A new method of generating diagnostic primers from draft bacterial genome sequences

Leighton Pritchard[1], Sonia Humphris[1], Gerry S Saddler[2],
Neil M. Parkinson[3], V. Bertrand[3], John G. Elphinstone[3] and **Ian K. Toth**[1]
[1]The James Hutton Institute, Invergowrie, Dundee, DD2 5DA, UK
[2]Science and Advice for Scottish Agriculture (SASA), Roddinglaw Road, Edinburgh, EH12 9FJ, UK
[3]Food and Environment Research Agency, Sand Hutton, York, YO41 1LZ, UK
Email: ian.toth@hutton.ac.uk

The James Hutton Institute

fera The Food and Environment Research Agency

The Scottish Government

Introduction

Dickeya spp. are plant pathogenic bacteria that cause major economic losses to ornamental and crop plants worldwide. Six species of *Dickeya* are currently defined. Of these, *D. solani* has recently emerged as a significant threat to potato production in Europe. In some countries *D. solani* has overtaken the more established bacterial pathogens *Pectobacterium atrosepticum* and *D. dianthicola* in terms of loss to disease (Toth *et al*., 2011 Plant Pathol. **60**, 385-9).

Reliable identification of *Dickeya* spp. is critical to control and monitoring of these pathogens.

We developed a computational pipeline that rapidly predicts discriminatory species-specific PCR primer sets, on the basis of draft genome sequences. We applied this to 16 draft and four complete *Dickeya* genome sequences, to generate over 700 primer sets predicted to discriminate between *Dickeya* at the species level, and validated a number of these sets in the laboratory (Pritchard *et al*. 2013. *Plant Pathol*. **62**, 587-96 – Paper 1).

Methods

Genome sequencing

16 *Dickeya* genomes were sequenced in this study (Roche/454) with others obtained from DNA databases.

Primer Prediction and Validation

The primer design pipeline is described fully in Pritchard *et al*. (2012. *PLoS ONE* **7**, e34498 – Paper 2) (Figure 1). Briefly, a training set of genome sequences was divided into *positive* (partitioned into classes I-V) and *negative* groups (Figure 1). 1000 primer sets were designed to each p*ositive* sequence (black markers), and tested for cross-hybridisation *in silico*. Primer sets amplifying all members of a single class (coloured markers) but not amplifying negative examples were retained as potentially diagnostic of that class. Predicted discriminatory primers were validated by qPCR against isolates not part of the training set.

Software

Primer design pipeline software may be downloaded at https://github.com/widdowquinn/find_differential_primers

