Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Economic Affairs FDEA
Agroscope Changins-Wädenswil Research Station ACW

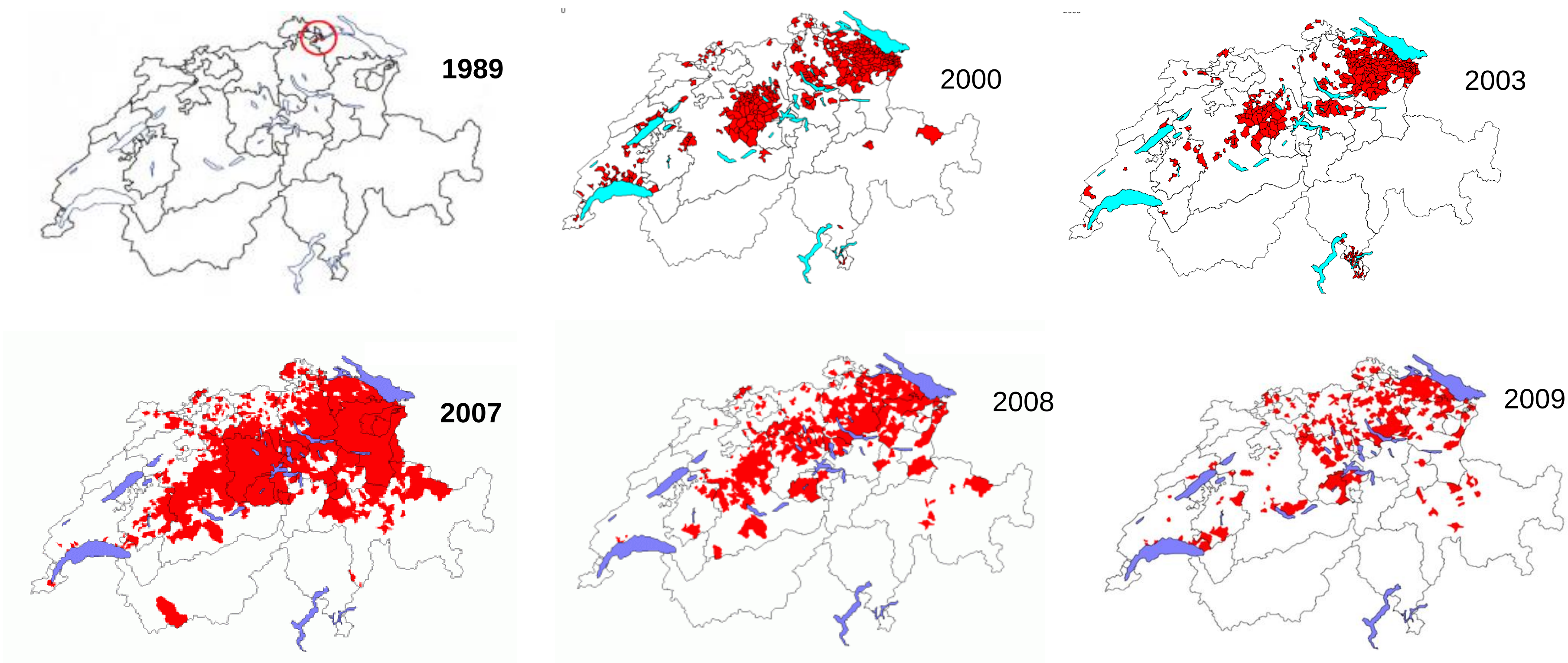
Agroscope

Source tracking and epidemiology of fire blight: exploiting the unexplored diversity of *E. amylovora*

Fabio Rezzonico

Fire Blight: history

- First reports in North East USA – 1780s
- First cases outside North America in 1910 (New Zealand)
- Arrived in Europe (UK) in 1950s
- First reported in Switzerland in 1989



What is source tracking?

- Identification of the spatial/temporal origin of a bacterial strain/population

  Where it comes from? When it's arrived?

- Requires characterization of the strain/population whose origin has to be assessed

  Morphological, biochemical or molecular methods

- Comparison with nearby strains/populations allows to establish the source of the strain/population under investigation

- The difficulty of the task increases if the diversity among strains and/or populations is low

Why source tracking?

- Identification of inoculum sources and reservoirs

  Targeted implementation of phytosanitary measures

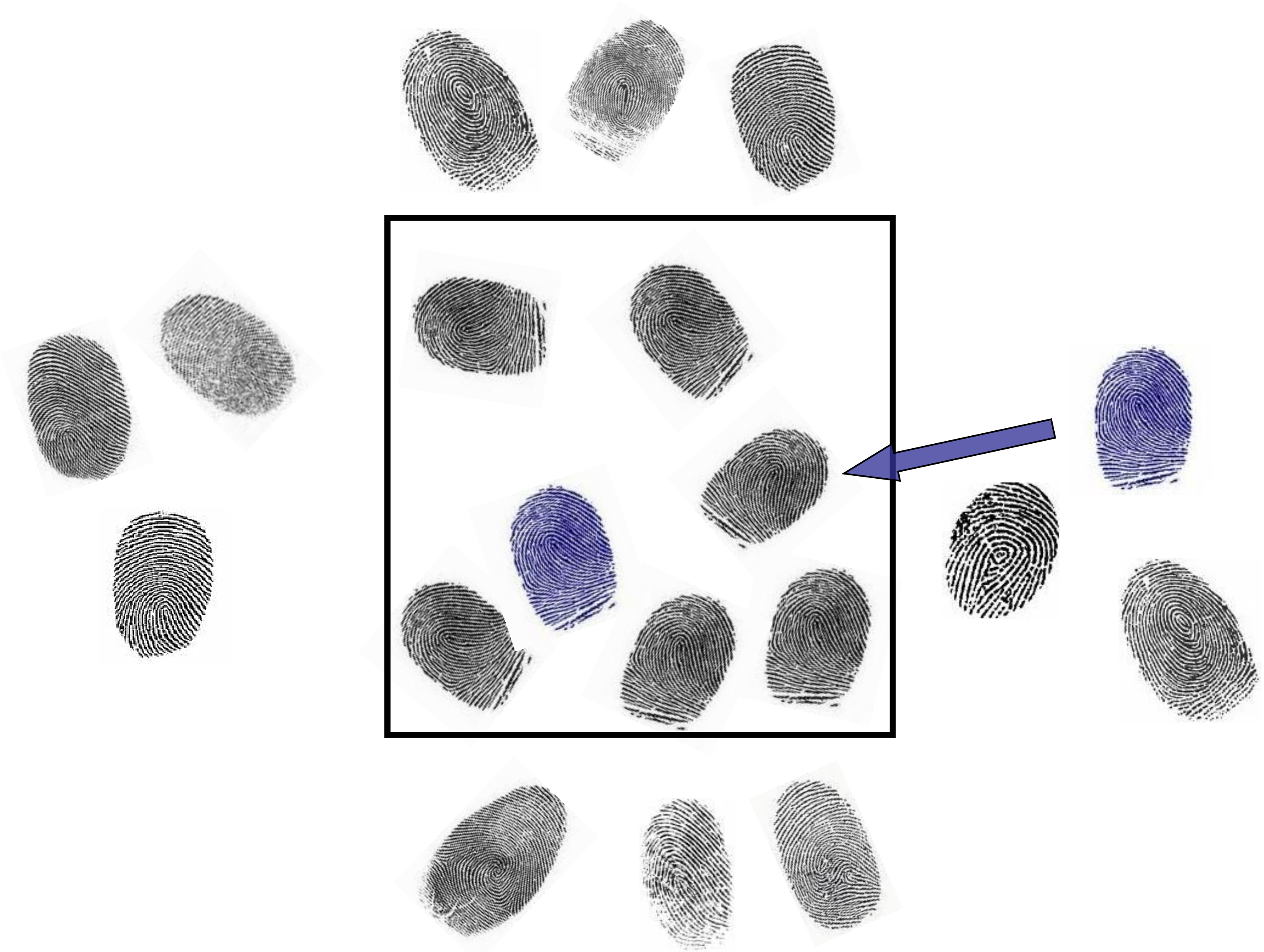
- Protection of orchards and nursery production

  In harmony with *Hochstamm* and other host plant cultivation (e.g., ornamentals)

- Understanding disease epidemiology

  Temporal and spatial spread of the disease

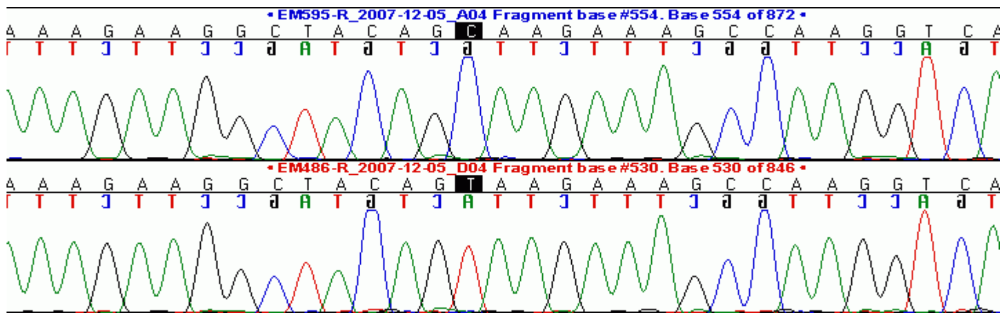
Source tracking: the principle



Molecular characterization

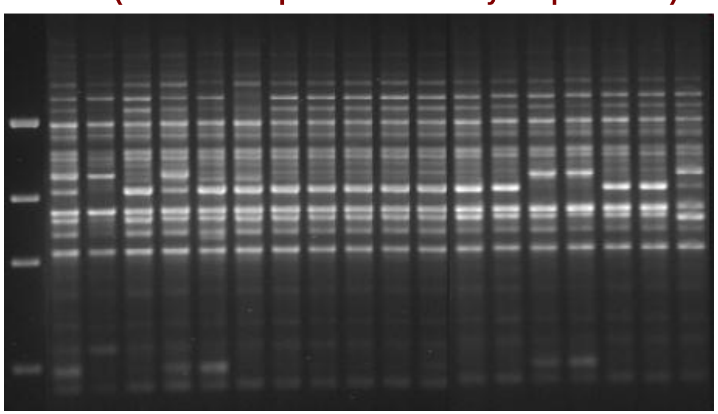
- Gene sequencing

Variation among DNA sequences: single nucleotide polymorphism (SNP), multiple locus sequence typing (MLST)

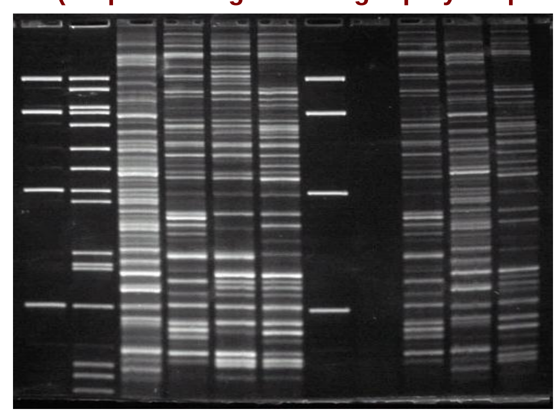


- Random molecular fingerprinting

RAPD (Random Amplification of Polymorphic DNA)

