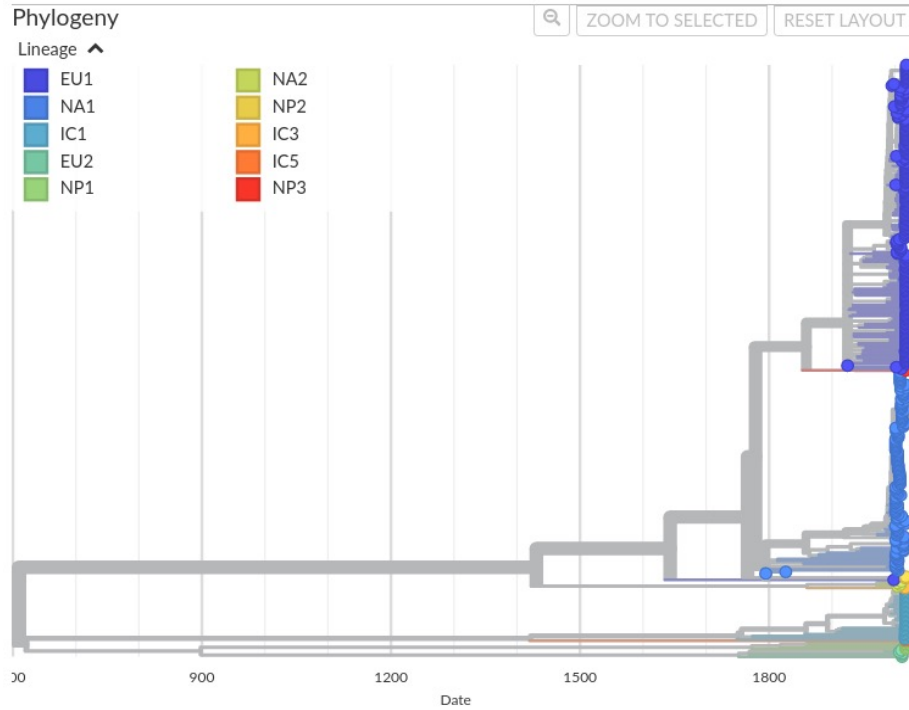
Tree Options: RECTANGULAR, RADIAL, UNROOTED, CLOCK, SCATTER

Branch Length: TIME, DIVERGENCE

Branch Labels: none

## Phytophthora ramorum

Maintained by Zachary Foster, Andrew Tupper, Nick Carleson and Niklaus Grunwald.  
Showing 371 of 371 genomes sampled between Oct 1794 and Jan 2019.



# Using **krisp** + SHERLOCK, we designed an assay for detection of *P. ramorum*

**Species level**

**Variant caller = GATK**

PR-102 - Ref. 

## Phyrar\_PR-102\_s0001:854172-854288 is diagnostic for Phytophthora

Reference : gGgatctttaTgagagTgattgcatTttcgacaaattcgagtgcgGggtcagcgc

*P. ramorum* : .....<T33>

non\_target: .....Y.....Y.....Y.....S.....<S7>.....Y..Y.....