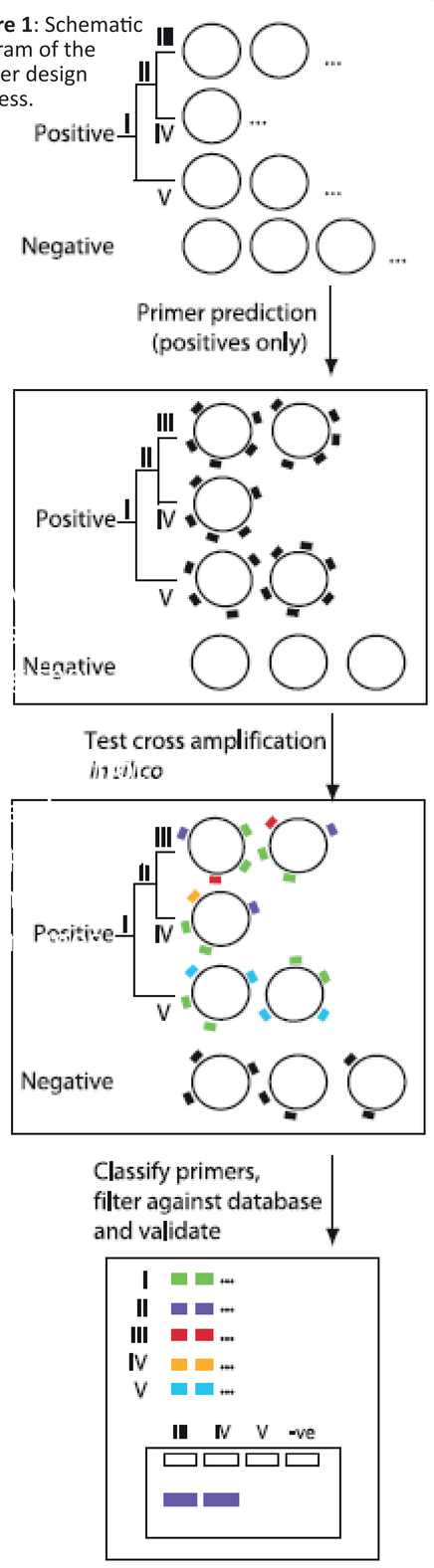

Figure 1: Schematic diagram of the primer design process.

Results

Table 1 shows counts of primers predicted to discriminate at the *Dickeya* isolate, species and genus levels. Fifteen predicted primer sets were selected randomly for validation: five each specific to *D. dianthicola*, *D. solani* and common (but not specific) to all *Dickeya* spp.

Each set generated a single PCR product from DNA in a panel of 13 representative *Dickeya* strains. Fourteen primer sets amplified DNA from the target species in conventional PCR. Predicted specificity of primers was confirmed for two *D. dianthicola* (DIA-A and DIA-C) and two *D. solani* (SOL-C and SOL-D) sets.

These sets were validated against a panel of 70 *Dickeya* strains. Table 2 shows specificity results for these primer sets. DIA-A and DIA-C amplified only isolates of *D. dianthicola* in the panel, as predicted. SOL-C and SOL-D amplified all panel isolates of *D. solani*, with a single false positive isolate, identified as *D. dadantii* by recA phylogenetic reconstruction.

		Amplification only			With hybridization oligo		
Identifier	Classification	Strain-specific	Species-specific	'Universal'	Strain-specific	Species-specific	'Universal'
Dch3533	*chrysanthemi*	0 (0)	136 (76)	9 (0)	0 (0)	139 (10)	9 (0)
Dch402	*chrysanthemi*	65 (52)	120 (69)	12 (0)	57 (10)	118 (14)	10 (0)
Dze1591	*chrysanthemi*	0 (0)	116 (69)	8 (0)	0 (0)	117 (9)	8 (0)
DzeP7246	*zeae*	24 (16)	128 (74)	5 (0)	21 (1)	129 (13)	5 (0)
Dze3531	*zeae*	17 (10)	111 (63)	5 (0)	16 (0)	117 (10)	6 (0)
Dze2538	*zeae*	13 (8)	114 (61)	9 (0)	12 (0)	113 (8)	9 (0)
DzeMK19	*zeae*	21 (17)	113 (63)	5 (0)	22 (3)	112 (9)	6 (0)
Dda586	*zeae*	36 (22)	127 (66)	5 (0)	30 (4)	131 (7)	5 (0)
Dda703	*paradisiaca*	0 (0)	342 (183)	10 (0)	0 (0)	347 (39)	11 (0)
Dpa2511	*paradisiaca*	4 (3)	342 (186)	10 (0)	4 (0)	346 (34)	12 (0)
Dda3937	*dadantii*	41 (24)	39 (26)	12 (0)	39 (4)	38 (7)	11 (0)
Ddf2976	*dadantii*	51 (26)	26 (18)	16 (0)	46 (5)	26 (5)	15 (0)
Ddi453	*dianthicola*	10 (9)	75 (59)	7 (0)	10 (2)	75 (9)	8 (0)
DdiIPO980	*dianthicola*	24 (16)	87 (74)	6 (0)	23 (1)	82 (9)	6 (0)
Ddi3534	*dianthicola*	23 (10)	79 (67)	8 (0)	22 (4)	76 (12)	9 (0)
DsoIPO2222	*solani*	0 (0)	57 (36)	10 (0)	0 (0)	56 (9)	8 (0)
DsoMK16	*solani*	0 (0)	55 (41)	13 (0)	0 (0)	52 (11)	11 (0)
DunkMK10	*solani*	0 (0)	57 (40)	15 (0)	0 (0)	54 (6)	13 (0)
DdiMK7	Unknown1	117 (83)	117 (83)	12 (1)	112 (19)	112 (19)	12 (0)
DunkP7247	Unknown2	312 (162)	312 (162)	12 (1)	310 (36)	310 (36)	12 (0)

