Making sense of *P. infestans* diversity data at national and international scales

Themes GMO & IPM

David Cooke *et al.*

The James Hutton Institute

PathogenTraits

Where & When?

Database

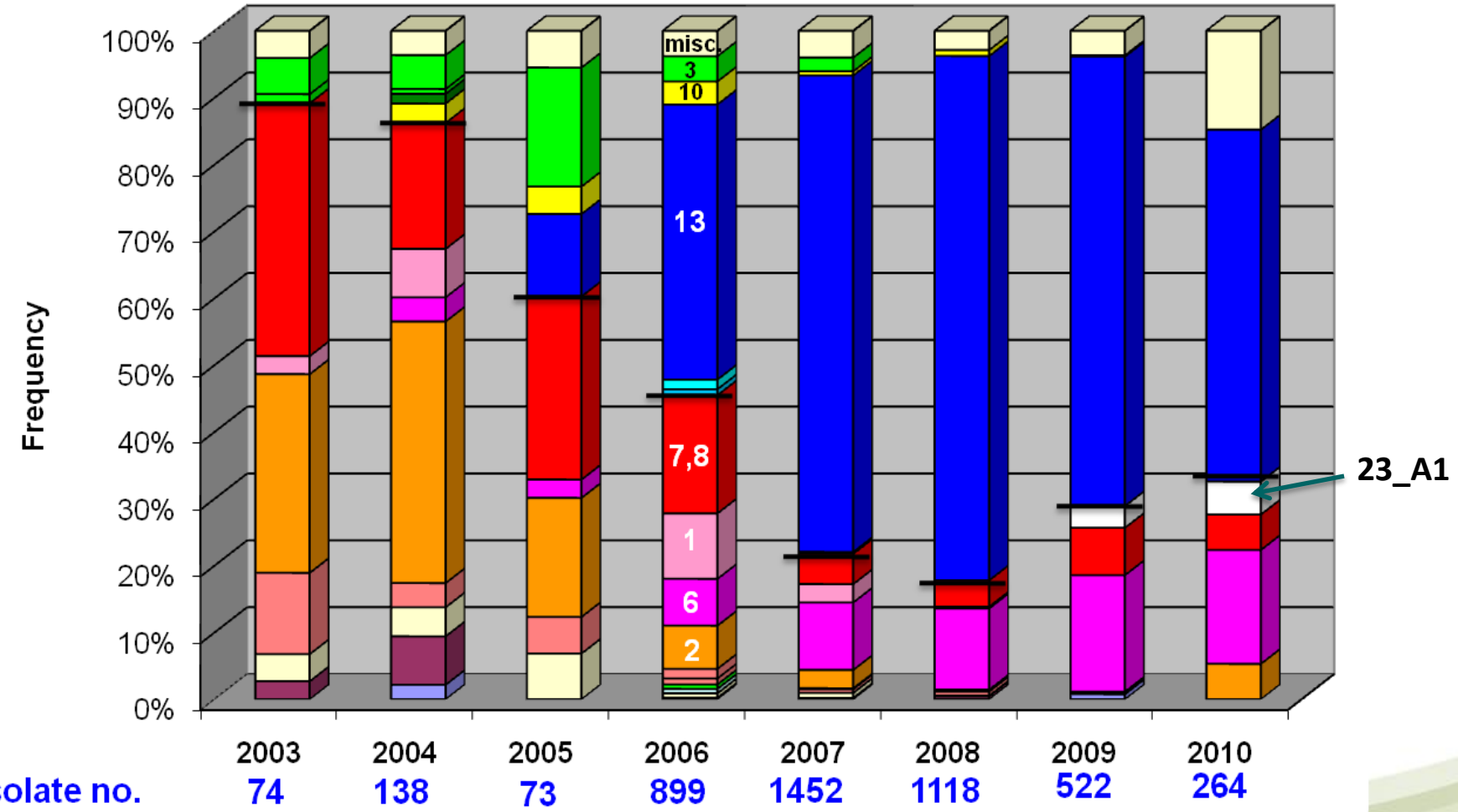
Mechanisms

- **DNA**
  - Markers
  - Gene sequence
  - Genome sequence
- **RNA**
  - NGS transcriptome
  - Microarray
- **Protein**
  - HPLC
  - Mass Spectrometry
- **Metabolites**
  - HPLC
  - Mass Spectrometry
- **Systems approach**
  - Modelling
  - Networks of responses
  - Key ‘hubs’ in interactome