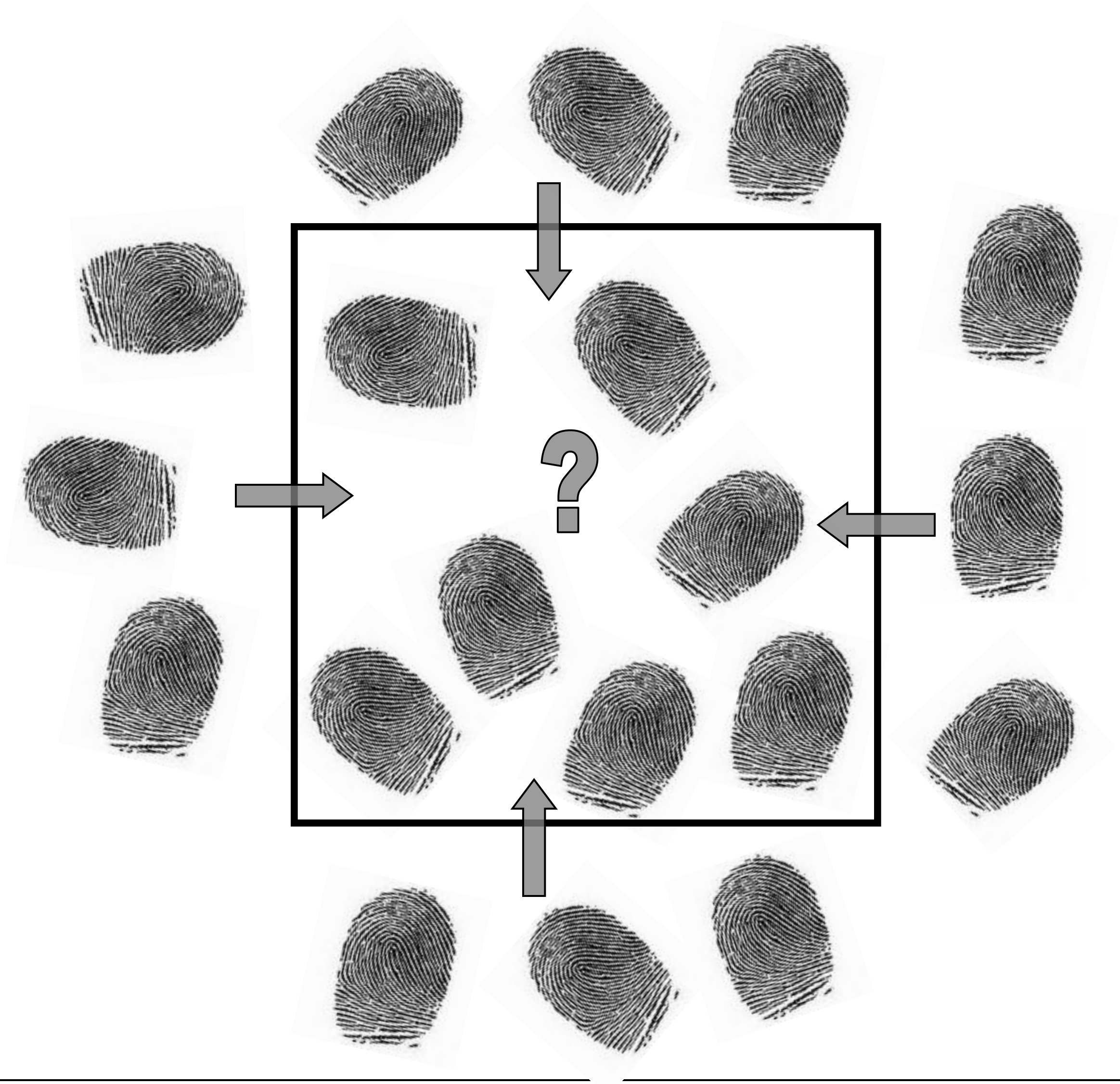


AFLP (Amplified fragment length polymorphism)



Agroscope

Source tracking: the *E. amylovora* case



Diversity in *Erwinia amylovora*

- Whole genome comparison among the 1 US + 5 EU strains sequenced
- Extremely low variability among strains (>99.99% identity)
- On average less than 100 SNP per strain were found in each genome

Per strain this corresponds to about one SNP every 36’000 to 146’000 bases

	CFBP1430	CFBP1232[T]	ACW56400	UPN527	01SFR-BO
Position					
2035	C	C	C	T	C
10754	G	G	G	A	G
11138	C	C	C	C	T
14465	C	C	A	C	C
22070	T	T	T	C	T
29924	A	G	A	G	G
46128	A	G	A	A	A
49615	A	G	A	A	A
52128	G	G	G	G	A
63783	T	T	G	T	T
68719	C	C	T	C	C
82840	G	G	A	G	G
91436	G	G	G	T	G
126068	T	C	T	C	C
142968	C	C	C	T	C
165475	C	T	C	C	C
172162	G	G	G	G	T
180211	T	C	T	C	C
188730	T	A	T	T	T
200206	C	C	C	T	C
217985	C	C	T	C	C
225771	C	C	C	C	T
233804	T	G	G	G	G
257568	G	G	G	A	G
271230	C	C	C	T	C
275858	T	C	C	T	T
293509	A	A	C	A	A
293515	G	G	A	G	G
296243	A	T	A	A	A
313323	G	G	G	C	G
⋮					
3646822	A	G	A	A	A
3676227	G	G	G	A	G
3731666	A	A	A	C	A
3751105	A	G	A	A	A
3751155	A	A	A	G	A
3795221	A	G	A	G	G
Total SNPs	35	26	94	106	35

New methods must be found to look for diversity in *E. amylovora*

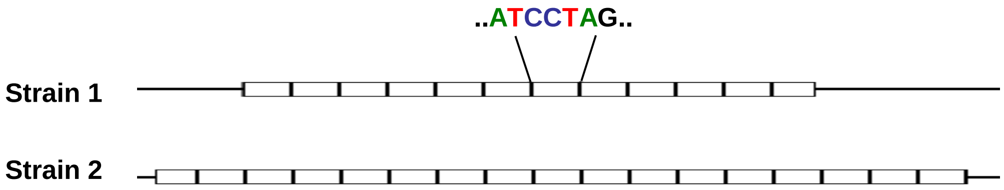
VNTRs & **CRISPRs**

Agroscope

Variable number tandem repeats (VNTRs)

- VNTRs are diversity hotspots in *E. amylovora*

  They consist of tandem repeats of a short repetitive DNA sequence which is present in variable number in different individuals/strains



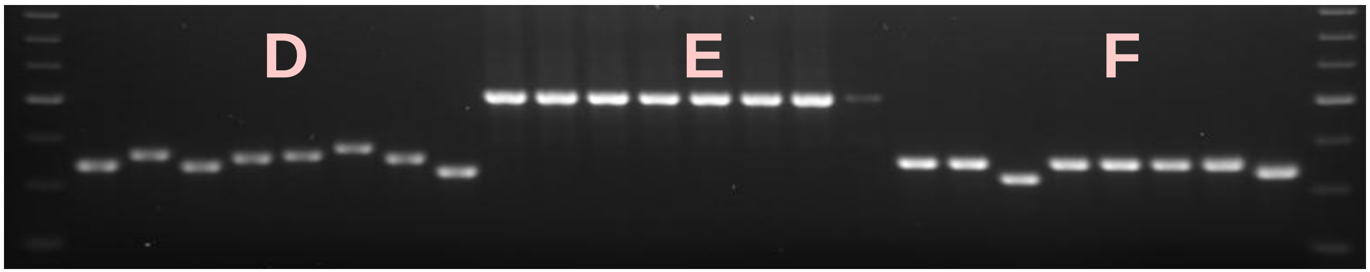
Localization of some VNTRs in CFBP 1430 genome

VNTR	Location	VNTR position		Length	Repeat length	Consensus sequence	Repeat number
A	pEA29	26041	26084	44	8	ATTACAGA	5
B	pEA29	14703	14726	24	8	TCAGCCTC	3
C	Chrom.	457256	457298	43	6	ATTGTT	9
D	Chrom.	1254077	1254165	89	6	TGGCAA	7
E	Chrom.	1535272	1535523	252	18	TTCCACCGCCGGAGCTGC	14
F	Chrom.	2668446	2668564	119	18	GGCAGCGTTAGTGCTAGT	6
G	Chrom.	2944996	2945030	35	6	TGATAT	5
H	Chrom.	3517853	3517922	70	9	GCGTGATAT	7
I	Chrom.	3591248	3591353	106	6	CTGGTT	16
J	Chrom.	3782324	3782457	134	9	GCTGTAA--TG	13

Variable number tandem repeats (VNTRs)

Courtesy of Tanja Dreo

Primers designed on flanking regions and test with a panel of strains of wolrdwide origin



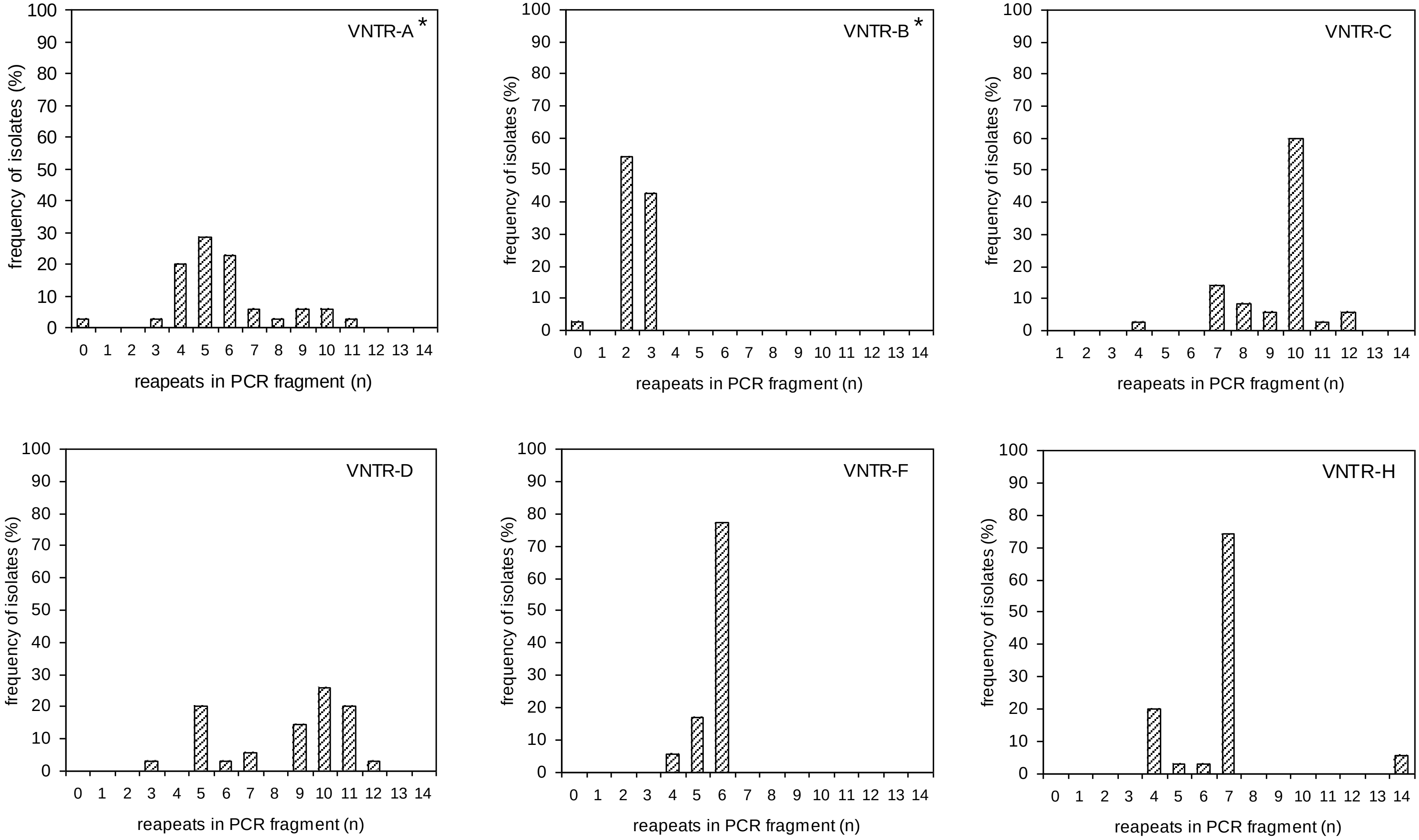
→ Some VNTRs are better than others: VNTRs with no or very low variability are discarded!