oligos: gggatctttatgagagtgattgcattttcg  
└──────────────────┘  
Left primer

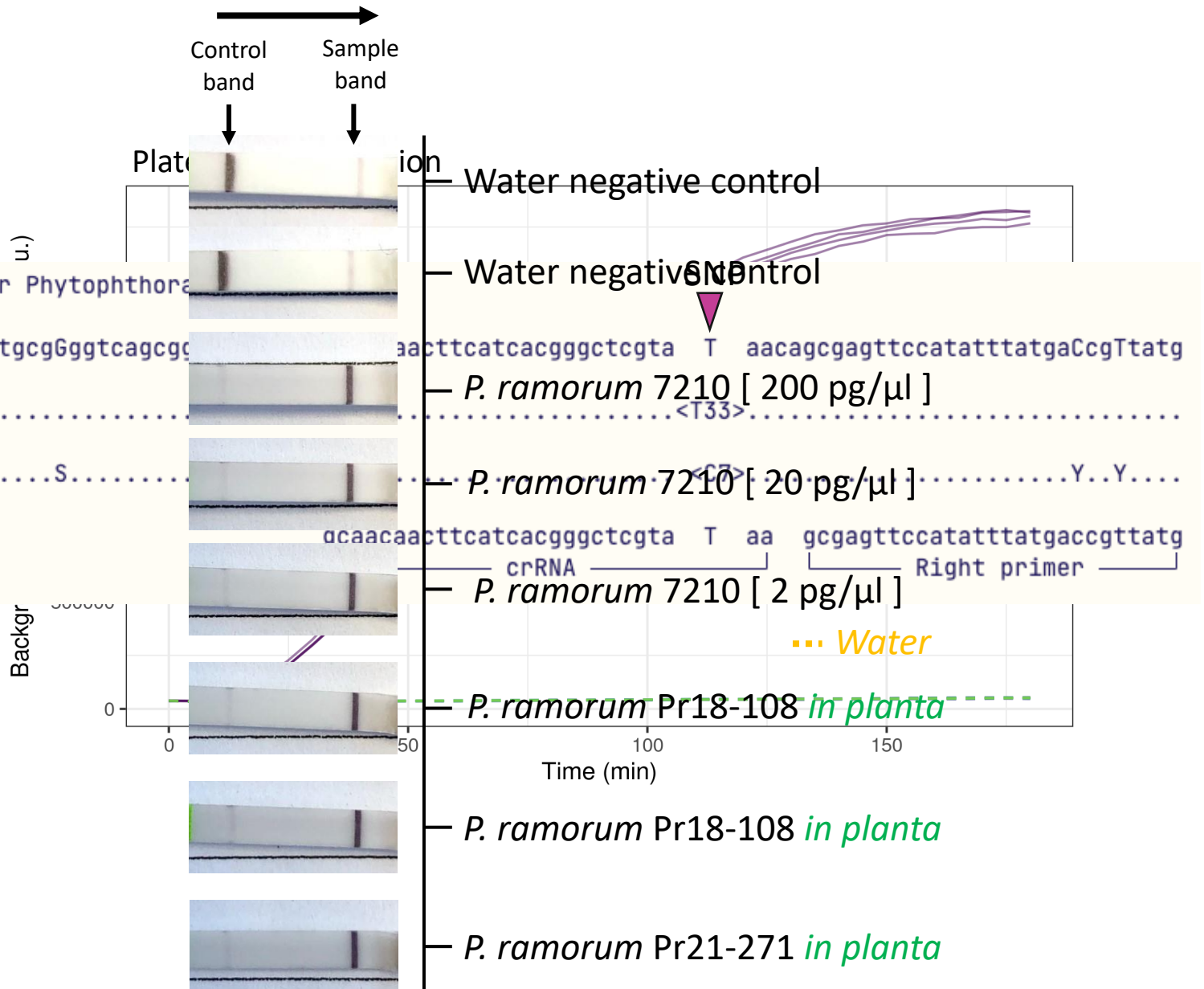
*P. lateralis* (N=1)

*P. syringae* (N=1)

