Population Diversity of Xanthomonas oryzae pv oryzicola TAL Effectors and their Candidate Targets

Katie Wilkins

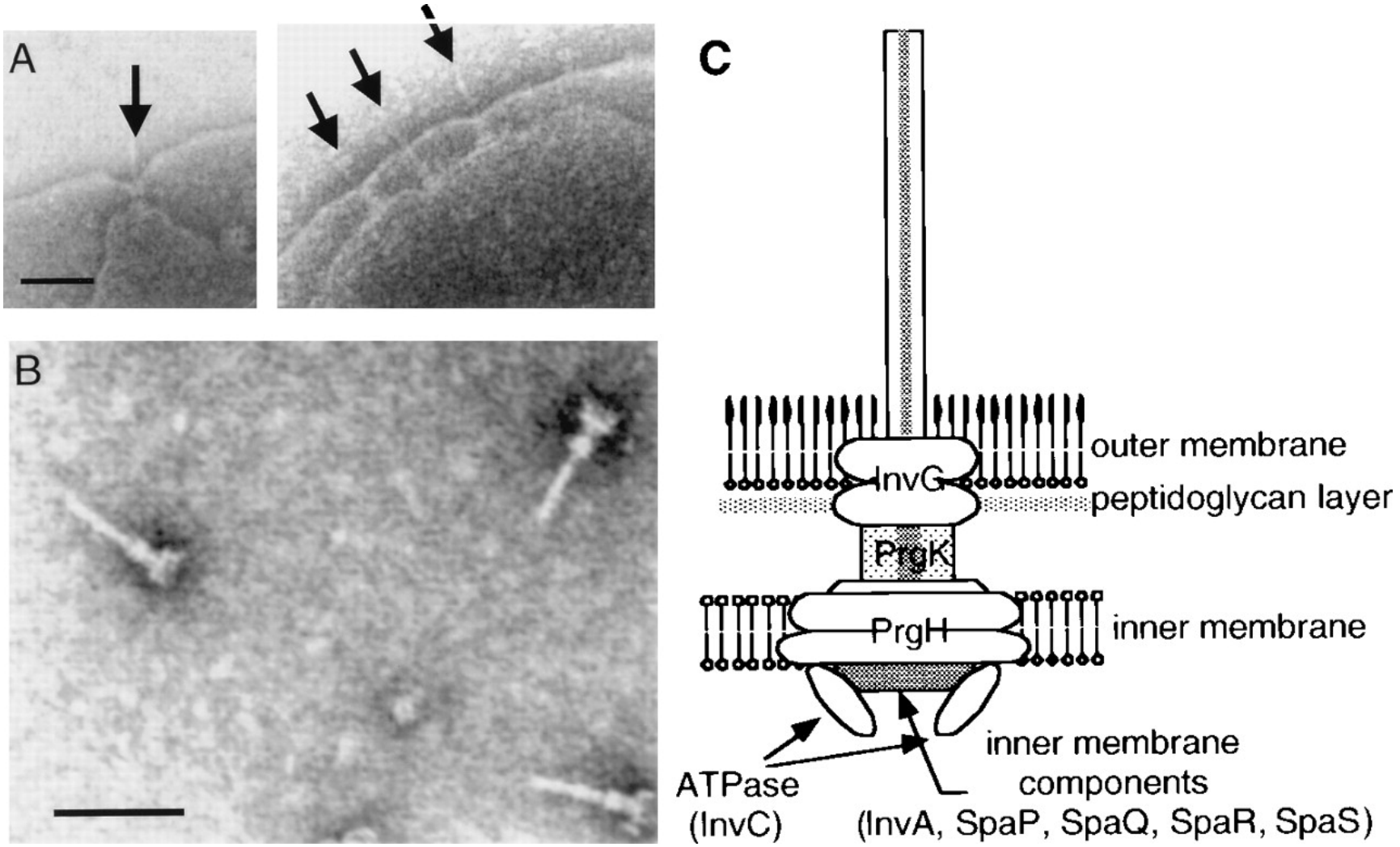
Outline

- Background
  - Biology
  - Computational problem
- Motivation
- Results
- Future work

Motivation

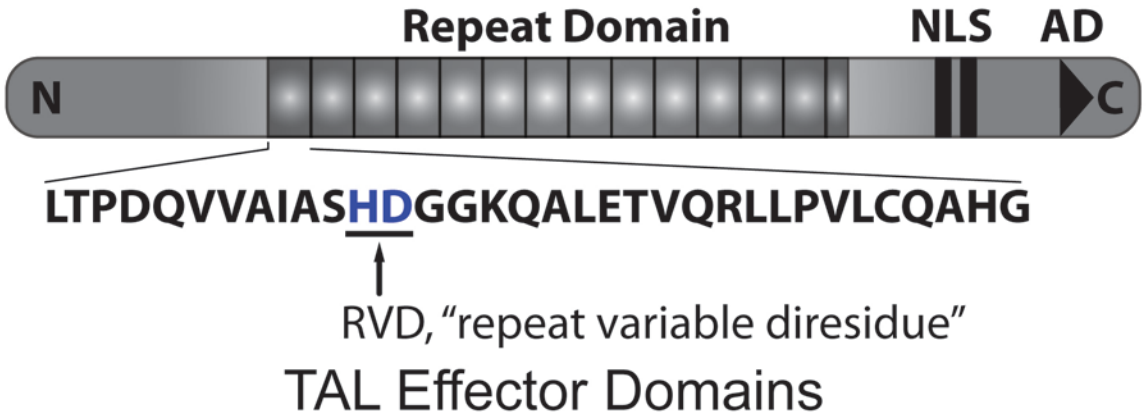
- *Xanthomonas* species cause diseases of >400 plant species