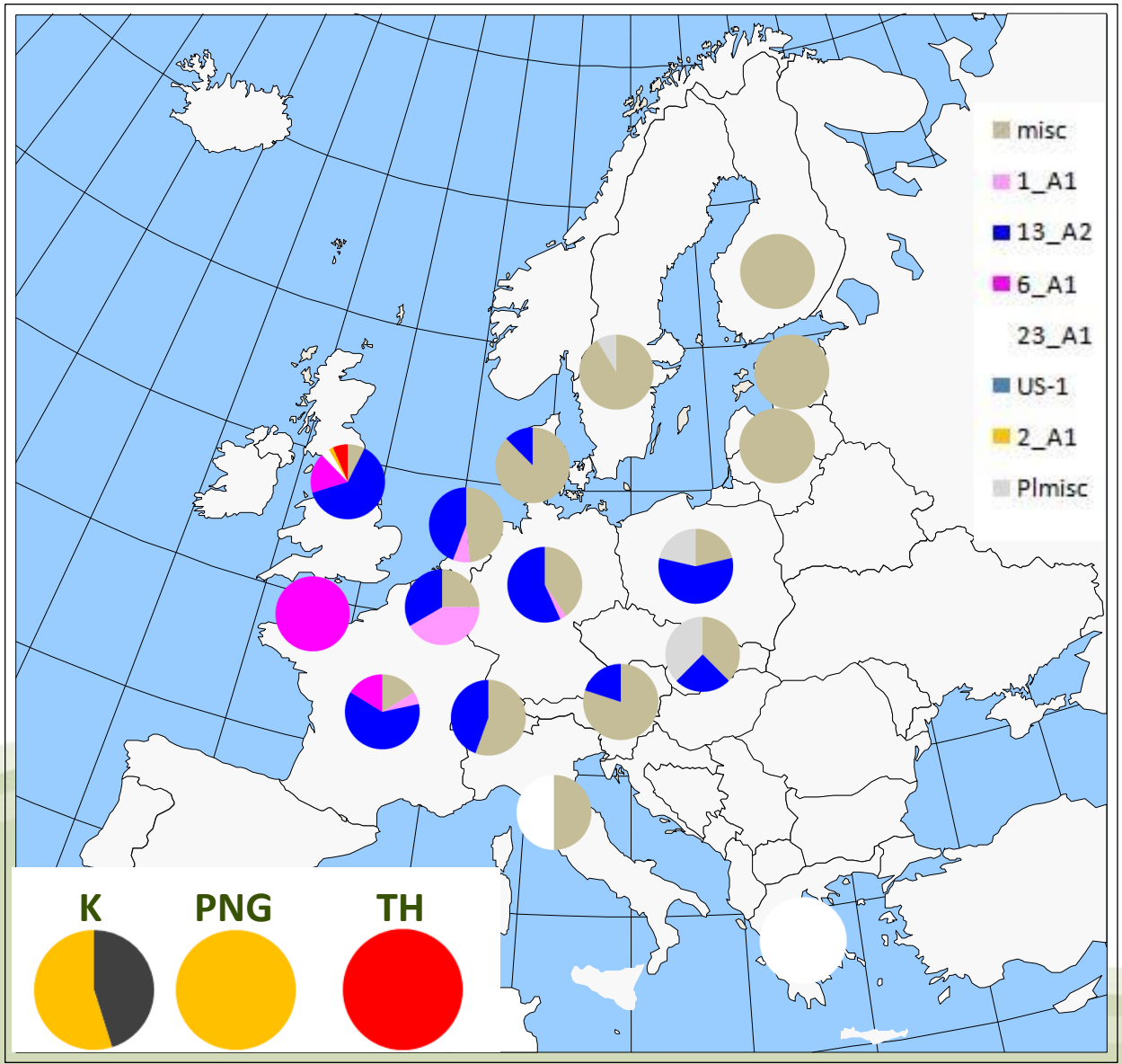
Outline

- GB & EU population status
- Database update & naming lineages/populations
- Progress on SSR markers and analysis
- Genomics and effector sequencing

GB *P. infestans* population update



Sample of EU *P. infestans* genotypes (2008-10) (*n*=350)



- 13_A2 still present in many areas (less due to reduced Metalaxyl use?)
- A high proportion of novel 'misc' types particularly in NE
- New genotype in PL, SK & SE
- 23_A1 on tomato in Italy & GK
- Pooling all EU data to come
- Thanks to Bayer (see poster) and Syngenta, Howard Hinds, Dolf deBoer, Vangelis Vellios

Database update

- EUCABLIGHT still active and can be updated using the P.exe tool developed by Jens Hansen & Poul Lassen (UoA)
- Integration underway into more comprehensive & redesigned ‘CropProblem’ database (cereal rust) with upgraded database techniques
- SSR data entry will be considerably easier in future with custom XL spreadsheets allowing upload of large datasets
- Improved mapping and analysis tools being developed for rust; aim to apply these to blight (funding needed)
- More discussion on Weds morning