**krisp\_vcf**

CSV

alignment



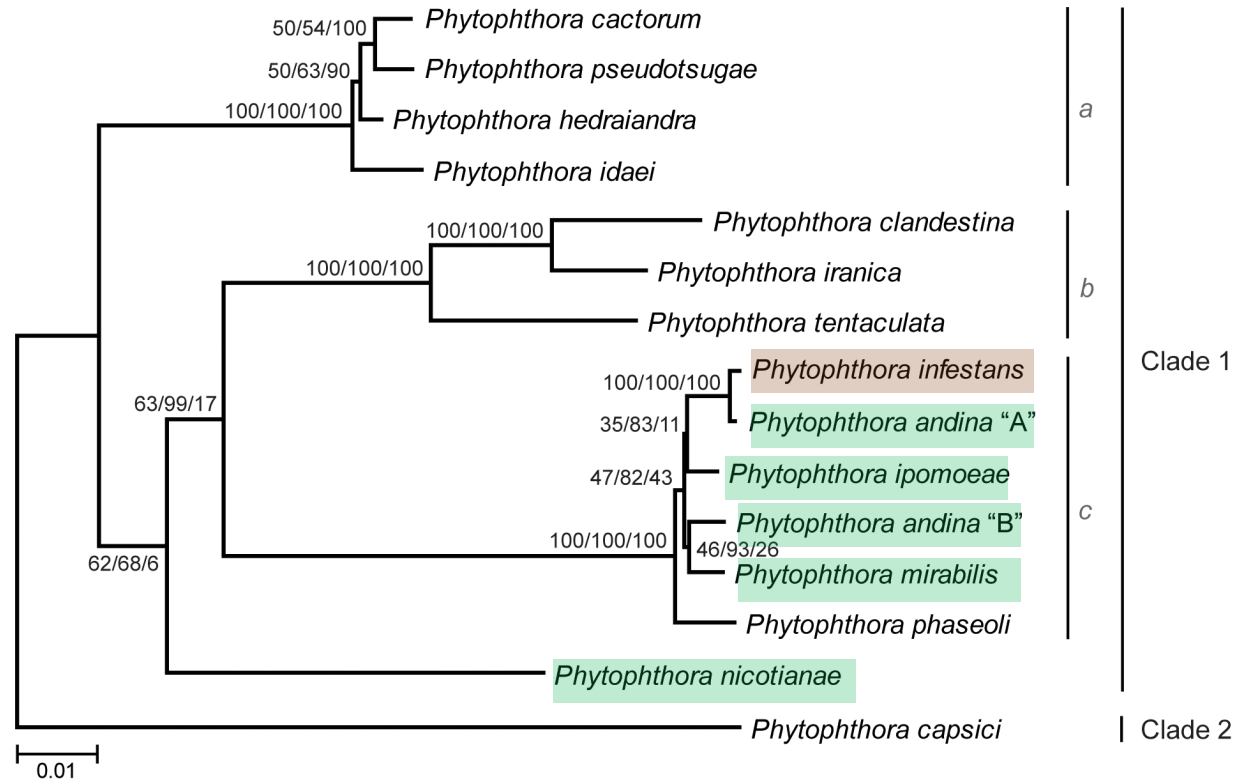




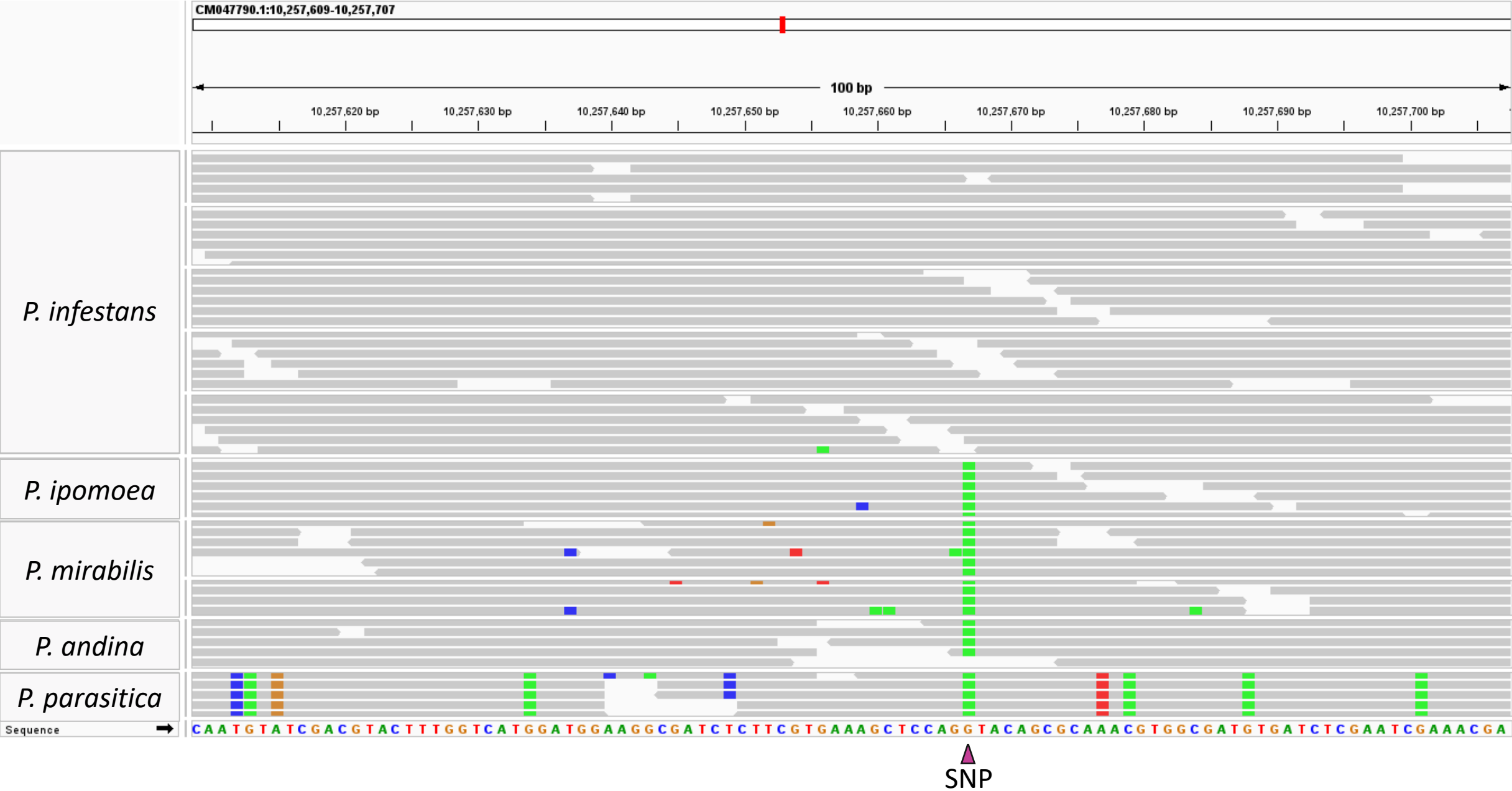
# A species level assay to distinguish *P. infestans*



Late blight