Localization of some VNTRs in CFBP 1430 genome

VNTR	Location	VNTR position		Length	Repeat length	Consensus sequence	Repeat number
A	pEA29	26041	26084	44	8	ATTACAGA	5
B	pEA29	14703	14726	24	8	TCAGCCTC	3
C	Chrom.	457256	457298	43	6	ATTGTT	9
D	Chrom.	1254077	1254165	89	6	TGGCAA	7
~~**E**~~	~~Chrom.~~	~~1535272~~	~~1535523~~	~~252~~	~~18~~	~~TTCCACCGCCGGAGCTGC~~	~~14~~
F	Chrom.	2668446	2668564	119	18	GGCAGCGTTAGTGCTAGT	6
~~**G**~~	~~Chrom.~~	~~2944996~~	~~2945030~~	~~35~~	~~6~~	~~TGATAT~~	~~5~~
H	Chrom.	3517853	3517922	70	9	GCGTGATAT	7
~~**I**~~	~~Chrom.~~	~~3591248~~	~~3591353~~	~~106~~	~~6~~	~~CTGGTT~~	~~16~~
~~**J**~~	~~Chrom.~~	~~3782324~~	~~3782457~~	~~134~~	~~9~~	~~GCTGTAA-TG~~	~~13~~

VNTRs: frequency distribution

Courtesy of Tanja Dreo



VNTRs: discrimination power

(collection of worldwide *E. amylovora* isolates)

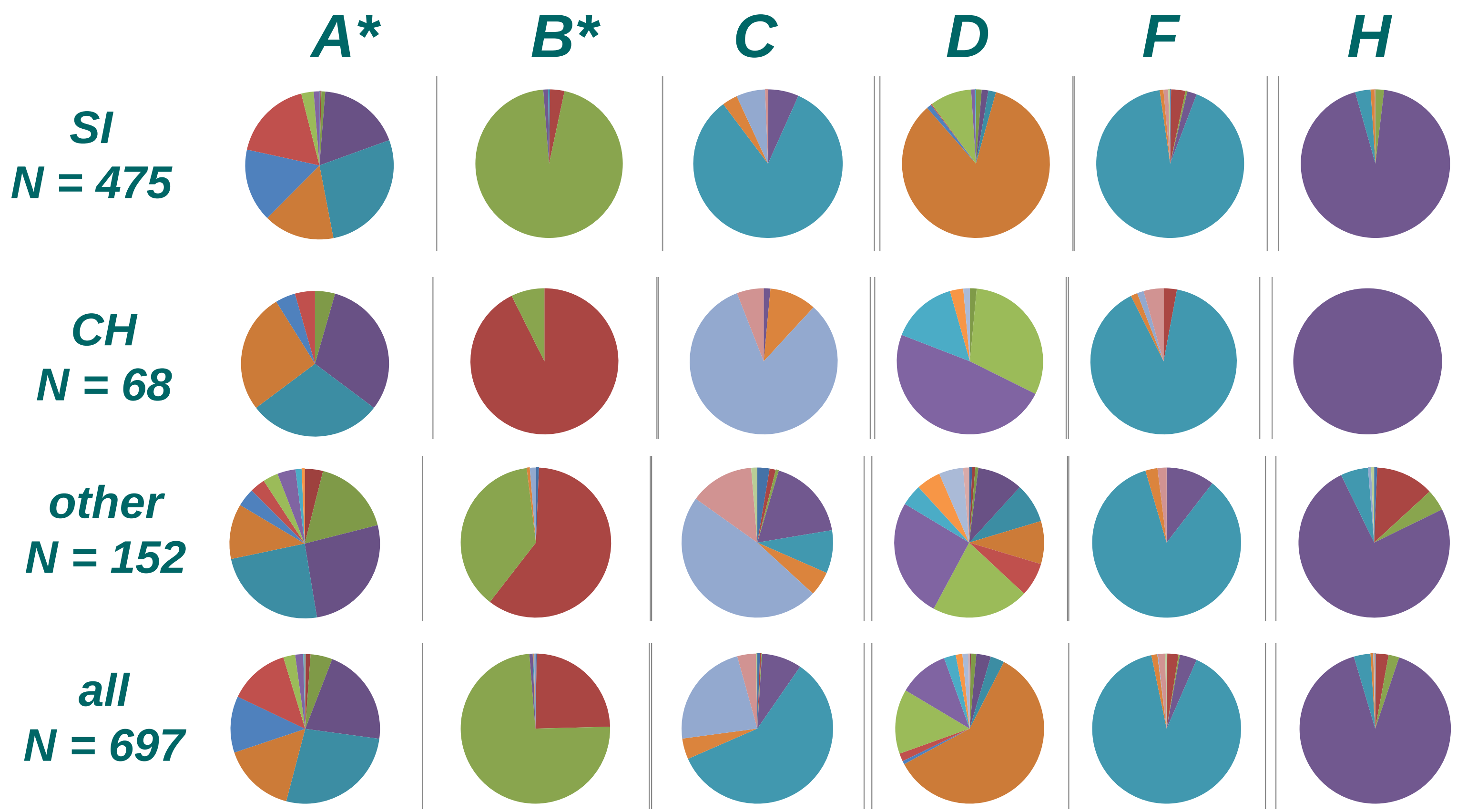
Distance between strains expressed as number of different VNTR-systems (max=6)

	IH3-1/colE1	LA092/pEU30	OR29/pEU30	JL 1189/pEU30	LA476	OR25/pEU30	JL1185	ACW56400	IL-5/colE1	Ea110R	UTFer2/pEU30	UTRJ2/pEU30	FAW63230	UPN527	Ea7/74	Ea4/82	Ea263	CFBP 3020	FAW63889	FAW63679	CFBP 3792	CFBP1430	Ea273	CFBP 3049	Leb B66/pEL60	Ea209	CFBP1232T	CFBP 3025	Leb A3/pEL60	01SFR-BO	CFBP 3098	CFBP 2301	Ea02	Ea153	FAW64132
IH3-1/colE1	0	5	5	5	5	5	5	5	5	5	5	5	5	5	5	6	6	5	4	5	4	5	4	4	5	5	5	5	5	5	4	5	5	4	4
LA092/pEU30		0	0	0	2	2	1	6	3	5	5	5	5	6	6	6	6	6	5	6	4	6	4	4	6	6	6	6	6	6	4	6	6	5	5
OR29/pEU30			0	0	2	1	1	6	3	5	5	5	5	6	6	6	6	6	5	6	4	6	4	4	6	6	6	6	6	6	4	6	6	5	5
JL 1189/pEU30				0	2	1	1	6	3	5	5	5	5	6	6	6	6	6	5	6	4	6	4	4	6	6	6	6	6	6	4	6	6	5	5
LA476					0	1	1	4	4	5	5	5	5	6	6	6	5	6	5	6	4	5	5	5	5	5	6	6	6	5	5	6	5	5	4
OR25/pEU30						0	0	5	4	5	5	5	5	6	6	6	6	6	5	6	4	5	5	5	5	5	6	6	6	5	5	6	5	5	4
JL1185							0	5	4	5	5	5	5	6	6	6	6	6	5	6	4	5	5	5	5	5	6	6	6	5	5	6	5	5	4
ACW56400								0	6	5	5	5	5	5	5	4	3	4	5	4	5	2	5	5	3	3	4	4	4	3	5	4	3	5	4
IL-5/colE1									0	4	5	4	4	5	5	4	4	5	4	5	4	5	3	3	5	5	5	5	5	5	3	4	5	3	4
Ea110R										0	2	2	3	4	5	4	4	4	3	4	3	4	3	3	3	3	4	4	4	4	3	4	4	3	3
UTFer2/pEU30											0	0	3	4	5	4	4	4	3	4	3	4	3	3	3	3	4	4	4	4	3	4	4	3	3
UTRJ2/pEU30												0	3	4	5	4	4	4	3	4	3	4	3	3	3	3	4	4	4	4	3	4	4	3	3
FAW63230													0	3	5	3	4	3	3	4	2	4	3	3	4	4	3	4	4	4	2	4	4	2	2
UPN527														0	3	4	4	2	3	3	3	3	3	3	3	3	2	3	3	3	2	3	3	2	2
Ea7/74															0	2	2	3	4	2	4	2	4	4	3	3	1	1	1	2	4	0	1	3	4
Ea4/82																0	1	3	4	2	4	3	4	4	3	3	2	1	1	2	4	1	2	3	4
Ea263																	0	3	4	2	4	3	4	4	3	3	2	1	1	2	4	1	2	3	4
CFBP 3020																		0	3	2	3	2	3	3	2	2	1	2	2	2	2	2	2	2	2
FAW63889																			0	2	2	3	2	2	3	3	2	3	3	3	2	3	3	2	2
FAW63679																				0	3	2	3	3	2	2	2	1	1	1	3	1	1	3	3
CFBP 3792																					0	3	2	2	3	3	3	4	3	3	2	3	3	2	2
CFBP1430																						0	3	3	1	1	2	2	2	1	3	2	1	3	2
Ea273																							0	0	3	3	3	3	3	3	1	3	3	2	2
CFBP 3049																								0	3	3	3	3	3	3	1	3	3	2	2
Leb B66/pEL60																									0	0	2	2	2	1	3	2	1	3	2
Ea209																										0	2	2	2	1	3	2	1	3	2
CFBP1232T																											0	1	1	3	2	1	2	1	2
CFBP 3025																												0	0	1	3	0	1	2	3
Leb A3/pEL60																													0	1	3	0	1	2	3
01SFR-BO																														0	3	1	0	3	2
CFBP 3098																															0	3	3	1	1
CFBP 2301																																0	1	2	3
Ea02																																	0	3	2
Ea153																																		0	1
FAW64132																																			0

➡ 35 strains and 27 different genotypes

VNTRs: comparison of two national populations

Courtesy of Tanja Dreo



➡ Substantial differences between populations in Slovenia and Switzerland

VNTRs: future approach

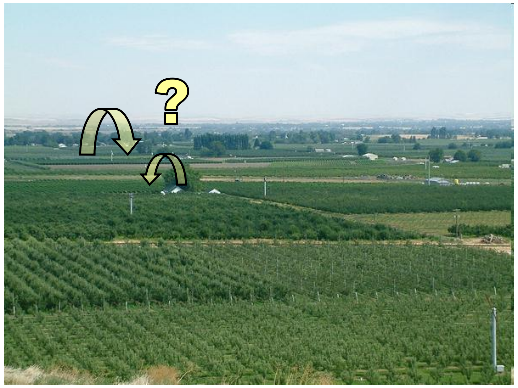
- Multiplexing of the PCR reactions and labeling with fluorescent primers

  Setup of a fragment analysis protocol will allow to go from the isolates to the results in just one PCR step, allowing automation

Isolates **Cell lysis** **Labelled multiplex PCR** **Fragment analysis** **Results**

- Testing discriminatory power at national, regional and local level

  Epidemiological studies and more appropriate phytosanitary measures



Clustered regularly interspaced short palindromic repeats (CRISPRs)

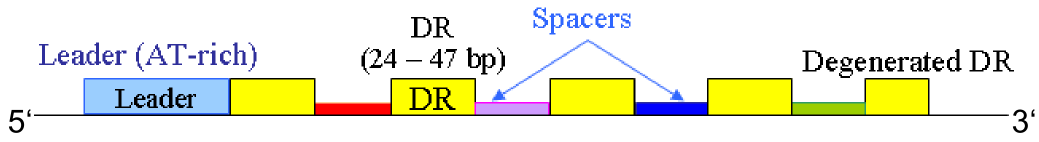
- CRISPRs are diversity hotspots in *E. amylovora*