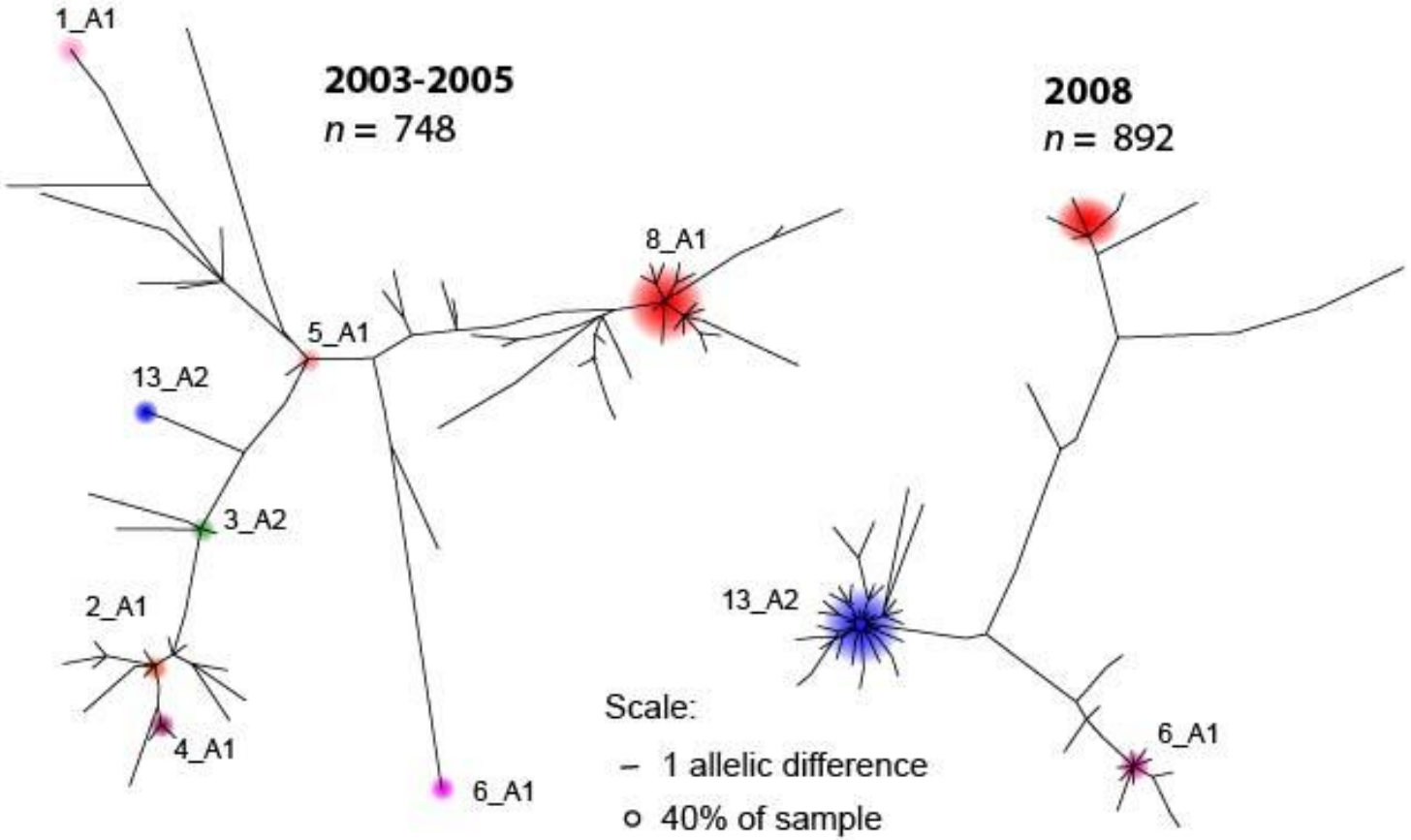
Genotype naming proposal

- New clonal lineages will emerge. But naming currently = *ad hoc*
- Single system would allow clearer scientific communication
- Ideally types should be defined by an objective method - calculated and applied *in silico*
- Is that possible? (see later)
- What marker system to use? SSRs best at the moment but sequencing may replace this in the future
- Naming convention (Cooke, Li and van der Lee)
  - EU_Y004_G0013_Y004_V001
  - EU_Y004_G0013_Y005_V002
  - EU_Y004_G0013_Y006_V003
  - CN_Y006_G0001_Y006_V001

SSR Analysis methods

- 12 markers in multiplex (PRI, JHI collaboration)
- Diploid and higher ploidies (multiple SSR peaks) in the same population makes population genetics difficult
- Challenges due to clonality and variants of clones
- A method that alleviates the ploidy problem based on Bruvo genetic distances published recently & implemented in R as POLYSAT (Clark & Jasieniuk *Molecular Ecology Resources 2011*)
- Minimum spanning trees & Principal component analysis output