# krisp\_vcf identified a list of 51 sites for a *P. infestans* species level assay



# A lineage level assay to distinguish among variants of concern

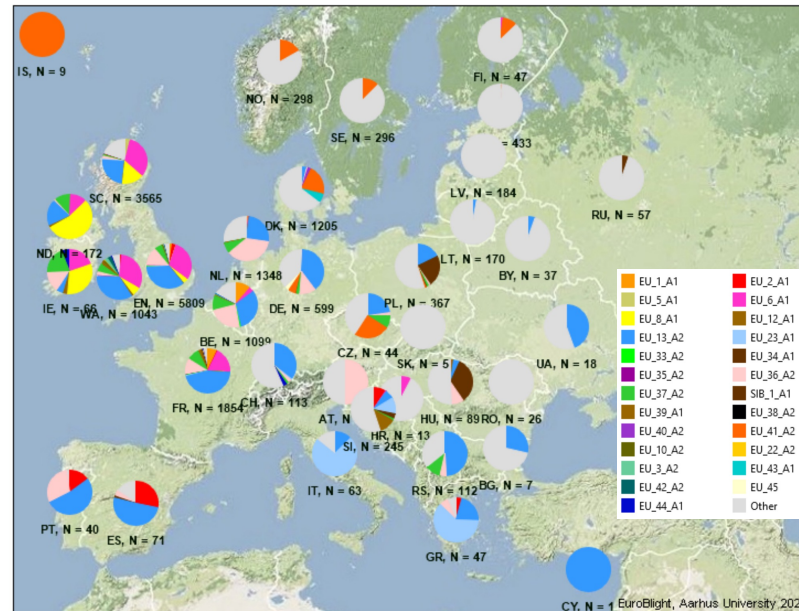
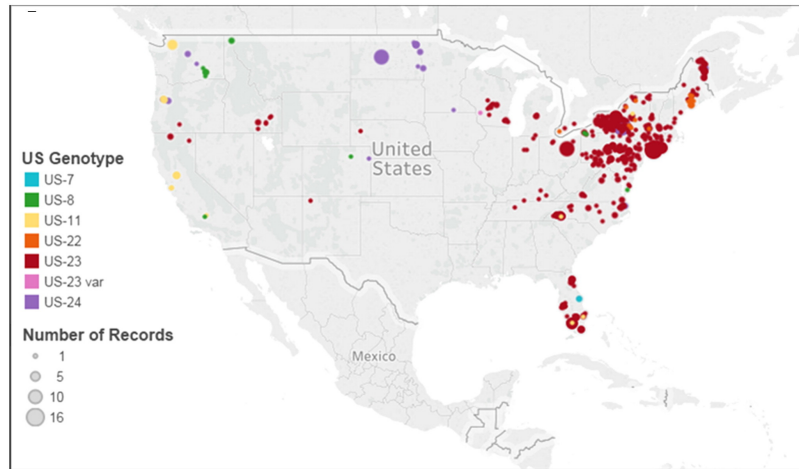
*Early detection*