Table 1: Counts of predicted diagnostic primer sets for *Dickeya* spp. Values in parentheses are counts of predicted primer sets when screened with BLASTN against all sequenced non-*Dickeya* Enterobacteriaceae, to eliminate potentially cross-amplifying primers.

		Number of isolates detected, by assay								
Test species	Isolates tested	DIC	DIA-A	DIA-C	SOL-C	SOL-D	PEC	ECH	ECA	ADE
D. dianthicola	7	**7**	**7**	**7**	0	0	**7**	**7**	0	**7**
'*D. solani*' (DUC-1)	16	**16**	0	0	**16**	**16**	**16**	**16**	0	**16**
DUC-2	5	**5**	0	0	0	0	**5**	**5**	0	**5**
DUC-3	1	**1**	0	0	0	0	**1**	**1**	0	**1**
D. dadantii	11	**11**	0	0	**1**	**1**	**11**	**11**	0	**11**
D. dieffenbachiae	6	**6**	0	0	0	0	**6**	**6**	0	**6**
D. chrysanthemi bv. *chrysanthemi*	7	**7**	0	0	0	0	**7**	**7**	0	**7**
D. chrysanthemi bv. *parthenii*	3	**3**	0	0	0	0	**3**	**3**	0	**3**
D. paradisiaca	1	**1**	0	0	0	0	**1**	**1**	0	**1**
D. zeae	11	**11**	0	0	0	0	**11**	**11**	0	**11**
New *Dickeya* species level clade (I)	1	**1**	0	0	0	0	**1**	**1**	0	**1**
New *Dickeya* species level clade (II)	1	**1**	0	0	0	0	**1**	**1**	0	**1**
Pectobacterium atrosepticum	1	**1**	0	0	0	0	**1**	0	**1**	0
P. carotovorum subsp. *carotovorum*	1	**1**	0	0	0	0	**1**	0	0	0
P. betavasculorum	1	**1**	0	0	0	0	**1**	0	0	0
P. carotovorum subsp. *odoriferum*	1	**1**	0	0	0	0	**1**	0	0	0
P. wasabiei	1	**1**	0	0	0	0	**1**	0	0	0
Pantoea agglomerans	1	**1**	0	0	0	0	0	0	0	0
Brenneria quercina	1	**1**	0	0	0	0	0	0	0	0
Erwinia amylovora	1	**1**	0	0	0	0	0	0	0	0

Table 2: Specificity of predicted real-time qPCR primer sets, compared with existing qPCR and conventional PCR assays. Positive results (critical threshold Ct <23) are shown in bold. The specificities of existing primer sets PEC, ECH, ECA and ADE are also confirmed.

Acknowledgements:
We wish to thank members of the Centre for Genomics Research at the University of Liverpool for sequencing 16 of the 20 Dickeya strains used in the study. The study was funded by the Scottish Government's Rural and Environment Science and Analytical Services Division (RESAS) (CR/2007/02) and the Agriculture and Horticulture Development Board (AHDB) through the Potato Council (R437).

Software Paper 1 Paper 2 Hutton

Conclusions

Our results demonstrate that highly-specific diagnostic primers can be designed on the basis of draft bacterial genome sequences. This enables rapid production of diagnostic tools by timely sequencing of representative isolates of pathogenic bacteria.

These primers are currently being tested as part of a EUPHRESCO initiative to standardise *Dickeya* and *Pectobacterium* testing across Europe.