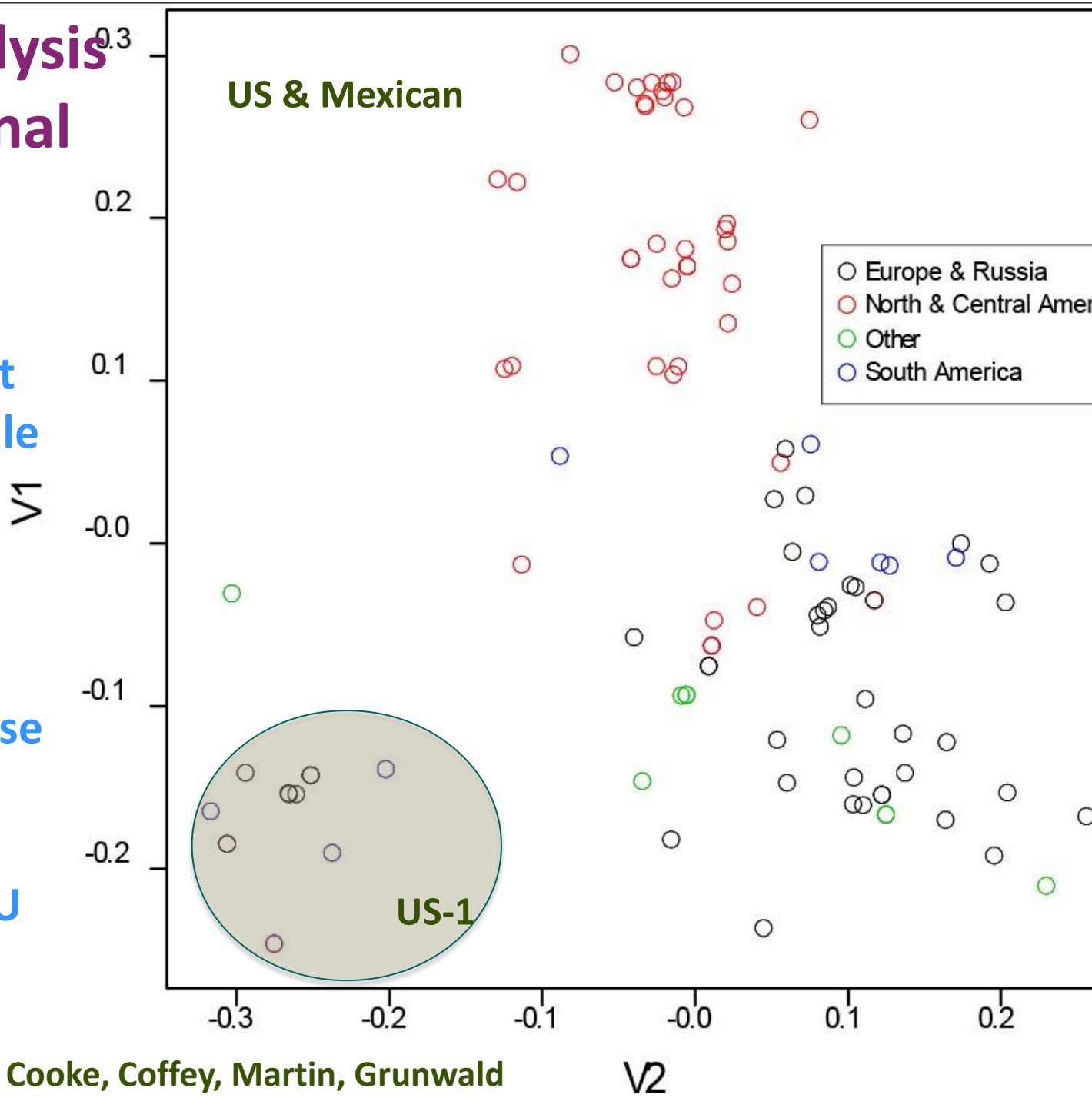
Bruvo distance and MINSP trees



Goss and Grunwald based on GB population data

POLYSAT analysis of international population (n=102)

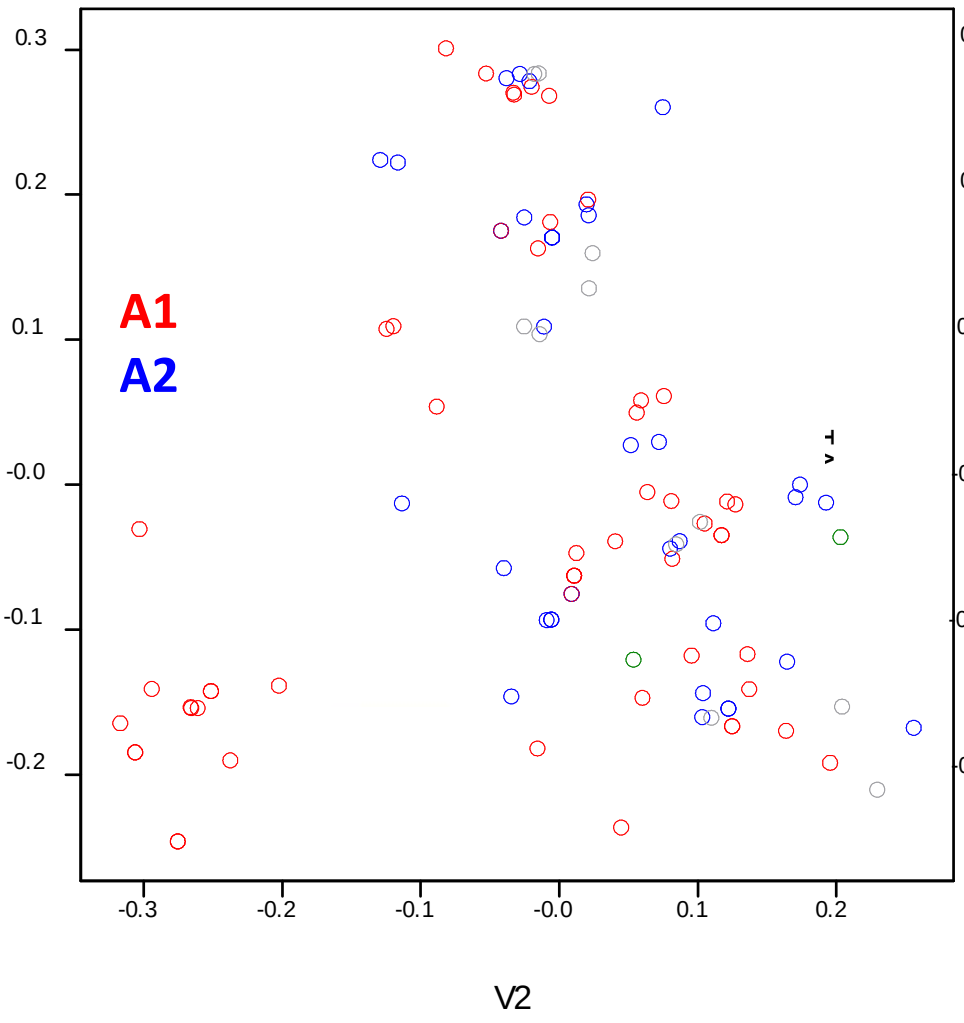
- Regional split at International scale
- Much diversity within regions
- US-1 now diverse
- MX and Andes contributed to EU diversity ?