Type III Effectors

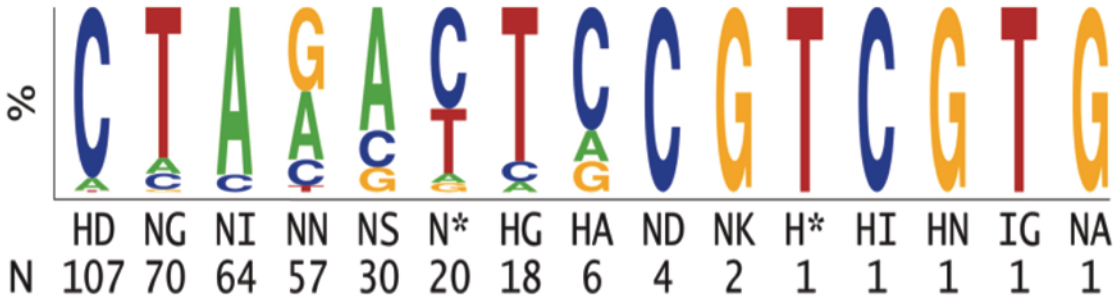


J E Galán, and A Collmer Science 1999;284:1322-1328

What is a TAL Effector?



TAL Effector Domains

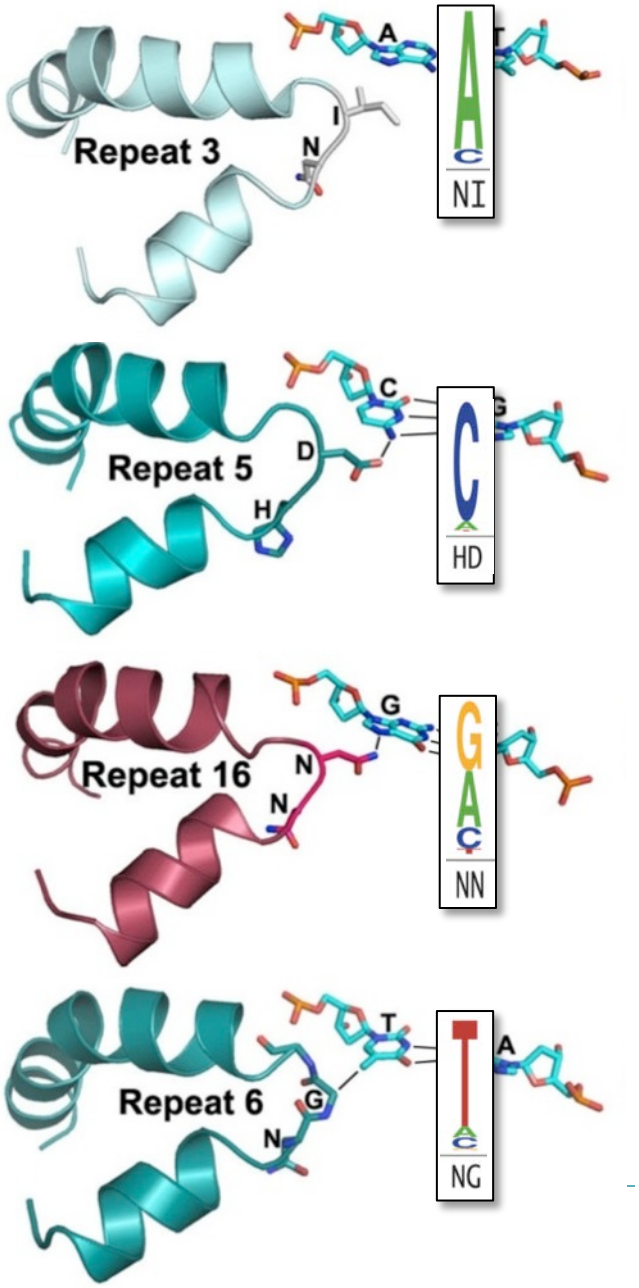


Observed RVD-Nucleotide Association Frequencies

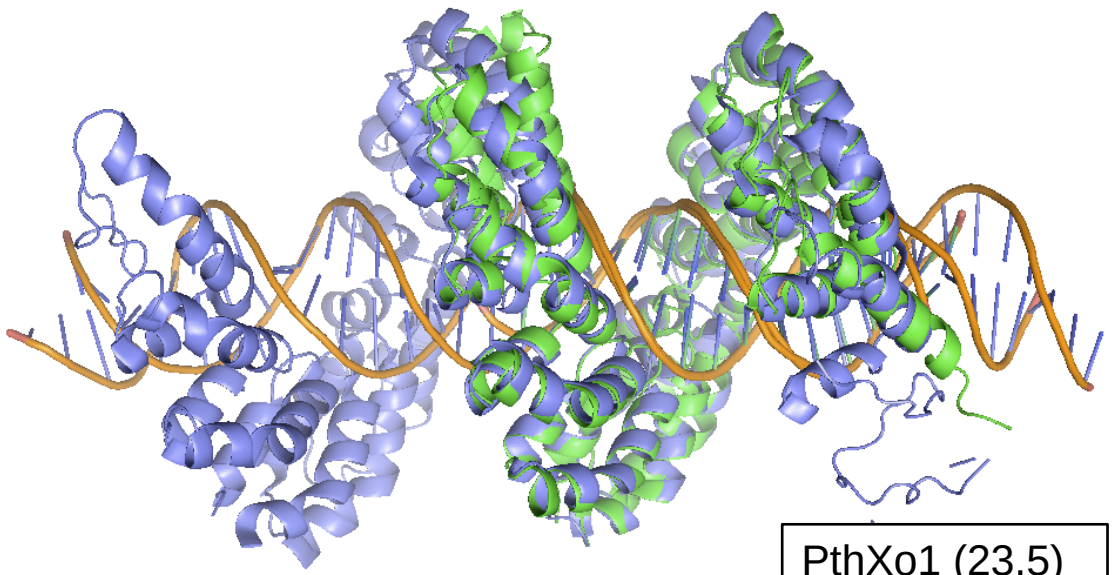
Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009 Dec 11;326(5959):1501. doi: 10.1126/science.1178817. Epub . PubMed PMID: 19933106.