  > Direct repeats (24-47 bp) separated by unique spacers of similar length

  > Spacer sequences match sequences in plasmids or phage genomes

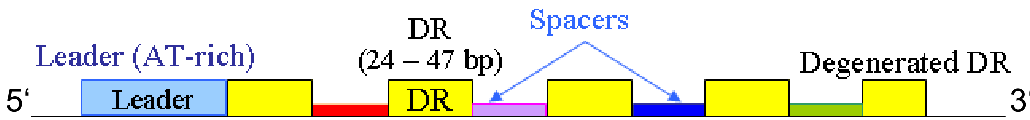
  > Together with the CRISPR assiociated (*cas*) genes is part of an RNA interference system present both in Archea and Bacteria

  > Three CRISPR regions (CRR) are present in *E. amylovora*

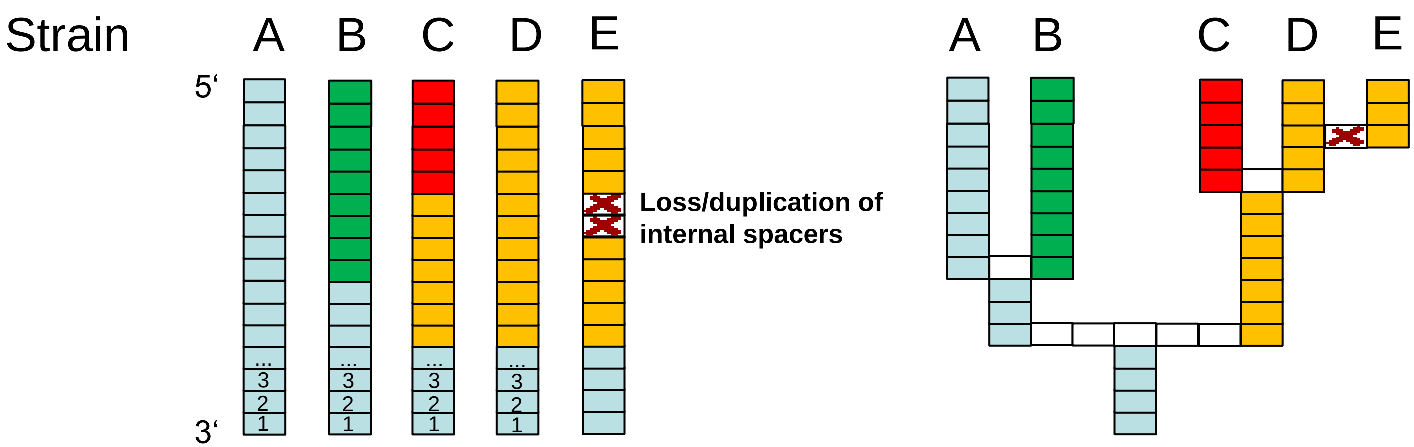


Agroscope

How are CRISPRs assembled?



- Actively evolving with new challenges (phages/plasmids)
  - > New spacers inserted polarily at the 5′ end of the cluster next to the leader
  - > Older spacers are frequently common whereas newer spacers are unique
  - > Chronological record of past encounters with foreign DNA
  - > CRR-1 and CRR-2 are active and evolving, while CRR-4 seems inactive



Strains that display recent incorporation of new spacers at 5’ end and/or deletions in the central regions of the array can not be ancestral

Typing results

Among the first 37 strains investigated

- CRR1 displayed 14 different genotypes
- CRR2 displayed 13 different genotypes
- CRR4 displayed 3 different genotypes

} 18 different combinations (CRR1, CRR2)

> No further discrimination (CRR4)

Total of 18 different genotypes

	01SFR-B0	FAW63579	Ea209/pEL60	CFBP 2301	CFBP3025	Ea4-82	CFBP 3020	ACW56400	Ea7/74	CFBP1232T	CFBP1430	Leb B66/pEL60	UPN527	CFBP3098	FAW64132	Leb A3/pEL60	FAW63230	Ea263
Genotype CRISPR-1	B	B	B	B	B	B	B	B	B	B	B	B	D	D	D	D	D	F
Genotype CRISPR-2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	d
Genotype CRISPR-3	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
CRISPR Genotype	2	2	2	2	2	2	2	2	2	2	2	2	4	4	4	4	4	6
CRISPR group	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia

	Ea02	Ea273	Ea153	FAW63889	Ea110R	CFBP3049	UTFer2/pEU30	UTRJ2/pEU30	CFBP 3792	IH 3-1/colE1	Ea7-96r	JL1185/pEU30	LA476/pEU30	OR29/pEU30	JL1189/pEU30	LA092/pEU30	OR25/pEU30	Ea6-96r	IL-5/colE1
Genotype CRISPR-1	C	A	E	D	B	B	G	G	G	L	N	H	H	I	I	I	J	M	K
Genotype CRISPR-2	c	a	b	l	e	f	g	g	b	k	n	h	h	h	h	h	h	m	i
Genotype CRISPR-3	α	α	α	α	α	α	α	α	α	α	γ	α	α	α	α	α	α	γ	β
CRISPR Genotype	3	1	5	16	7	8	9	9	10	15	18	11	11	12	12	12	13	17	14
CRISPR group	Ia	Ia	Ia	Ia	Ib	Ib	II	II	II	IH	R	III	III	III	III	III	III	R	R

Typing results

Among the first 37 strains investigated

- CRR1 displayed 14 different genotypes
- CRR2 displayed 13 different genotypes
- CRR4 displayed 3 different genotypes

18 different combinations (CRR1 + CRR2)

> No further discrimination (CRR4)

Total of 18 different genotypes

	01SFR-B0	FAW63579	Ea209/pEL60	CFBP 2301	CFBP3025	Ea4-82	CFBP 3020	ACW56400	Ea7/74	CFBP1232T	CFBP1430	Leb B66/pEL60	UPN527	CFBP3098	FAW64132	Leb A3/pEL60	FAW63230
Genotype CRISPR-1	B	B	B	B	B	B	B	B	B	B	B	B	D	D	D	D	D
Genotype CRISPR-2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
Genotype CRISPR-3	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
CRISPR Genotype	2	2	2	2	2	2	2	2	2	2	2	2	4	4	4	4	4
CRISPR group	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia
PCR ribotyping										1	1		1				

	Ea263	Ea02	Ea273	Ea153	FAW63889	Ea110R	CFBP3049	UTFer2/pEU30	UTRJ2/pEU30	CFBP 3792	IH 3-1/colE1	Ea7-96r	JL1185/pEU30	LA476/pEU30	OR29/pEU30	JL1189/pEU30	LA092/pEU30	OR25/pEU30	Ea6-96r	IL-5/colE1
Genotype CRISPR-1	F	C	A	E	D	B	B	G	G	G	L	N	H	H	I	I	I	J	M	K
Genotype CRISPR-2	d	c	a	b	l	e	f	g	g	b	k	n	h	h	h	h	h	h	m	i
Genotype CRISPR-3	α	α	α	α	α	α	α	α	α	α	α	γ	α	α	α	α	α	α	γ	β
CRISPR Genotype	6	3	1	5	16	7	8	9	9	10	15	18	11	11	12	12	12	13	17	14
CRISPR group	Ia	Ia	Ia	Ia	Ia	Ib	Ib	II	II	II	IH	R	III	III	III	III	III	III	R	R
PCR ribotyping			1												3	3	3	3		4

> CRISPR-grouping correlate with PCR-ribotyping data (McManus & Jones, 1995; Donat *et al*., 2007)

Typing results

Among the first 37 strains investigated

- CRR1 displayed 14 different genotypes
- CRR2 displayed 13 different genotypes
- CRR4 displayed 3 different genotypes

} 18 different combinations

> No further discrimination

Total of 18 different genotypes

	01SFR-B0	FAW63579	Ea209/pEL60	CFBP 2301	CFBP3025	Ea4-82	CFBP 3020	ACW56400	Ea7/74	CFBP1232T	CFBP1430	Leb B66/pEL60	UPN527	CFBP3098	FAW64132	Leb A3/pEL60	FAW63230	Ea263	Ea02
Genotype CRISPR-1	B	B	B	B	B	B	B	B	B	B	B	B	D	D	D	D	D	F	C
Genotype CRISPR-2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	d	c
Genotype CRISPR-3	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α	α
CRISPR Genotype	2	2	2	2	2	2	2	2	2	2	2	2	4	4	4	4	4	6	3
CRISPR group	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia	Ia
PCR ribotyping										1	1		1						
PFGE Genotype	6			1		2			1	1	3A		4					2	

	Ea273	Ea153	FAW63889	Ea110R	CFBP3049	UTFer2/pEU30	UTRJ2/pEU30	CFBP 3792	IH 3-1/colE1	Ea7-96r	JL1185/pEU30	LA476/pEU30	OR29/pEU30	JL1189/pEU30	LA092/pEU30	OR25/pEU30	Ea6-96r	IL-5/colE1
Genotype CRISPR-1	A	E	D	B	B	G	G	G	L	N	H	H	I	I	I	J	M	K
Genotype CRISPR-2	a	b	l	e	f	g	g	b	k	n	h	h	h	h	h	h	m	i
Genotype CRISPR-3	α	α	α	α	α	α	α	α	α	γ	α	α	α	α	α	α	γ	β
CRISPR Genotype	1	5	16	7	8	9	9	10	15	18	11	11	12	12	12	13	17	14
CRISPR group	Ia	Ia	Ia	Ib	Ib	II	II	II	IH	R	III	III	III	III	III	III	R	R
PCR ribotyping	1												3	3	3	3		4
PFGE Genotype	other	1		4			4		G									

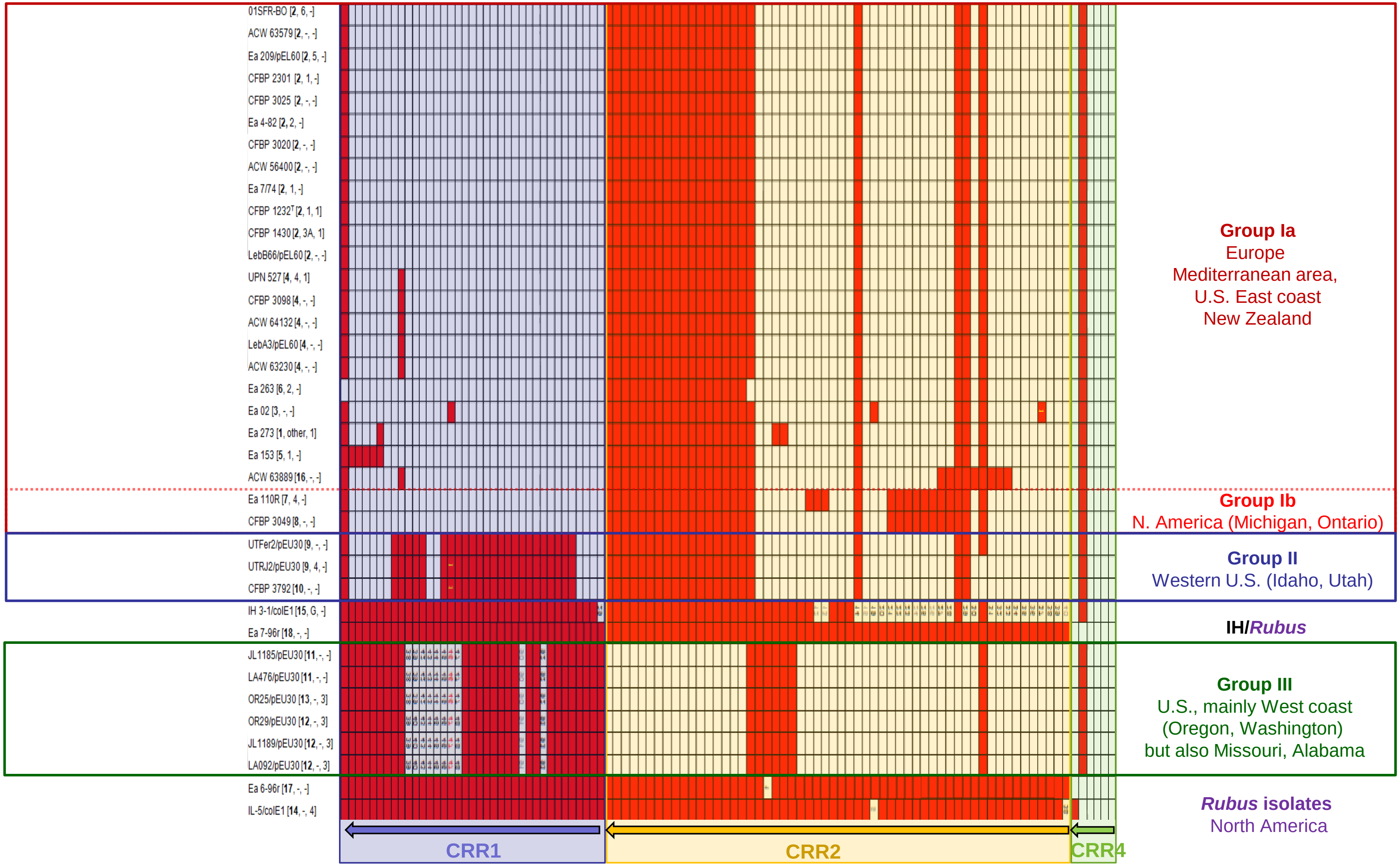
> CRISPR-grouping correlate with PCR-ribotyping data (McManus & Jones, 1995; Donat *et al*., 2007)