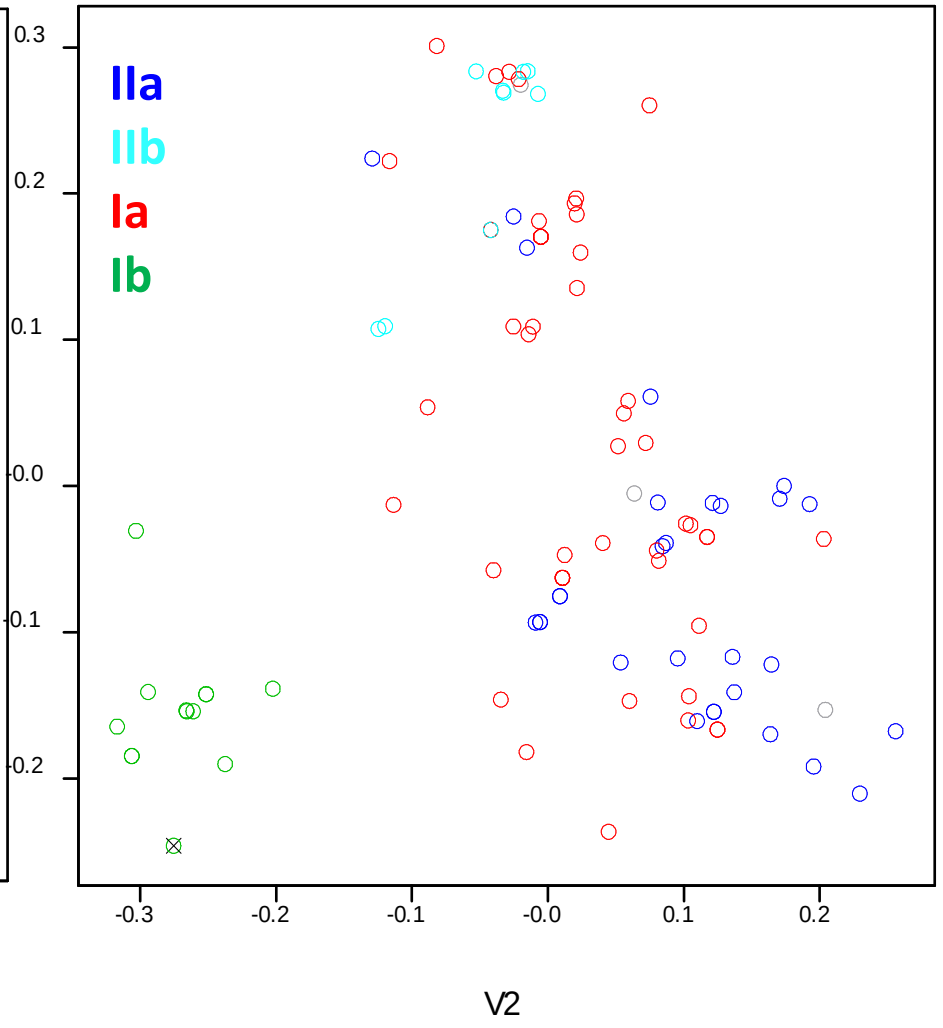


Cooke, Coffey, Martin, Grunwald

Global data by mating type & mtDNA

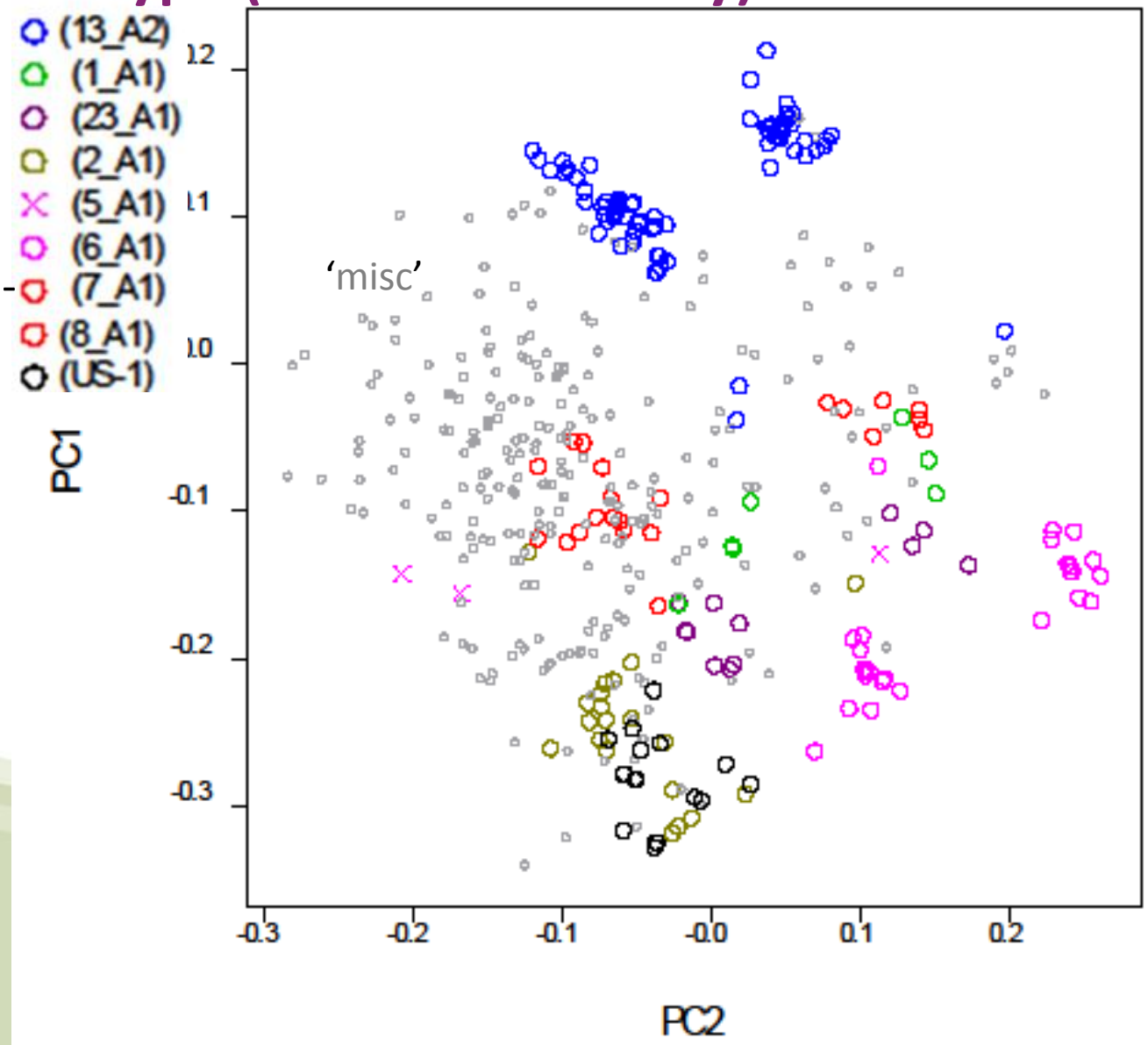