Repeats form a superhelix that tracks the DNA major groove

Residue 13 contacts explain RVD specificities



Repeat structure is highly conserved ("LEGO blocks")



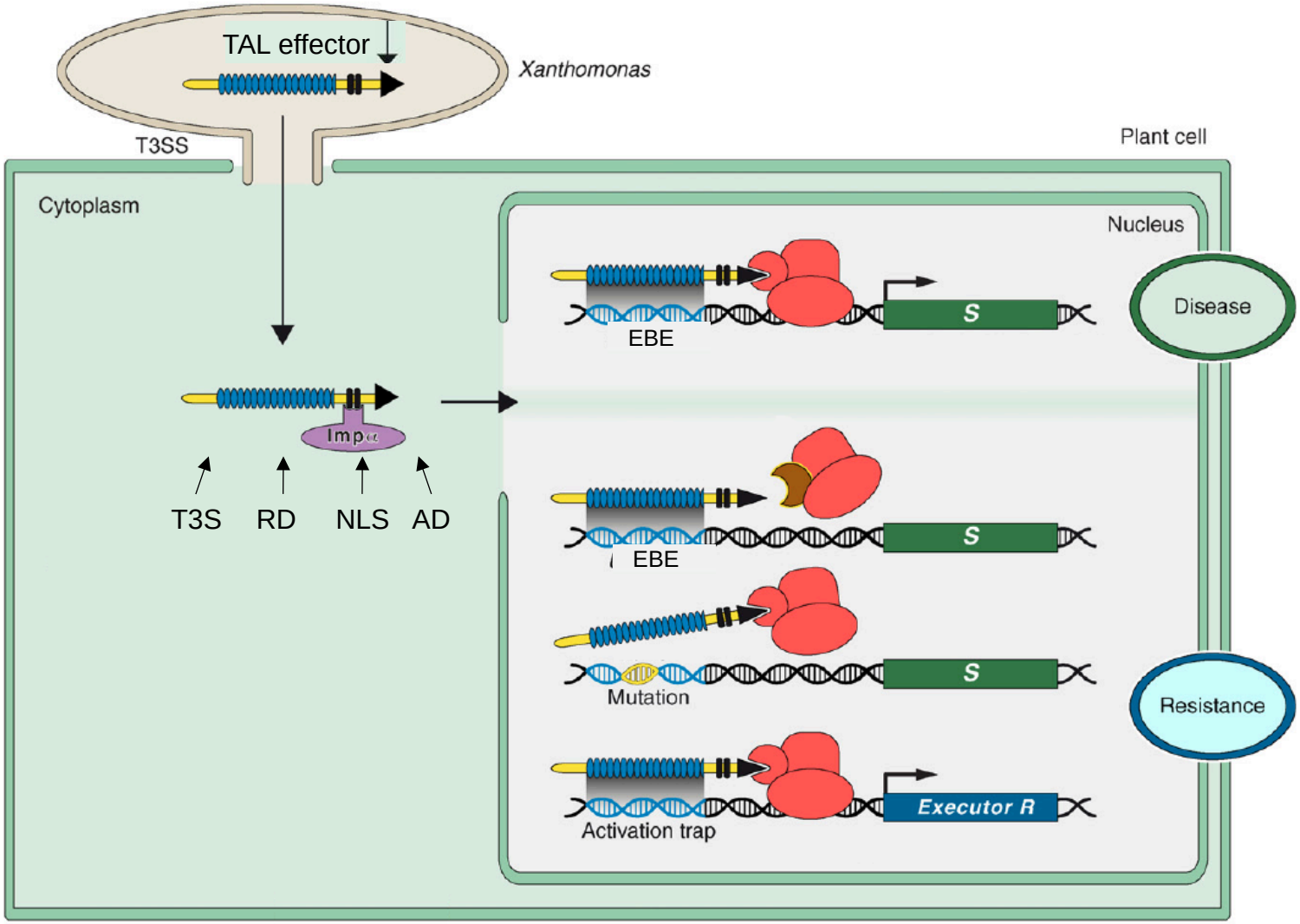
PthXo1 (23.5)
dHax3 (11.5)
Repeats 1-11
r.m.s.d ~0.8Å

PthXo1 and dHax3

dHax3 structure: Deng *et al*. (2012) Science

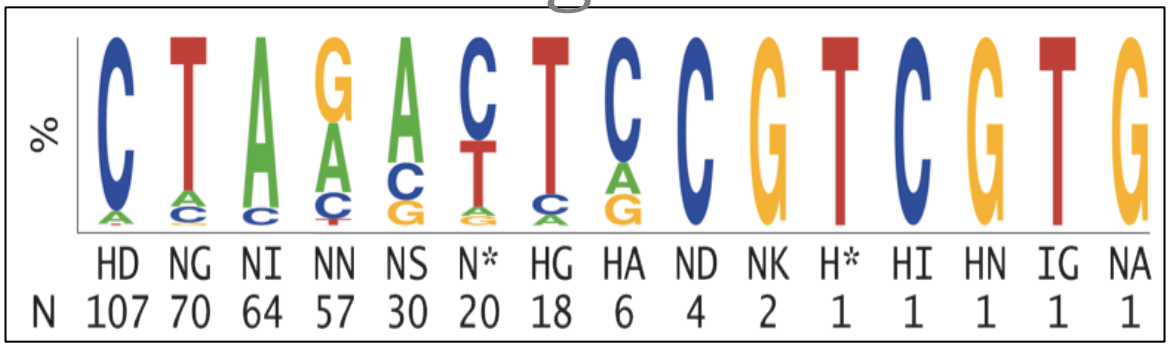
Amanda Mak

TAL effectors in disease



Bogdanove et al. (2010) *Curr. Opinion Plant Biol*

Predicting TAL Effector Function



RVD	Total	A	C	G	T
HD	107	0.084	0.858	0.025	0.033
NI	64	0.841	0.109	0.025	0.025
NG	70	0.102	0.102	0.038	0.758
NN	57	0.357	0.151	0.436	0.057
NS	30	0.625	0.205	0.145	0.025
N*	20	0.070	0.520	0.070	0.340
HG	18	0.075	0.125	0.025	0.775
HI	1	0.025	0.925	0.025	0.025
Other	0	0.250	0.250	0.250	0.250