> Strains with the same PFGE genotype (Jock *et al*., 2002) could be separated using CRISPR analysis

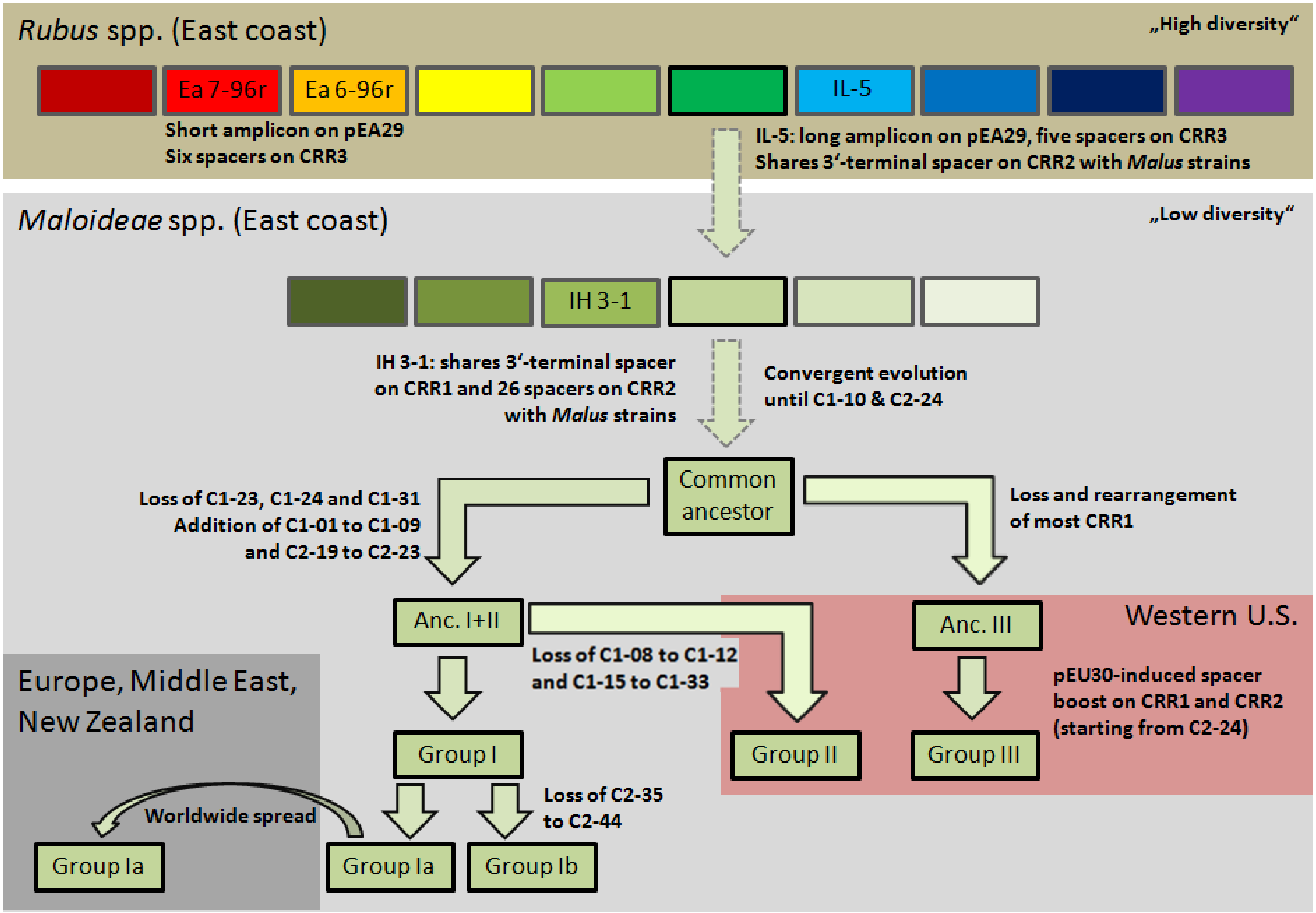
Agroscope

Clustering of strains

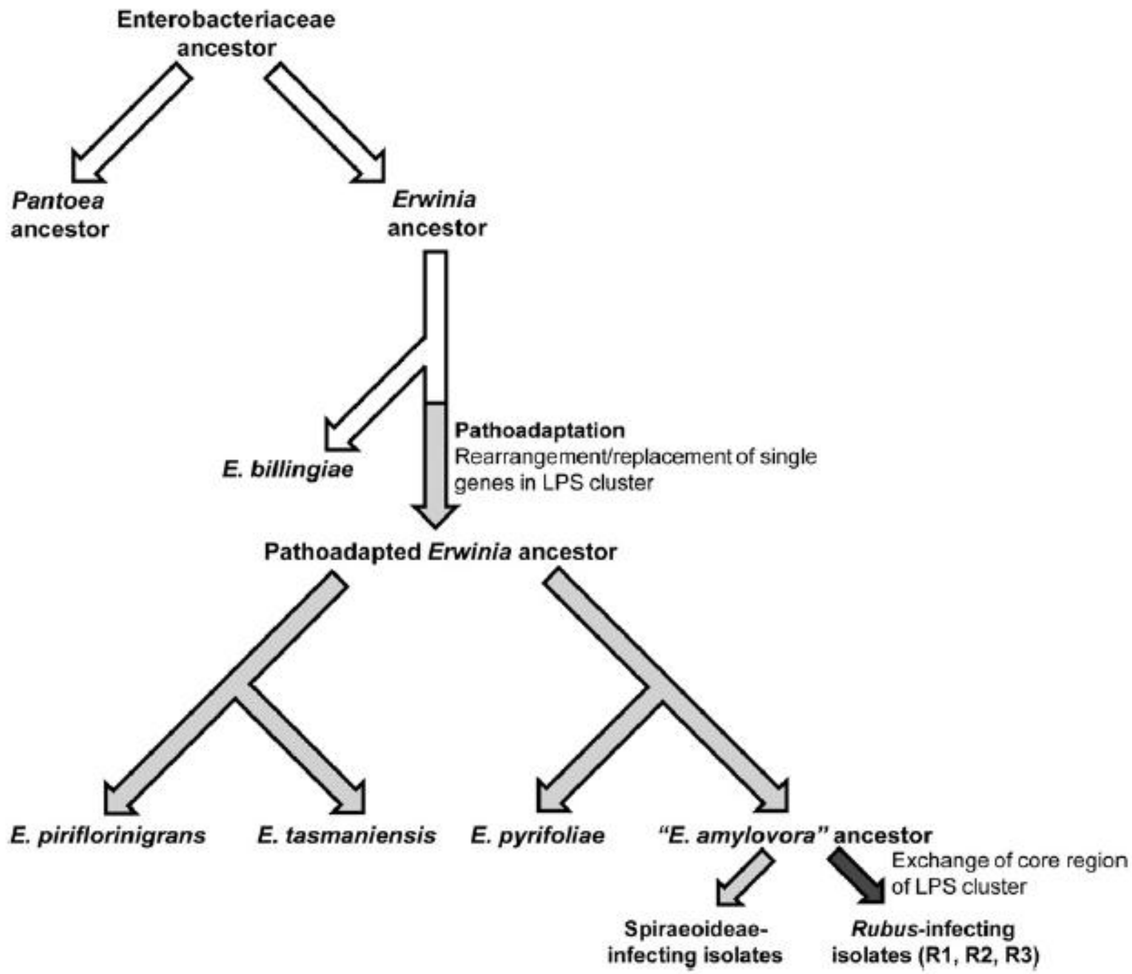
Correlation to geographic origin



Original model for *E. amylovora* evolution based on CRISPRs

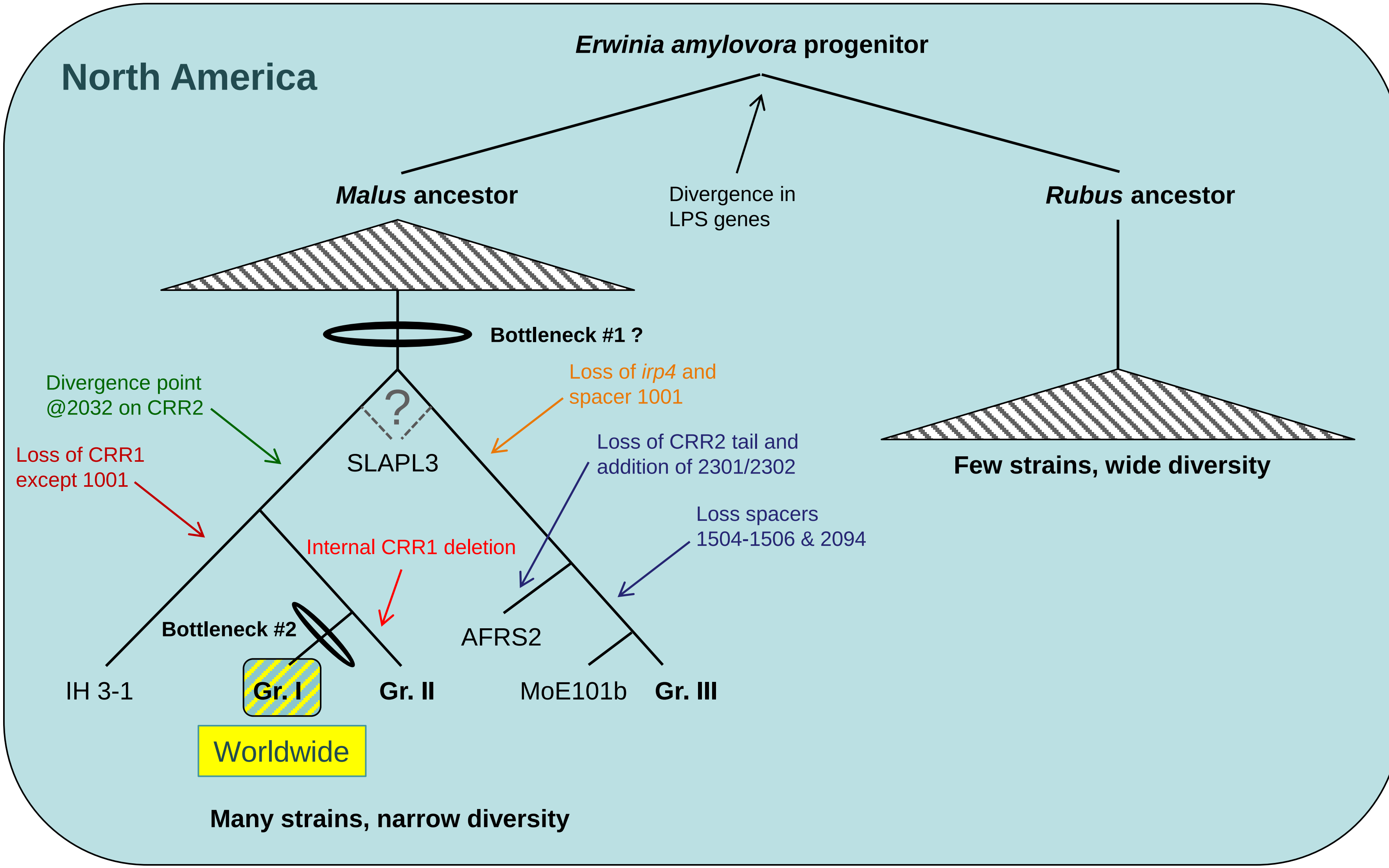


Genome analysis reveals unique LPS cluster for *Rubus*-infecting *E. amylovora* strains