The James Hutton Institute





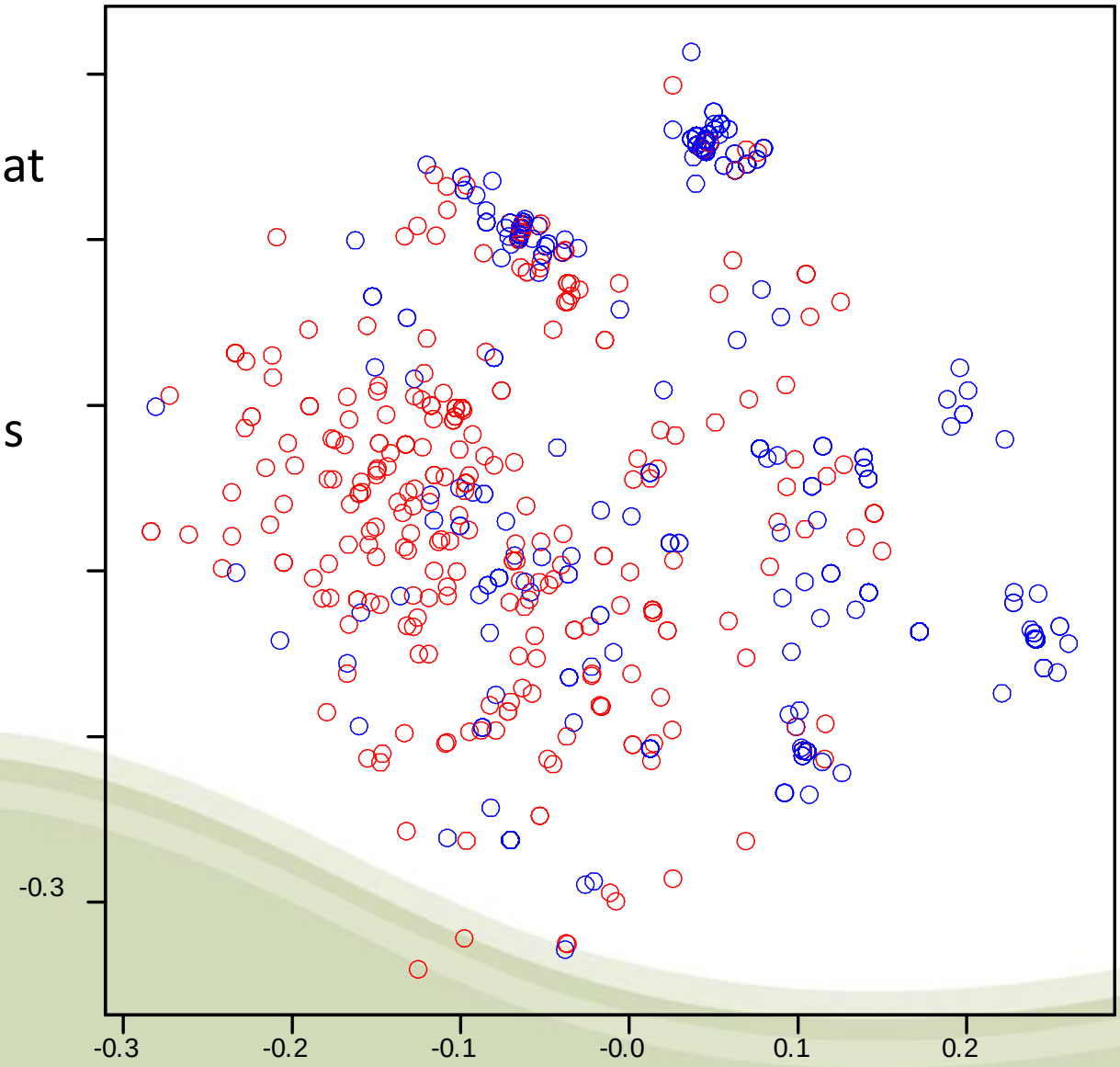
POLYSAT - Global population plus 1300 European isolates Labelled by genotype (defined manually)