$$score = \sum -\log(RVD : nt.observed.frequency)$$

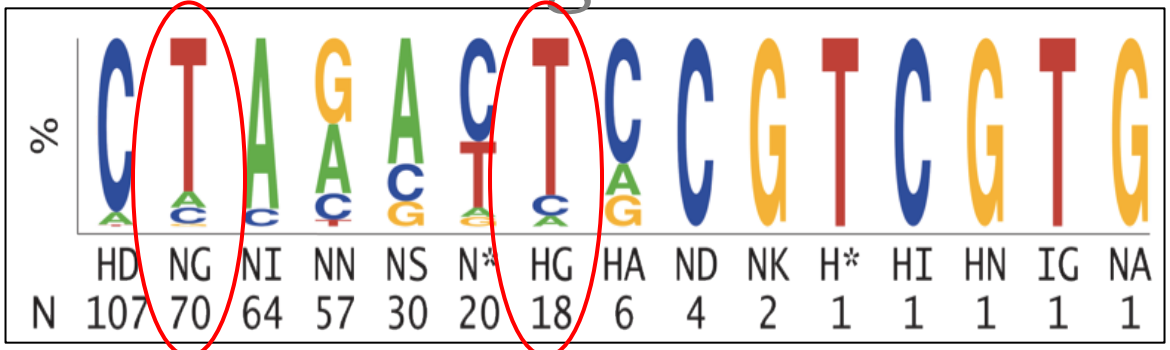
Observed RVD-nt association frequencies

Convert to probabilities

Sum the –log probability for each RVD-nt pair in an alignment

Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009 Dec 11;326(5959):1501. doi: 10.1126/science.1178817. Epub . PubMed PMID: 19933106. Slide from Adam Bogdanove.

Predicting TAL Effector Function



RVD	Total	A	C	G	T
HD	107	0.084	0.858	0.025	0.033
NI	64	0.841	0.109	0.025	0.025
NG	70	0.102	0.102	0.038	0.758
NN	57	0.357	0.151	0.436	0.057
NS	30	0.625	0.205	0.145	0.025
N*	20	0.070	0.520	0.070	0.340
HG	18	0.075	0.125	0.025	0.775
HI	1	0.025	0.925	0.025	0.025
Other	0	0.250	0.250	0.250	0.250

$$score = \sum -\log(RVD : nt.observed.frequency)$$

Observed RVD-nt association frequencies

Convert to probabilities

Sum the –log probability for each RVD-nt pair in an alignment

Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009 Dec 11;326(5959):1501. doi: 10.1126/science.1178817. Epub . PubMed PMID: 19933106. Slide from Adam Bogdanove.

Is it really that simple?

- Many false positive predictions of natural TAL effector binding sites
- Occasionally designer TAL effectors don't work

Factors influencing binding

- Location in genome
- Context effects
- Epigenetic status
- Chromatin status

Factors influencing binding

- Location in genome
- Context effects
- Epigenetic status
- Chromatin status

Naïve Bayes classifier construction

- Features:
  - Score
  - Relative Score
  - Rank
  - Transcriptional Start Site
  - Translational Start Site
  - TATA Box Location
  - Y Patch Location

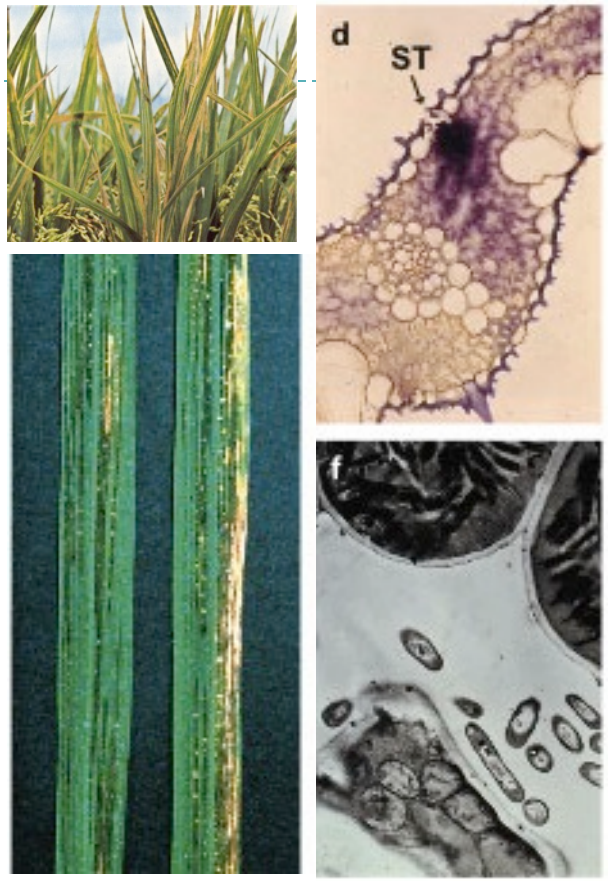
Naïve Bayes classifier performance

- Leave-one-out cross validation

Accuracy	Precision	Recall	F measure	MCC	AUC
.89	.88	.92	.90	.77	.88

- "The best classifier that resulted…eliminates 85% (17/20) of the falsely predicted targets at a cost of less than 10% (2/24) of the real ones."