*Late detection*



*Phytophthora infestans*



## Mefenoxam

US-01 = Insensitive

US-08 = Insensitive

US-11 = Sensitive

US-22 = Sensitive

US-23 = Sensitive

US-24 = Sensitive

# We identified variants to distinguish among *P. infestans* lineages

## CM047797.1:9488157-9488247 is diagnostic for US-11

SNP

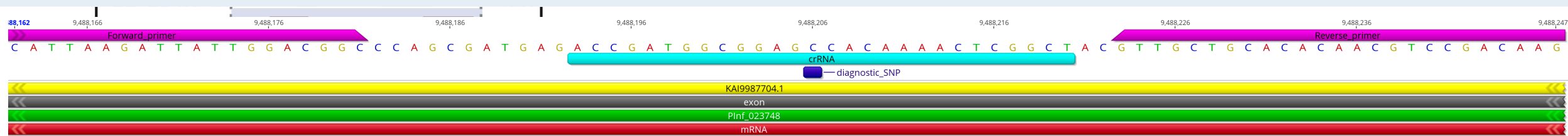


Reference: cgtagcattaagattattggacggcccagcgatgagaccgatggcggag C cacaaaactcggctacgttgctgcacacaacgtccgacaag  
 US-01 (7): ..... T7 .....  
 US-08 (5): ..... T5 .....  
 US-11 (4): ..... <C4> .....  
 US-22 (6): ..... T6 .....  
 US-23 (8): ..... T8 .....  
 US-24 (3): ..... T3 .....  
 EU\_13\_A2 (6): ..... T6 .....  
 EU\_36\_A2 (6): ..... T6 .....

oligos: cgtagcattaagattattggacggc  
 Left primer

accgatggcggag C cacaaaactcggct  
 crRNA

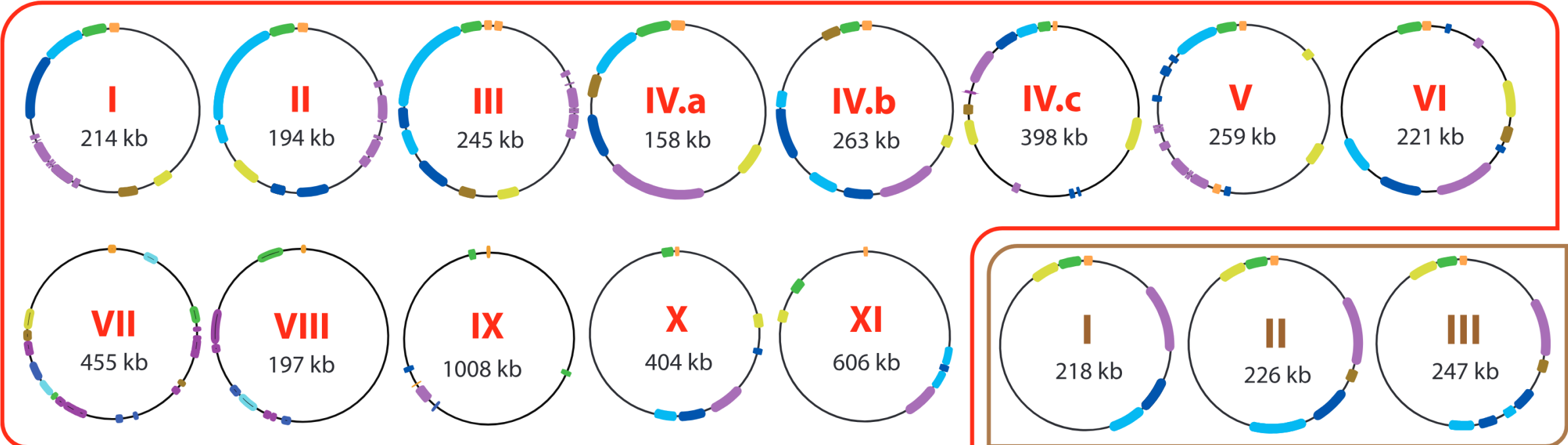
gttgctgcacacaacgtccgacaag  
 Right primer