- Reasonably well defined but split in each unexplained
- Add all ‘misc’ types the ‘diversity’ gaps are filled with recombinants
- Overlaps need to be looked at more carefully
- This clustering may not be suited to defining clones? MINSP trees better



The same data labelled by GB versus the rest of Europe

- Some overlap
- Also regions that appear quite distinct
- Further analysis needed



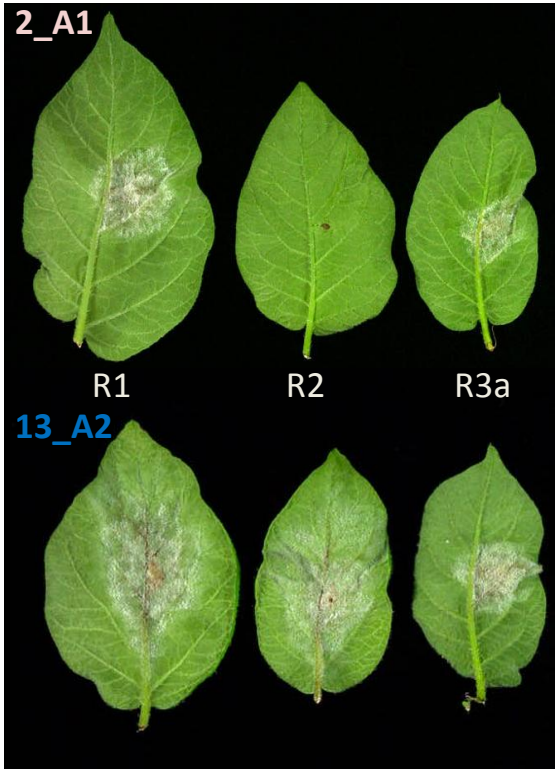
Clustering approaches have potential

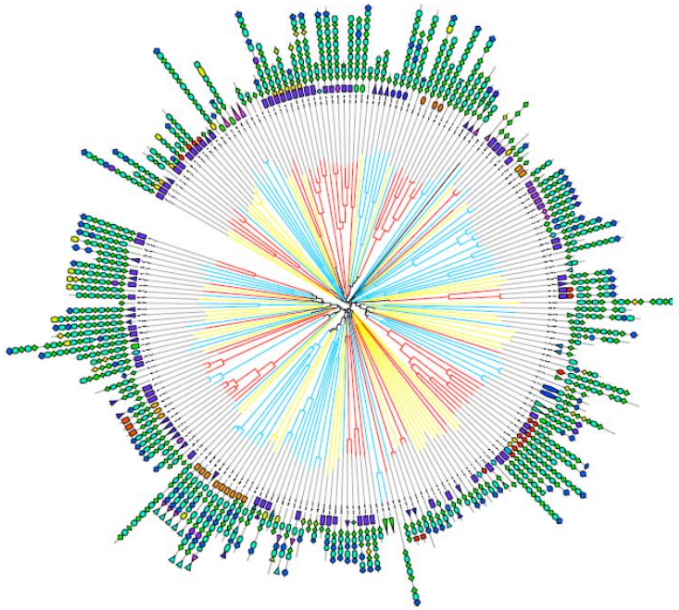
Dendroscope viewing tool looks useful

Bootstrapping tools needed to test statistical support for such analysis

Effector diversity

- Understanding effector diversity important for deployment new blight resistances (WUR method promising)
- Which effectors? 500+ RXLRs alone. At JHI candidates based on functional and localisation assays





Haas et al., *Nature* 2009

How to study effector diversity?

- PCR & test gene by gene
  - JHI testing new software (L. Pritchard) to design hundreds of primer pairs specific to single RXLR effector genes.
  - Presence/absence in genome
  - Sequence variation – define function
  - Collaborations needed
- Next Generation sequencing of DNA or cDNA transcripts

Example - RXLR gene PITG_06478 (RX5)