Bacterial blight of rice



X. oryzae pv. oryzicola

"Xoc"

(leaf streak)

26 TAL effectors (strain BLS256)

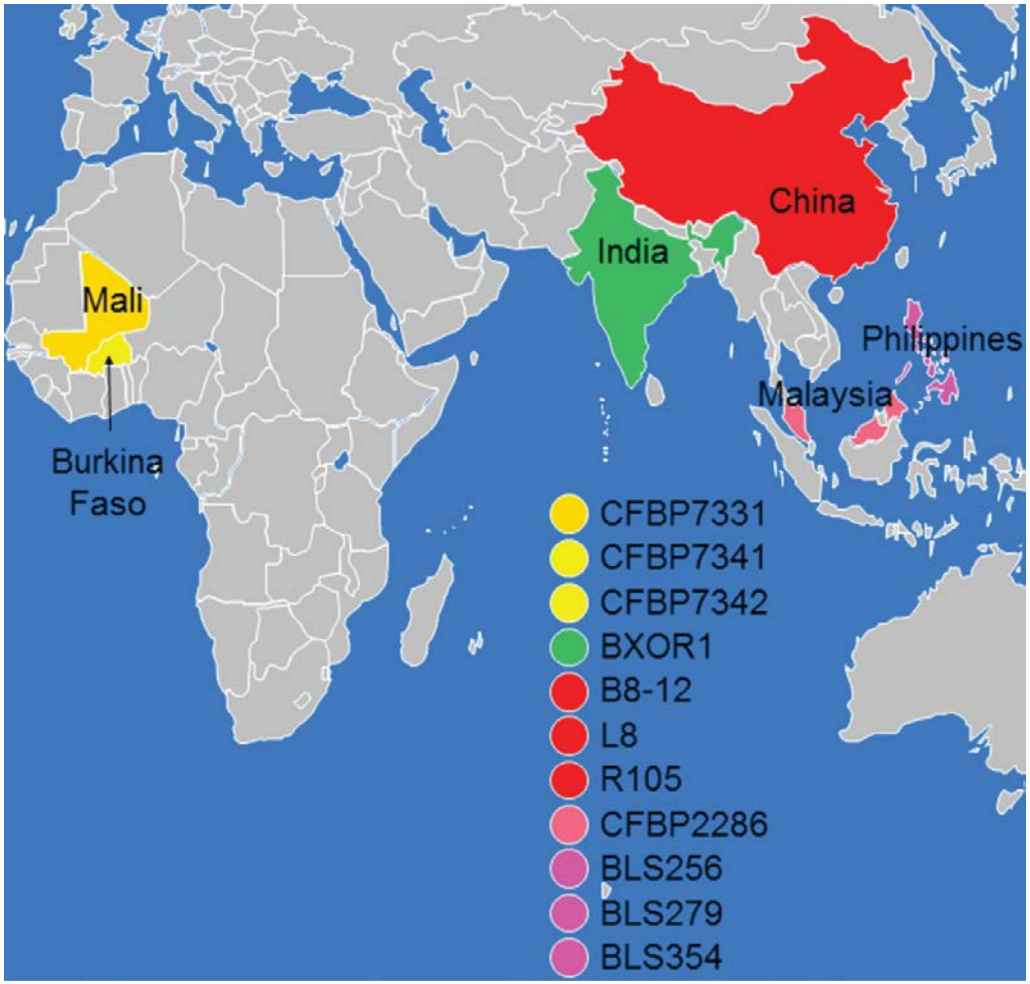
Bogdanove *et al.* (2011) J. Bacteriol.

Objectives

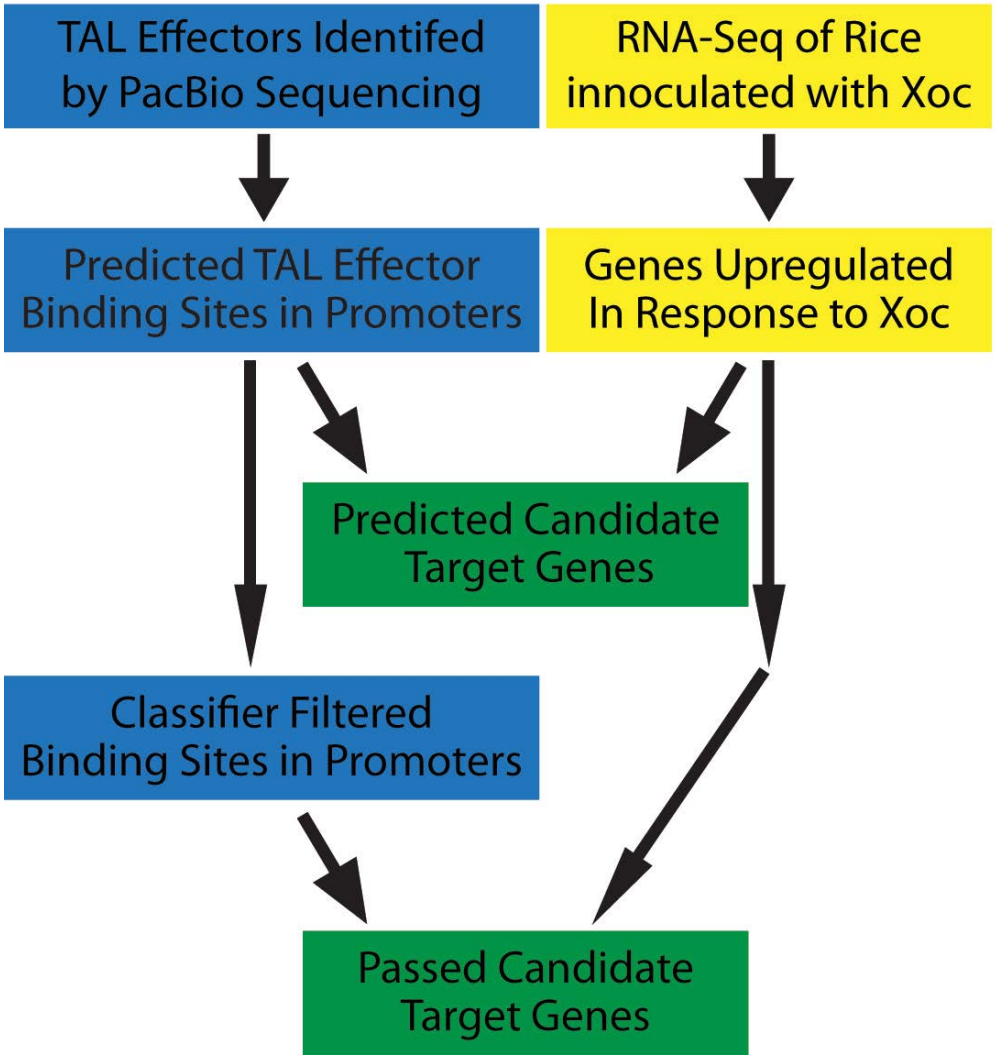
- Compare TAL effector content and rice transcriptional response to 10 geographically diverse isolates of Xoc
- Identify TAL effectors and candidate targets that suggest strategies for developing broadly resistant rice varieties
- Determine whether Xoc TAL effectors are horizontally transferred