Putative lipase

# A lineage/variant level assay to distinguish among plasmids of concern

## Tumor-inducing plasmids (pTi)



# We found variants to distinguish types of oncogenic plasmids

crRNA

GACCAAGCGAAGGATCTCAATAAACCCGCTTTCAAAAAGTGGACCAAACCCGCCAGATCCTATTCCCATTAGCTTGAATAGGTCCLCAATCATGAGGGAAA : Type Ia/b  
 .....A..... : Type III


**Primer statistics:**

Direction	Penalty	Sequence	Tm	Gc Percent	Self An	and Th	Hairpin Th
Forward	5.27932	GACCAAGCGAAGGATCTCAATAAAC	60.22068	44.0	0.0		0.0
Reverse	5.99488	CCTCATGATTGGGACCTATTCAAGC	61.49488	48.0	6.53393		0.0

**Pair statistics:**

Penalty	Compl Any Th	Compl End Th	Product Size	Product Tm
11.2742	3.07148	0.0	96	81.42821


**Tumor inducing plasmid**



I  
214 kb

: Type III

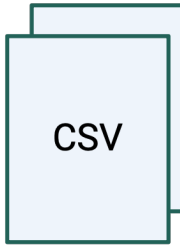
N = 60 plasmids



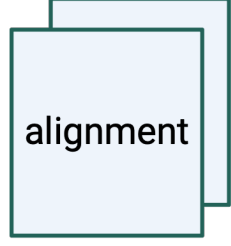
II  
194 kb

N = 19 plasmids

**krisp\_fasta**



CSV



alignment

**Wet lab validation in progress...**



## Collaborators

Jeff Chang, Jared LeBoldus, David Gent, Jessie Uehling, Sarah Navarro, Alex Weisberg, Richard Hamlin, Frank Martin, Howard Judelson, Rosie Bradshaw, Tom Sharpton, Chris Smart, Bill Fry, Hector Lozoya-Saldaña, Silvia Restrepo, Sylvia Fernández-Pavía, Alan Kanaskie, Ellen Goheen, Everett Hansen, Simone Prospero, Nari Williams, Rebecca McDougal, Jenifer Parke, Richard Bélanger, Jerry Weiland, Carolyn Scagel, Inga Zasada, Dimitre Mollov, Andy Jones, Thomas Jung, Patricia Manosalva, Nari Williams, Brent Kronmiller, ...

## Grünwald lab

Caroline Press, Val Fieland, Karan Fairchild, Zach Foster, Nick Cauldron, Martha Sudermann, Fernanda Iruegas-Bocado, Ricardo Alcalá-Briseño, Camilo Parada-Rojas, (Meg Larsen), (Brian Knaus), (Erica Goss), (Zhian Kamvar), (Javier Tabima), (Jonah Brooks), (Kim Graham), (Shankar Shakya), (Alex Weisberg), (Felipe Albornoz), (Jane Stewart), (Sydney Everhart) ...

## Funding



- OAN, HRI
- ODA, ODF, CDFA
- NNCRC
- USFS PSW
- J. Frank Schmidt