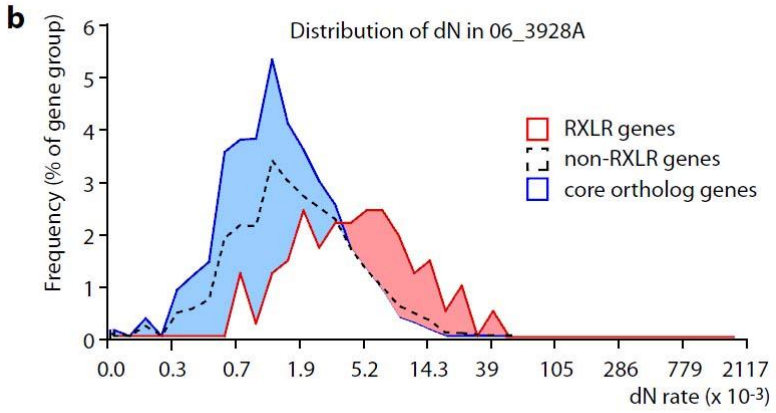
The James Hutton Institute

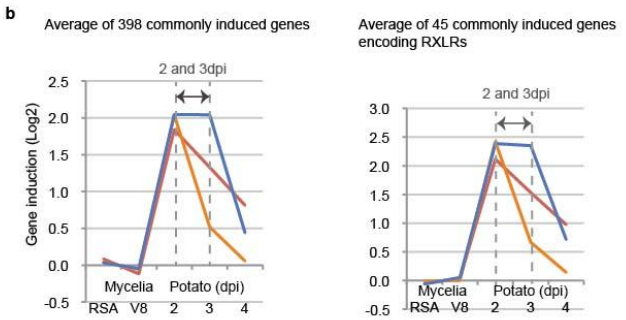
Presence of SNPs in PITG_06478 effector Gene (RX5)													
	SNP position (bp)												
Isolate	333	336	348	350	386	391	488	515	519	634	646	751	752
NL99013	C	A	R	R	R	T	C	A	C	G	T	W	Y
PITG_06478	C	G	G	G	G	T	C	A	C	A	T	A	T
NL03132	C	G	G	G	G	T	C	A	C	R	T	A	T
2009_7654A	C	G	G	G	G	T	C	A	C	R	T	A	T
NL07367	C	G	G	G	G	T	C	A	C	G	T	A	T
NL04255	C	G	G	G	G	T	C	A	C	G	T	A	T
NL03020	C	G	G	G	G	T	C	A	C	A	T	A	T
NL01008	C	G	G	G	G	T	C	A	C	A	T	A	T
EC-3527	C	G	G	G	G	T	C	A	C	R	T	A	T
2006_3936C2	C	G	G	G	G	T	C	A	C	R	T	A	T
2006_3928A	C	G	G	G	G	T	C	A	C	R	T	A	T
2006_3884B	C	G	G	G	G	T	C	A	C	R	T	A	T
T30-4	G	A	A	A	A	A	T	C	C	G	T	T	C
NL01155	G	A	A	A	A	A	T	C	C	G	T	T	C
NL03114	G	A	A	A	A	A	T	C	C	G	Y	T	C
NL04223	G	A	A	A	A	A	T	C	C	G	Y	T	C
NL04297	G	A	A	A	A	A	T	C	C	G	T	T	C
NL02037	G	A	A	A	A	A	T	C	C	G	T	T	C
2004_10477B	G	A	A	A	A	A	T	C	C	G	T	T	C
T30-4	G	A	A	A	A	A	T	C	C	G	T	T	C
MP_622	G	A	A	A	A	A	T	C	C	G	T	T	C
2006_4024E	G	A	A	A	A	A	T	C	C	G	T	W	Y
MP_618	G	A	A	A	A	A	T	C	C	G	T	T	C
SE_03087	G	A	A	A	A	A	T	C	C	G	T	T	C
2006_4332	G	A	A	A	A	A	T	C	C	G	Y	T	C
2004-7804B	G	A	A	A	A	A	T	C	C	G	T	T	C
2006_4244E	S	A	R	R	A	W	Y	M	C	G	T	W	Y
2003_25_3_1	S	R	R	R	A	W	Y	M	C	R	T	W	Y
IPO01900	S	R	R	R	G	T	Y	M	Y	R	T	W	Y
88069	S	R	R	R	R	T	C	A	C	G	Y	W	Y
80029	S	R	R	R	R	W	Y	M	C	R	T	W	Y
IPO98014	S	R	R	R	R	W	Y	M	C	R	T	W	Y
NL00026	S	R	R	R	R	W	Y	M	C	G	T	W	Y
NL05548	S	R	R	R	R	W	Y	M	C	G	T	W	Y
88133	S	R	R	R	R	W	Y	M	C	R	T	W	Y
NL02085	S	R	R	R	R	W	Y	M	C	R	T	W	Y
2006_4168C	S	R	R	R	R	W	Y	M	C	G	Y	W	Y
2006_3992G	S	R	R	R	R	W	Y	C	C	G	T	W	Y
2006_4256B	S	R	R	R	R	W	Y	M	C	G	Y	W	Y
2006_3984C	S	R	R	R	R	W	Y	M	C	R	T	W	Y
2006-4012F	S	R	R	R	R	W	Y	M	C	G	T	W	Y
2006-3888A	S	R	R	R	R	W	Y	M	C	R	T	W	Y
2006_4100A	S	R	R	R	R	W	Y	M	C	R	T	W	Y
2006_4232E	S	R	R	R	R	W	Y	M	C	G	Y	W	Y

- 13 SNPs identified
- Seven replacement changes observed
- Homozygous and heterozygous forms found in European population
- Starting to examine the evolutionary history of individual and collective RXLR diversity
- Spatial and temporal patterns will be important too so we need to build database that holds such info – merging bioinformatics with phylogeography

P. infestans 13_A2 genomics

- 06_2928A sequenced and compared to five others
- Many SNPs identified –RXLRs in particular under +ve selection
- Expanded effector repertoire & many CNVs seen
- Gene expression patterns of 13_A2 differ from other isolates





Cano, Rafaelle, Kamoun The Sainsbury Lab

Conclusions & Thanks

The James Hutton Institute

- We have not yet made sense of all *P.infestans* diversity !
- *P. infestans* population diverse but structured
- Analysis tools improving
- Database updates and better interpretation at regional and Intnl scale required
- We need to understand the drivers of popn change better better links to fundamental research on effectors and R-genes)

- Thanks to all current and future data submitters and collaborators, friends and colleagues