Data

- PacBio Sequencing of Xoc genomes
- RNA-Seq of Xoc inoculated rice



Workflow for candidate target identification



TAL effector sequences of the 10 Xoc strains

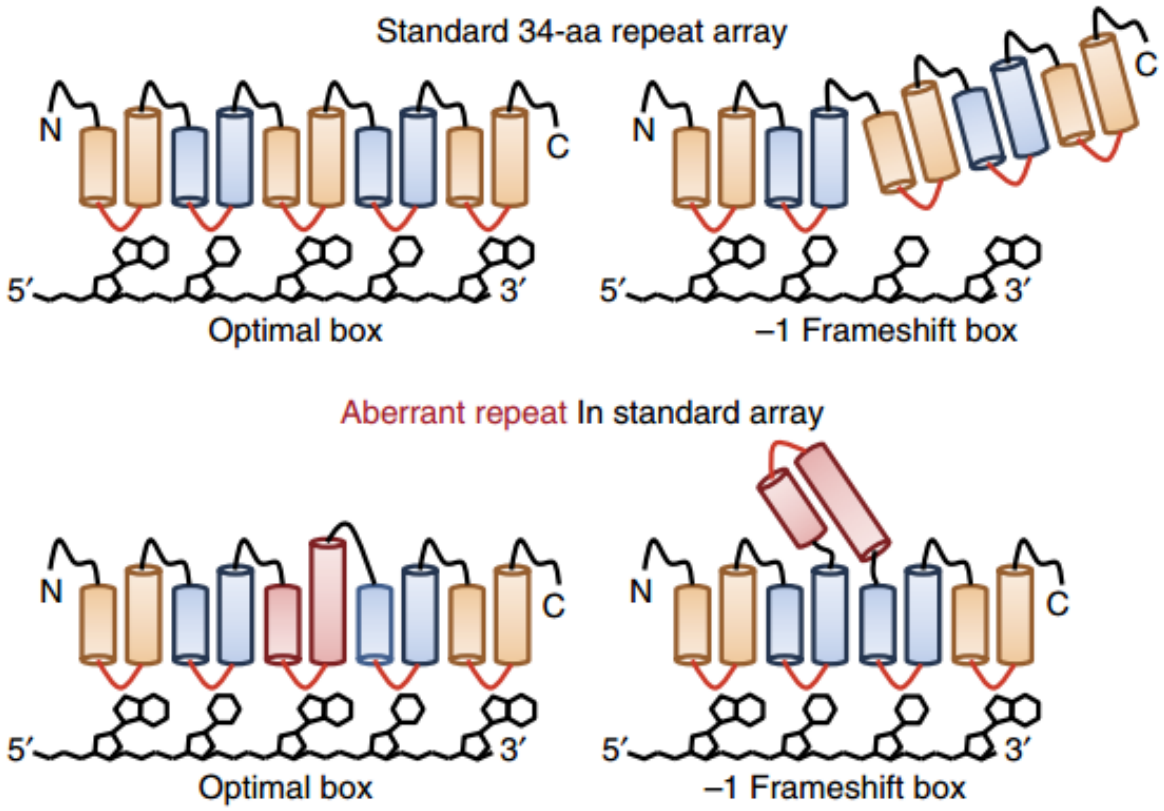
- 250 TAL effectors plus 8 pseudogenes with recognizable TAL effector repeat regions
- 241 complete TAL effectors represent 99 unique binding specificities (different in RVD or repeat length)

Conservation of unusual TAL effectors suggest functionality

- One TAL effector per strain has a truncated C-terminus with 95% or greater AA sequence identity to Tal2h C-terminus
  - 1 pseudogene, 2 of only 6 RVDs, the rest with repeat of 28bp

Conservation of unusual TAL effectors suggest functionality

- 6 with atypical repeat (Richter et al.) and 4 36bp repeats



Richter et al. (2014) *Nat Commun*

TAL effector conservation and known TAL effector targets up-regulated by the Xoc strains

Strain	From	Tal1a	Tal1b	Tal1c	Os07g06970	Tal2a	Os02g43760	Tal2b	Tal2c	Os03g03034	Tal2d	Os04g49194	Tal2e	Tal2f	Tal2g	Os01g52130	Os06g46500	Tal2h	Tal3a
BLS256	Philippines	0	0	0	11.61	0	7.18	0	0	54.23	0	474.73	0	0	0	219.92	100.32	-(AD) 0	0
BLS279	Philippines	X	0	0	12.24	2	1.00	1	0	60.92	4	28.66	X	X	0	197.56	179.52	-(AD) 1	0
CFBP2286	Malaysia	X	0	0	8.26	0	3.58	1	0	24.02	0	211.50	0	1	0	94.45	77.16	-(AD) 0	0
B8-12	China	X	0	0	9.46	0	5.34	1	1	110.79	4	56.25	2	X	0	223.22	151.25	X	0
L8	China	X	0	0	23.07	0	9.92	1	1	197.12	4	60.52	2	X	0	215.34	214.25	-(AD) 1	0
RS105	China	X	0	0	9.45	0	4.60	1	X	1.00	X	1.00	X	X	0	157.95	136.29	X	0
BXOR1	India	X	0	X	1.00	X	1.00	2	3	48.32	3	179.16	X	X	0	103.06	97.75	X	0
CFBP7331	W. Africa	7	X	X	1.00	X	1.00	6	4	85.65	3	120.57	X	X	1	177.22	208.31	X	1
CFBP7341	W. Africa	7	X	X	1.00	X	1.00	6	4	53.64	3	139.41	X	X	1	163.82	175.28	X	1
CFBP7342	W. Africa	X	1	X	1.00	X	1.00	6	X	15.45	3	292.41	X	X	1	204.80	180.73	X	1