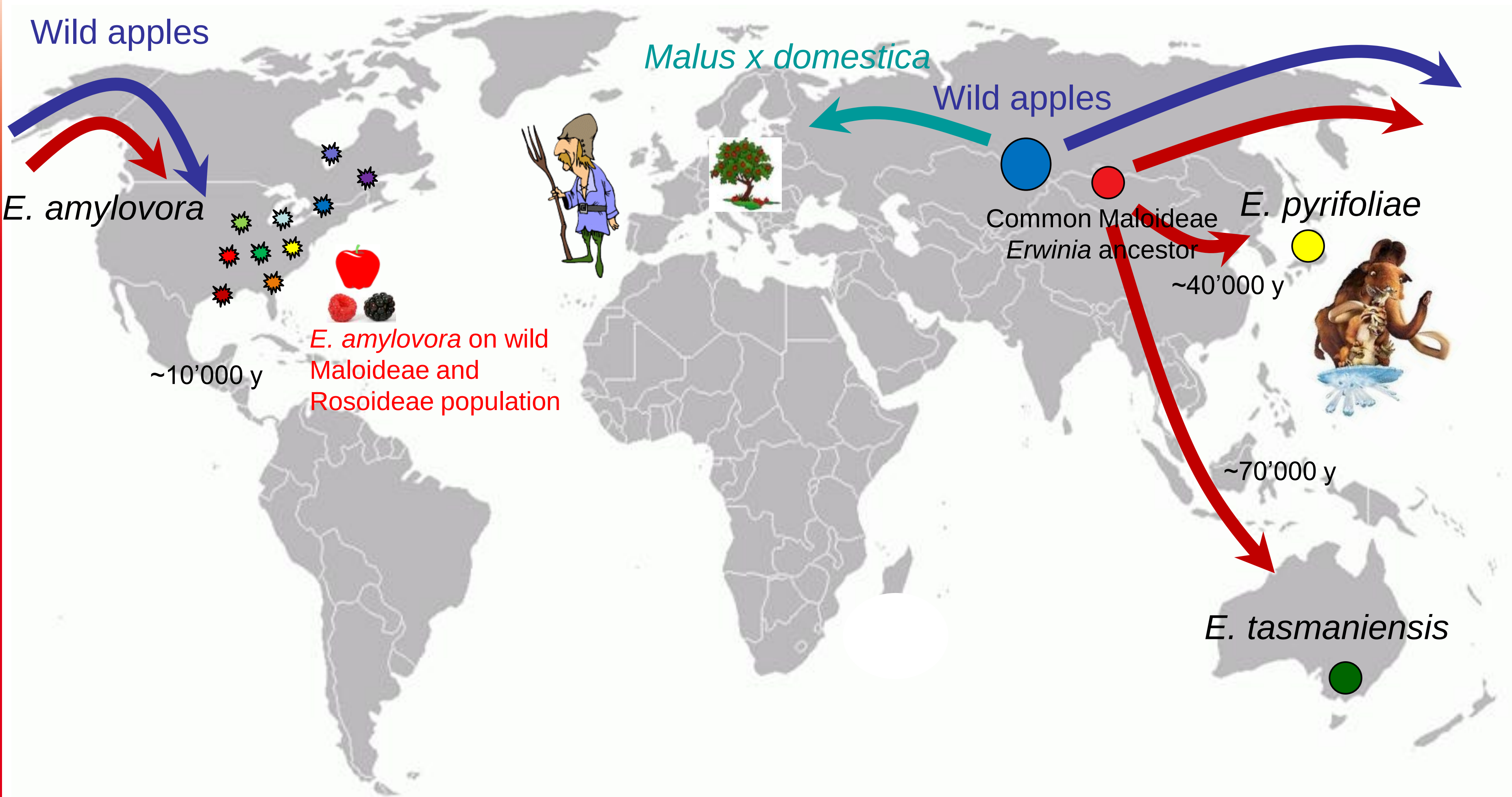
➡ *Malus* infecting isolates can not derive from *Rubus* infecting isolates

Evolutionary history of *E. amylovora*

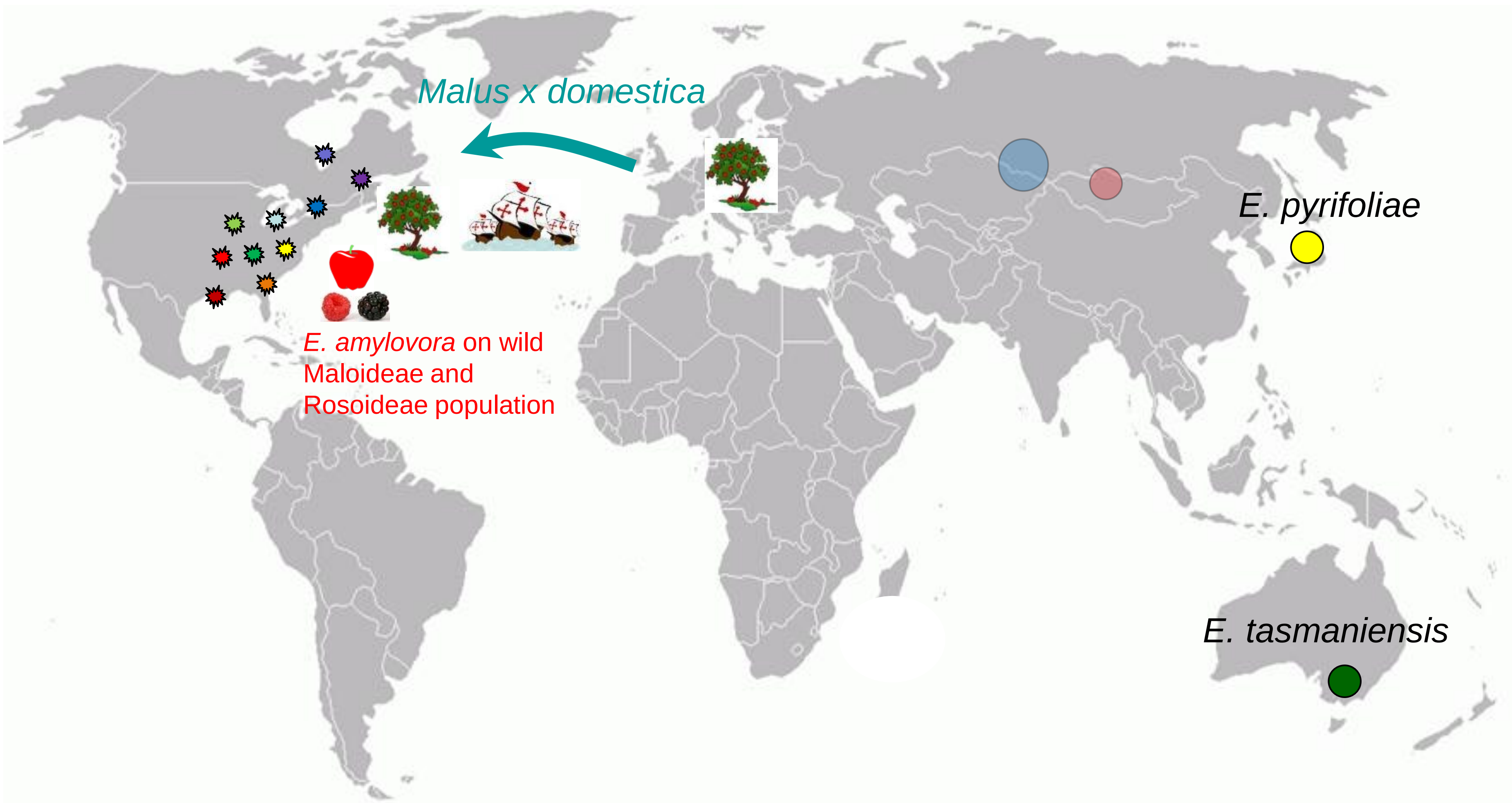


CRISPR-based dissemination model for *E. amylovora*

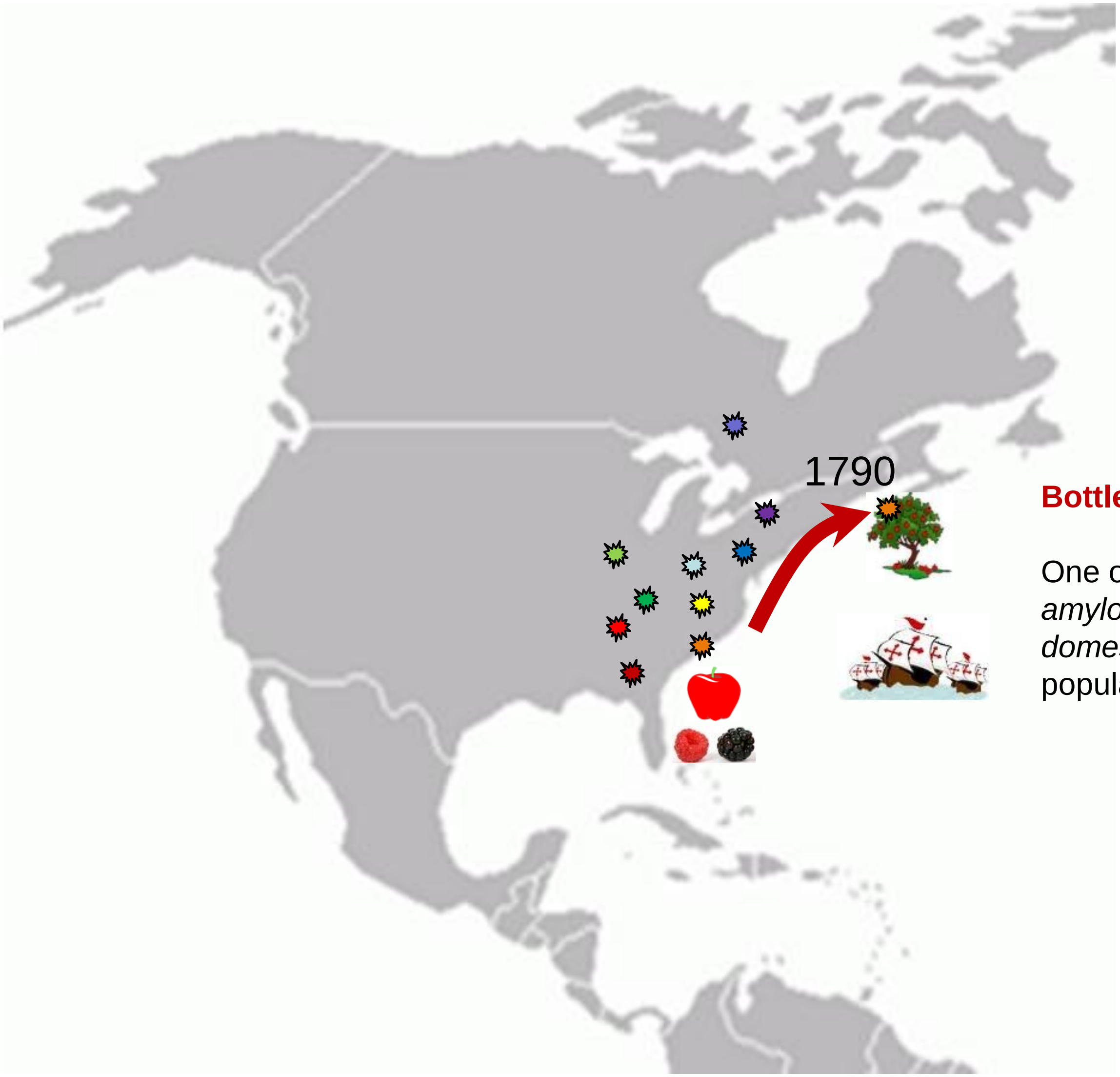
Molecular clock calibrated on whole genome data