Strain	From	Tal3b	Os02g34970	Os05g27590	Os07g36430	Tal3c	Os02g47660	Os03g07540	Tal4a	Os03g37840	Tal4b	Os09g32100	Tal4c	Os06g37080	Tal5a	Os02g15290	Tal5b	Tal6	Os01g31220	Os09g29100	Os12g42970
BLS256	Philippines	0	243.56	217.74	86.24	0	20.81	705.86	0	43.85	0	43.85	0	109.91	0	148.30	0	0	12.86	39.21	41.08
BLS279	Philippines	0	219.18	237.55	162.31	0	20.66	658.77	0	40.17	0	40.17	0	113.09	0	146.22	0	0	11.96	31.10	25.61
CFBP2286	Malaysia	0	77.78	115.06	64.53	0	11.36	326.54	X	1.00	1	1.00	1	45.34	0	103.51	0	1	1.00	25.76	13.19
B8-12	China	0	265.41	254.94	168.33	0	23.95	797.24	0	42.68	0	42.68	0	118.77	0	132.80	0	0	8.37	16.00	32.86
L8	China	0	172.66	233.61	134.39	0	27.70	625.95	0	29.78	1	29.78	0	117.09	0	172.70	0	0	22.41	86.74	20.36
RS105	China	0	161.28	223.77	135.07	0	17.76	585.38	0	35.48	0	35.48	0	85.01	0	117.09	0	0	6.29	13.16	16.52
BXOR1	India	0	92.27	74.34	81.79	1	12.14	252.90	1	24.55	3	24.55	5	42.92	2	173.96	0	1	2.04	15.61	8.94
CFBP7331	W. Africa	0	8.75	81.47	11.40	X	1.00	7.04	1	36.01	X	36.01	X	26.30	16	1.00	X	6	1.00	5.14	26.58
CFBP7341	W. Africa	0	11.26	74.93	13.85	X	1.00	5.35	1	37.92	X	37.92	4	38.67	X	1.00	X	6	1.00	1.00	24.29
CFBP7342	W. Africa	1	163.06	155.67	207.48	X	17.85	56.69	2	38.25	X	38.25	5	51.19	X	1.00	0	6	1.00	1.00	28.57

Strain	From	Tal7	Tal8	Tal9a	Tal9b	Os01g51040	Tal10	Tal11a	Tal11b	Tal12	non-BLS256 TAL#1	non-BLS256 TAL#2	non-BLS256 TAL#3	non-BLS256 TAL#4	non-BLS256 TAL#5	non-BLS256 TAL#6	non-BLS256 TAL#7	non-BLS256 TAL#8	non-BLS256 TAL#9	non-BLS256 TAL#10	non-BLS256 TAL#11	unique TAL effectors
BLS256	Philippines	0	0	0	0	30.46	0	0	0	0	X	X	X	X	X	X	X	X	X	X	X	0
BLS279	Philippines	0	0	0	3	14.66	0	0	0	2	X	X	X	3	X	X	X	X	X	X	X	0
CFBP2286	Malaysia	0	0	0	0	23.94	0	0	0	0	0	X	X	2	X	X	X	X	X	X	X	0
B8-12	China	0	0	0	3	15.47	0	0	0	0	X	X	0	0	X	-(AD) 0	X	X	X	X	X	0
L8	China	0	0	0	3	36.39	0	0	0	0	X	X	0	0	X	X	X	X	X	X	X	1
RS105	China	Ψ 0	0	0	3	16.20	0	0	0	0	X	X	0	X	X	-(AD) 0	X	X	X	X	X	0
BXOR1	India	0	X	1	X	1.00	X	2	0	1	0	X	0	X	X	X	2	X	0	0	X	3
CFBP7331	W. Africa	2	X	1	X	0.58	X	2	2	X	X	0	X	X	X	X	X	-(AD) 0	0	0	Ψ 0	3
CFBP7341	W. Africa	2	X	1	X	0.60	X	2	2	2	X	0	X	X	0	X	X	Ψ 0	0	0	Ψ 0	1
CFBP7342	W. Africa	3	X	1	X	1.00	X	X	2	2	X	X	1	X	4	X	0	X	0	X	X	6

TAL Effector Conservation

- Conserved TAL effectors
  - Could have important roles in virulence (ie may target S genes)
  - Could be used in an R gene trap for broad resistance
- Only 5 TAL effector groups have an ortholog in every strain which is neither pseudogenized nor diverged by more than 5 BSRs from the others
- These include the groups with BLS256 Tal2g and Tal11b, as well as Tal3a, Tal3b, and Tal9a