Post-Ice age



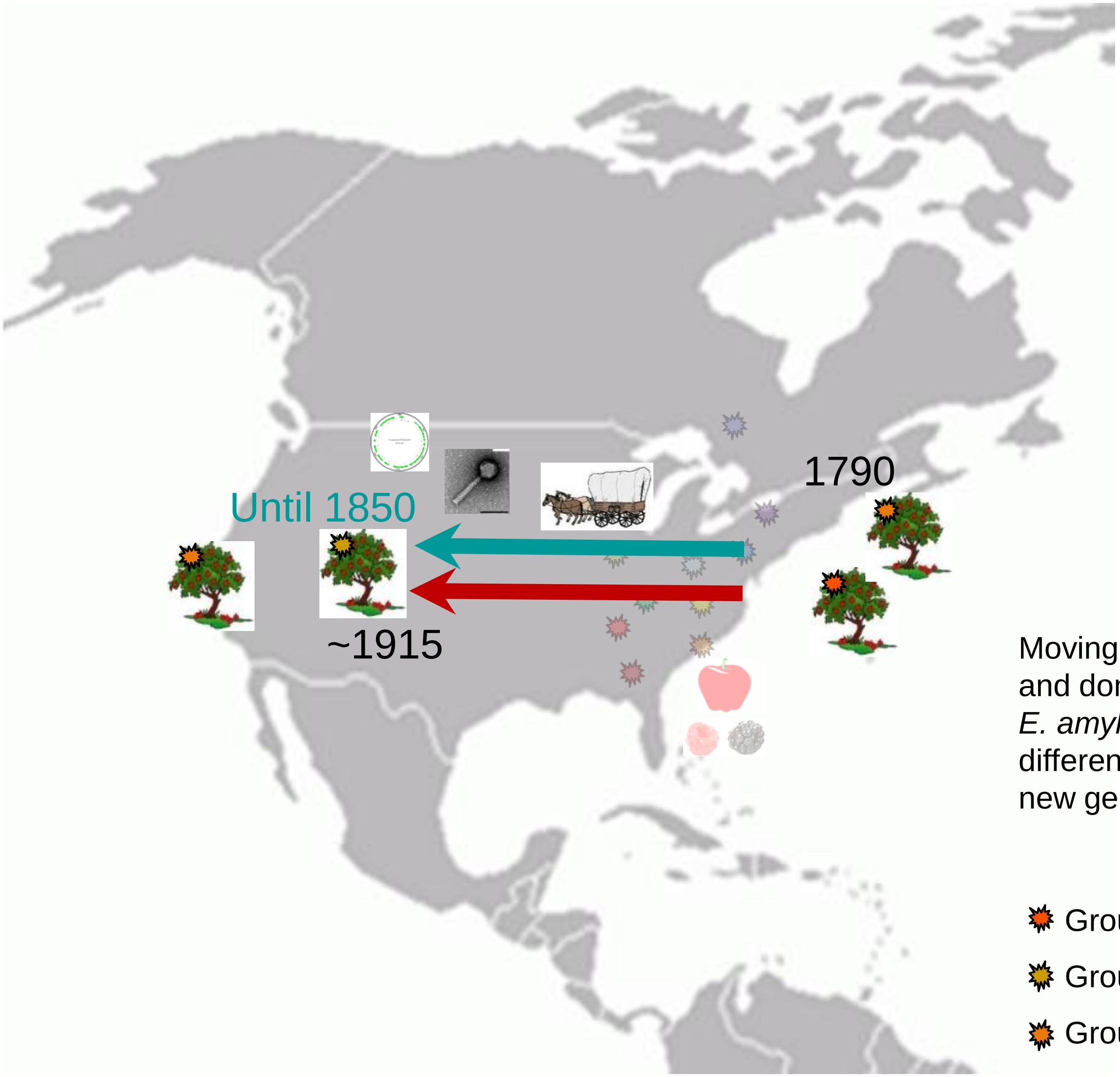
17th-18th century



Bottleneck #1

One or few major genotypes of *E. amylovora* are transferred on *Malus x domestica* from wild native *Maloideae* population

17th-18th century



Moving westward with settlers and domesticated apple trees *E. amylovora* slowly differentiates or incorporates new genotypes from wild plants

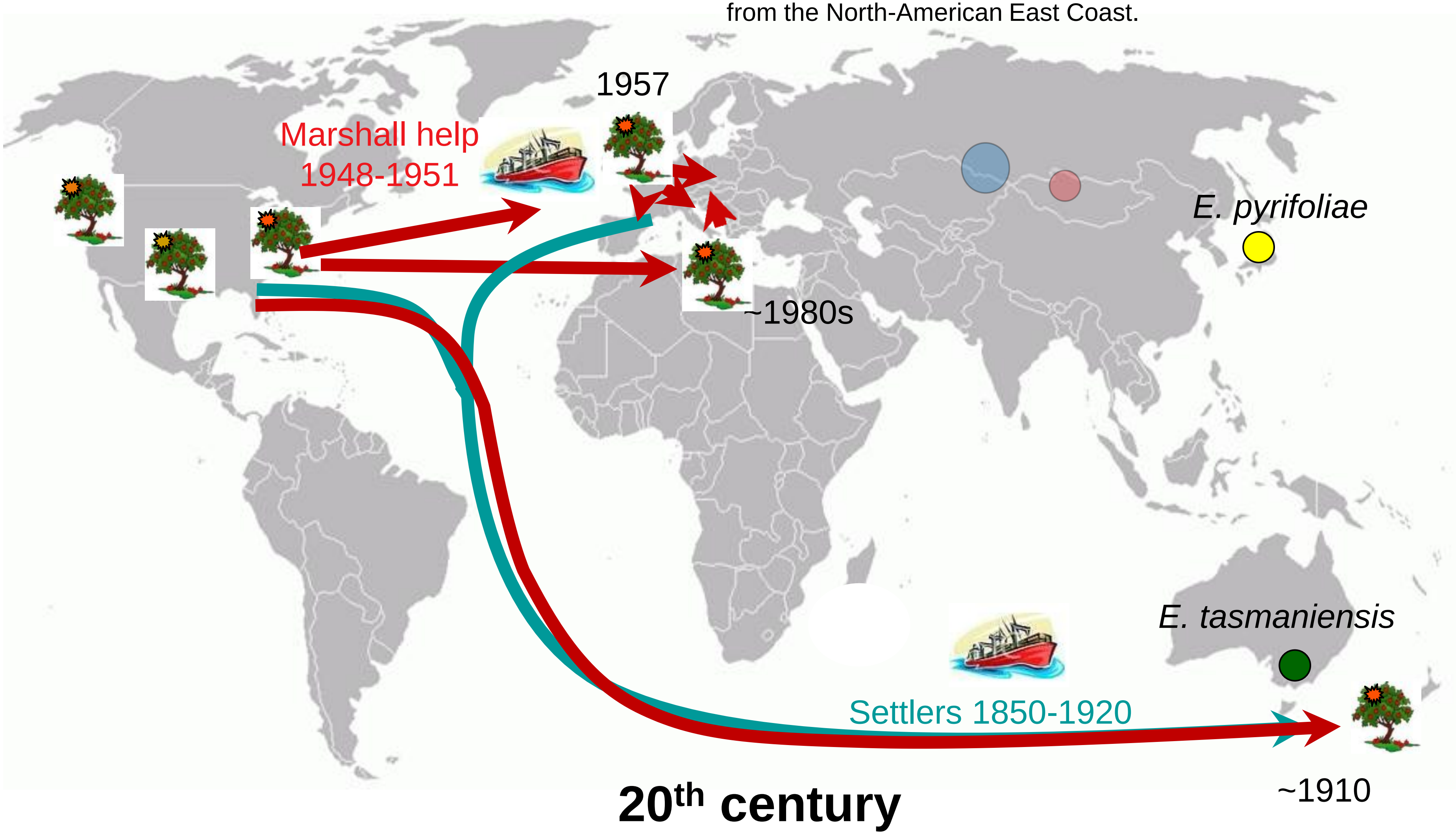
- Group I
- Group II
- Group III

19th century

Agroscope

Bottleneck #2

A single genotype (**belonging to Gr. Ia**) is found throughout the world, probably derived from few dissemination events originating from the North-American East Coast.



Worldwide dispersal of *E. amylovora* was shaped by two evolutionary bottlenecks

Rezzonico *et al*. (2011) AEM 77:3819-3829

- Diversity of *E. amylovora* strains isolated from *Rubus* spp. in North America is much higher if compared to strains isolated from *Maloideae* worldwide.

  > Group III genotype is closer to the common ancestor of *Maloideae* type of *E. amylovora*
  > Strain from southeastern US are closer to the to ancestral Gr. III genotype
  > IH 3-1 genetically intermediate between strains on undomesticated plants and fruit tree isolates

- Diversity on *Malus* spp. is higher in the center of origin (North America)

  > *Maloideae* strains enriched on *Malus domestica* from the pool present on undomesticated plants
  > Yet unexplored diversity must be present on undomesticated plant in the center of origin

- CRISPR-based clustering agrees with hypothesis of first outbreak in Europe/NZ caused by the dissemination of a single genotype from the U.S.