Candidate target conservation

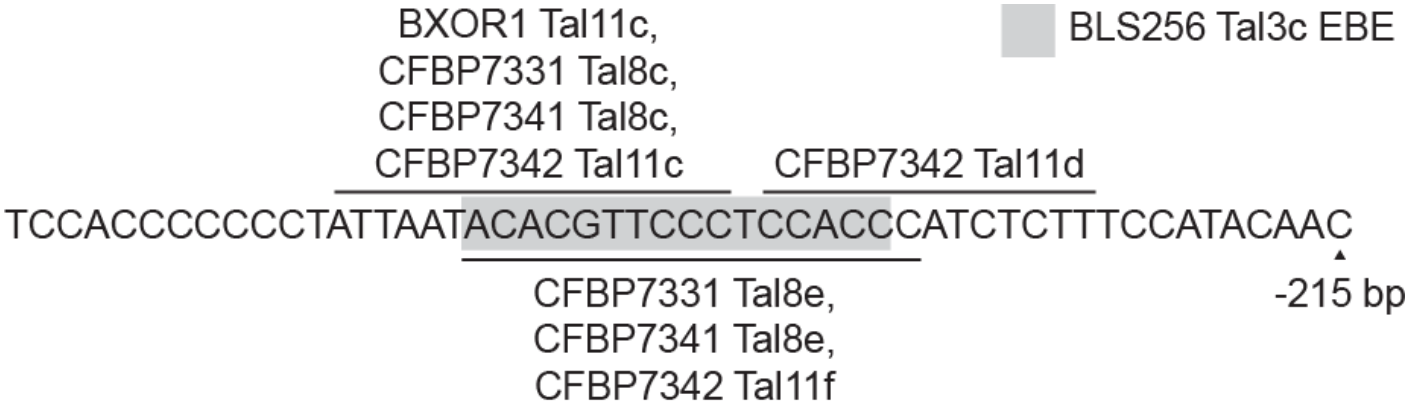
- Tal3c targets are up-regulated even in strains where Tal3c is not conserved

Strain	From	Tal3c	Os02g47660	Os03g07540
BLS256	Philippines	0	20.81	705.86
BLS279	Philippines	0	20.66	658.77
CFBP2286	Malaysia	0	11.36	326.54
RS105	China	0	17.76	585.38
B8-12	China	0	23.95	797.24
L8	China	0	27.70	625.95
BXOR1	India	1	12.14	252.90
CFBP7331	W. Africa	-1	1.00	7.04
CFBP7341	W. Africa	-1	1.00	5.35
CFBP7342	W. Africa	-1	17.85	56.69

- May represent convergent targeting of an S gene

Candidate target conservation

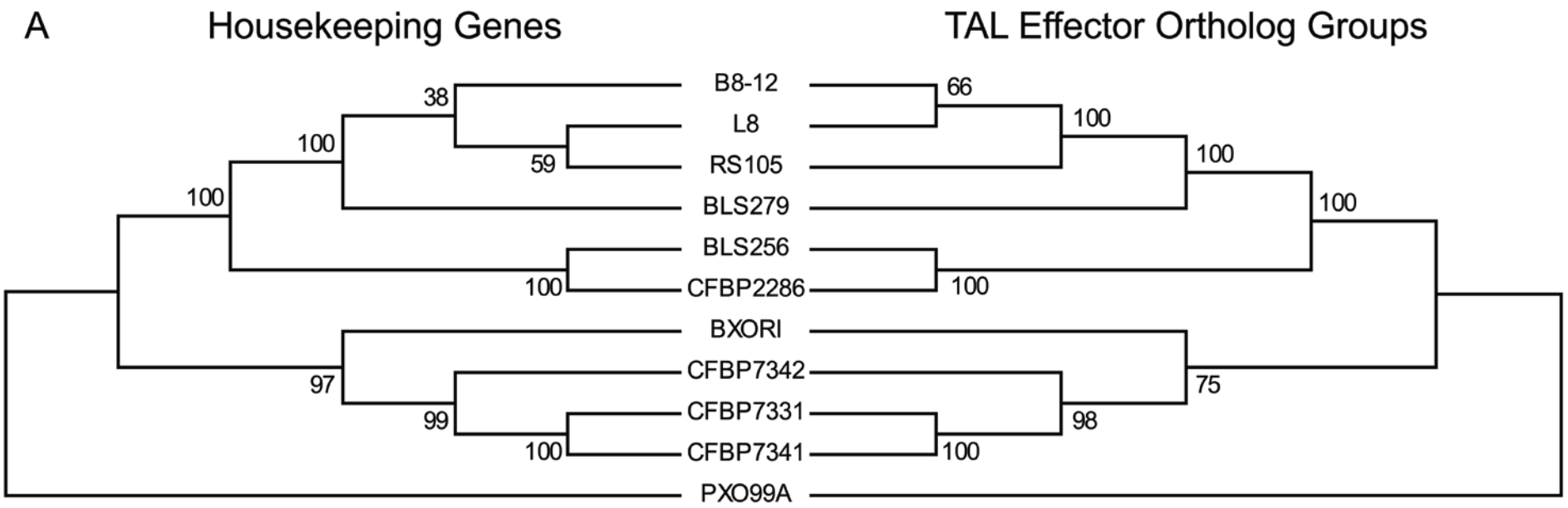
- There is a TAL effector found only in CFBP7342 with a passing EBE in Os02g47660
- There are also two TAL effectors in all of the African strains with passing EBEs in Os02g47660
- Both of these predicted EBEs overlap the Tal3c EBE



TAL effector distribution

- Xoc TAL effectors are far more conserved than Xoo TAL effectors, particularly within Asian or African strains
- Within the three sequenced Xoo genomes, at most 21% of their TAL effectors are conserved at the BSR level between any two strains
- Within the Asian Xoc strains, a minimum of 25% of TAL effectors and an average of 51% are conserved across every pair
- Within the African Xoc strains, a minimum of 32% of TAL effectors an average of 57% are conserved across every pair

TAL effector distribution



The tree created using the TAL effector orthologs is only a marginally significantly better fit for the TAL effector ortholog alignment than the tree created using the housekeeping genes (Kishino–Hasegawa test p-value = .054)
=> No horizontal gene transfer

Conclusions

- Xoc TAL effectors are highly conserved but TAL effector content distinguishes Asian and African strains of Xoc
- The five TAL effectors shared by all strains include the two BLS256 TAL effectors with a known role in virulence
- There is some evidence of convergent evolution leading to the targeting of one gene by different TAL effectors
- Xoc TAL effectors are largely vertically transmitted
- TAL effector-based resistance strategies appear promising

Future work

- Features that might influence binding:
  - Location in genome
  - Context effects
  - Epigenetic status
  - Chromatin status
- Can we use what we learn to modify EBEs without regard for genomic context?