  > East-coast type strain(s) closely related to group Ia genotype present in Europe

Conclusions: VNTRs vs CRISPRs

- Both VNTRs and CRISPRs analysis are better suited to analyze the diversity of *E. amylovora* strains with respect to the molecular methods used so far

- VNTRs display more diversity and are suitable for source tracking at national, regional and local level

- CRISPRs show smaller diversity, but comprehend substantial chronological information that makes their use ideal for population analysis on a global scale

Acknowledgements

ACW Wädenswil: Theo H.M. Smits, Bea Schoch, Cosima Pelludat, Jürg & Beatrice Frey, Brion Duffy

NIB Ljubljana: Tanja Dreo

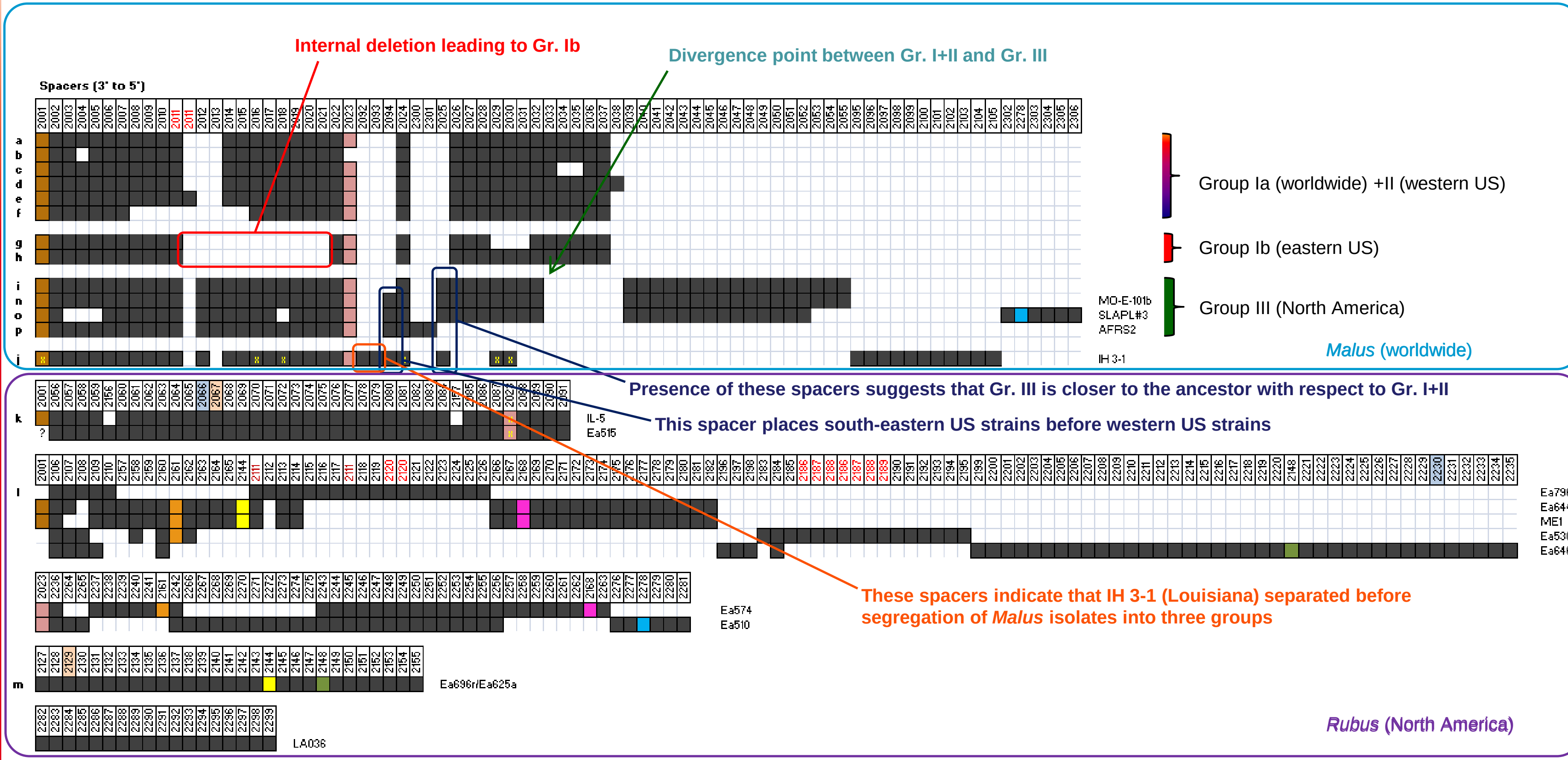
ETH Zürich: Yannick Born

Michigan State University: George Sundin, Gayle McGhee

Agroscope

CRISPR Region 2

What does it tell us?

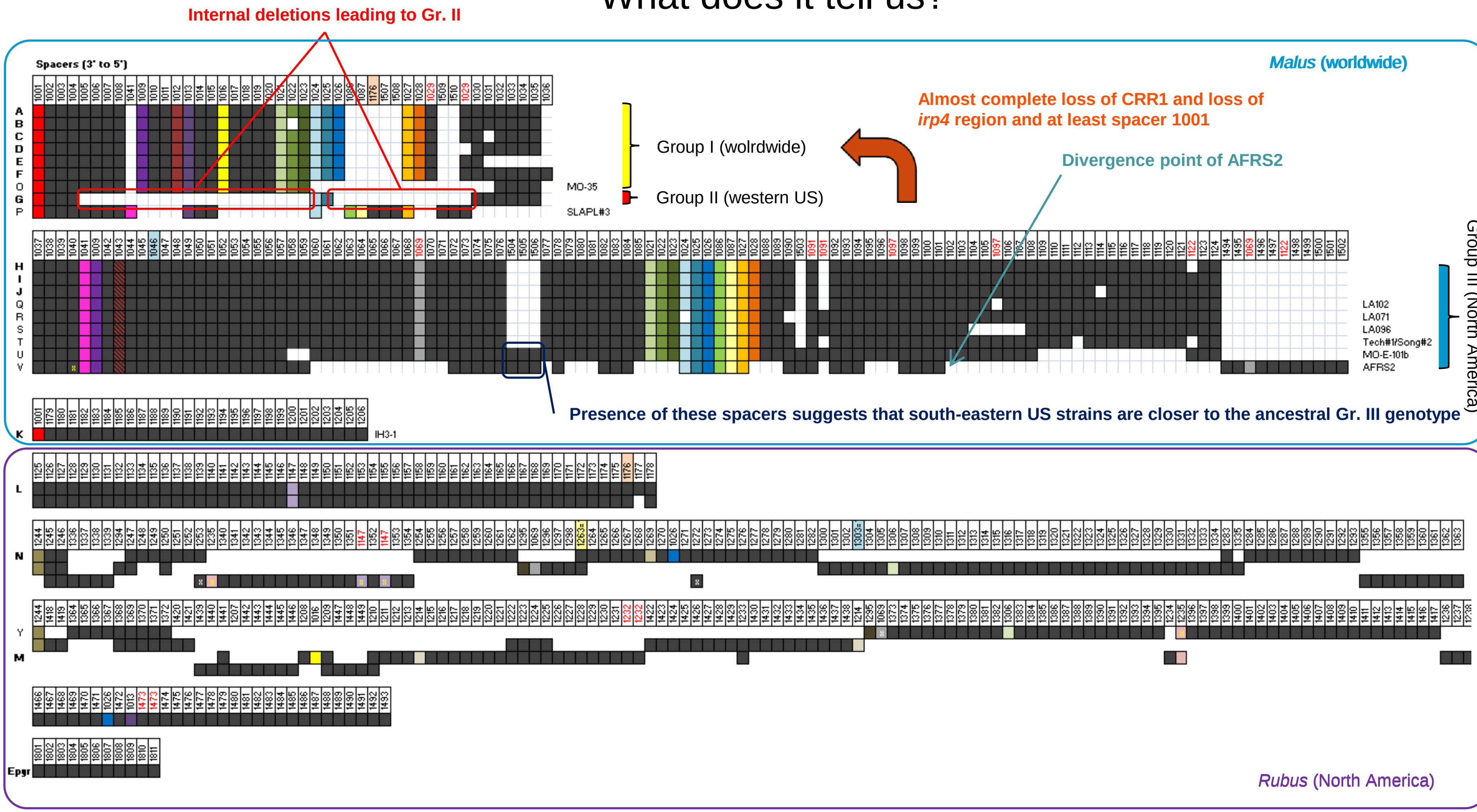


Diversity among *Rubus* isolates much higher than in *Malus* isolates

Agroscope

CRISPR Region 1

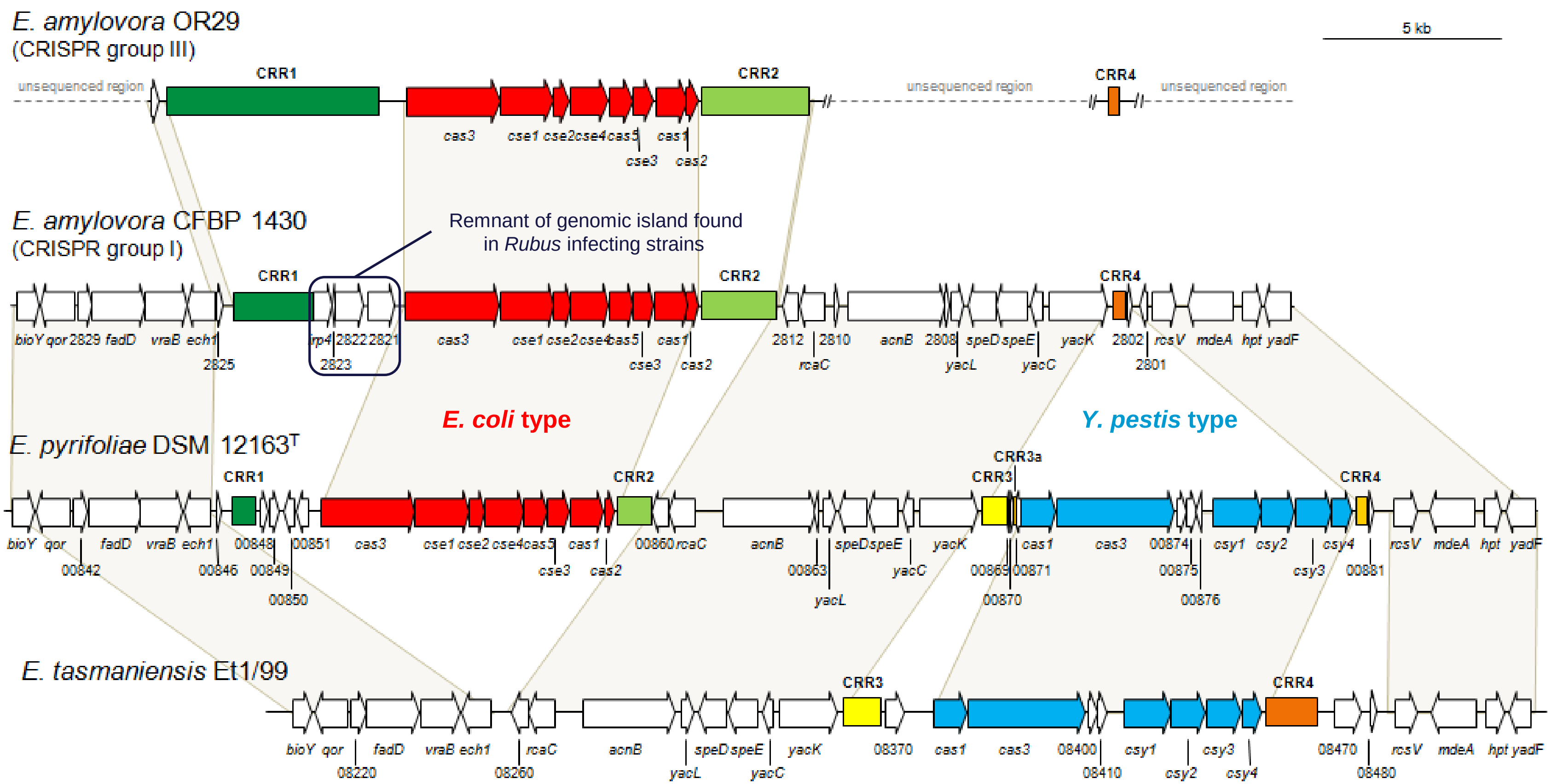
What does it tell us?



Diversity among *Rubus* isolates much higher than in *Malus* isolates

Agroscope

Structure of cas/CRISPR-regions in *Erwinia* spp.



Three separate CRISPR regions in *E. amylovora*

> PCR from the flanking genes and sequencing of the amplicon

> Cumulative data (spacer absence/presence) converted into a binary matrix

Agroscope

Evolution of *Erwinia* and *Pantoea*

HYPOTHESIS
based on genome data