environmental microbiology

Environmental Microbiology (2021) **23**(4), 2035–2053 doi:10.1111/1462-2920.15339

Comparative genomics of *Fusarium oxysporum* f. sp. *melonis* strains reveals nine lineages and a new sequence type of AvrFom2

Fatemeh Sabahi,[1,2] Mara de Sain,[2] Zia Banihashemi[1] and Martijn Rep[2]*

[1]*Department of Plant Protection, College of Agriculture, Shiraz University, Shiraz, Iran.*

[2]*Molecular Plant Pathology, University of Amsterdam, Amsterdam, 1098 XH, The Netherlands.*

Summary

***Fusarium oxysporum* f. sp. *melonis* (Fom) is one of the most important pathogens of melon worldwide. In this study, we investigated the genomic diversity of Fom. One of the aims was to find clues for the origin(s) and dispersal of clonal lineages and races of Fom. We therefore included a large number of Fom strains from Iran, where melon has been cultivated for at least 5000 years. In 33 new genome sequences of Fom strains from different geographical regions of Iran and across the world, 40 new candidate effector genes were identified. Presence/absence of candidate effector genes and phylogenetic analyses resolved nine Fom lineages. The presence of a highly similar set of effector genes in some distant lineages is suggestive of horizontal chromosome transfer, a process known to occur in the *Fusarium oxysporum* species complex. Race 1.2, which breaks both *Fom1* and *Fom2* resistance genes, occurs in three of the nine lineages, two of which are predominant in Iran. We also identified a new sequence type of the *AVRFom2* avirulence gene in one lineage. Expression of this sequence type during melon infection and genetic complementation suggest that this sequence type is not recognized by the Fom2 resistance protein.**

Introduction

Musk melon, *Cucumis melo* L., is a typical fruit vegetable and is endemic in subtropical and warm temperate regions. Fusarium wilt of melon, a disease which is difficult to control, is caused by *Fusarium oxysporum* f. sp. *melonis* (Fom) Snyder and Hans (Armstrong and Armstrong, 1978). Fom is one of the most widespread and economically important *formae speciales* in the *Fusarium oxysporum* species complex (FOSC) (Kim *et al*., 1993). Fom includes nine known vegetative compatibility groups (VCGs): 0130–0138 (Jacobson and Gordon, 1988; Jacobson and Gordon, 1990; Katan *et al*., 1994) and four races - 0, 1, 2 and 1.2 (Risser *et al*., 1976). Two dominant resistance (R) genes in musk melon, *Fom1* and *Fom2*, have a key role in the control of Fusarium wilt. These R genes are widely used in commercial melon cultivars and were identified through map-based cloning (Joobeur *et al*., 2004; Brotman *et al*., 2013; Oumouloud *et al*., 2013). Generally, the product of an R gene is an immune receptor that recognizes the product of an *AVR* gene in a pathogen, which most commonly encodes a small *in planta* secreted protein, also called 'effector'. This recognition activates a signal transduction cascade which leads to an effective defence response (Hammond-Kosack and Jones, 1997; Martin *et al*., 2003).

Race 0 of Fom is recognized by both *Fom1* and *Fom2*, race 1 by *Fom2*, race 2 by *Fom1* and race 1.2 by neither resistance gene (Risser *et al*., 1976). Schmidt *et al*. (2016) found that Fom2 recognizes the product of the avirulence gene *AVRFom2*, which is present in race 0 and race 1, but absent in the race 2 and race 1.2 strains investigated in that study.

Genomic analysis of *Fusarium oxysporum* (Fo) revealed an association of an upstream miniature impala (mimp) transposable element with effector genes (Schmidt *et al*., 2013). In the genomes of 59 strains of different *formae speciales*, 104 candidate effectors have been identified by searching in the proximity of the Terminal Inverted Repeat (TIR) sequence of mimps, and the presence/absence pattern of these effectors was shown to be predictive of the host of a pathogenic strain (Van Dam *et al*., 2016). Initial genomic comparison of different *formae speciales* of the FOSC indicated a relatively high genetic variability between Fom strains (van Dam *et al*., 2018).

While a wide global distribution has been reported for Fom, including Europe, North America, Israel, Japan and South Africa, a more limited distribution pattern has been

observed for Fom races (Jacobson and Gordon, 1991; Katan *et al*., 1994; Namiki *et al*., 1998; Schreuder *et al*., 2000). To determine the possible origin(s) of Fom and find patterns of its dispersal, Iran is of special interest. Melon has been cultivated in Iran for more than 5000 years (Zohary and Hopf, 2000; Mirtalebi *et al*., 2013), and there is no report of melon seed import into Iran. In contrast, the available evidence supports a history of melon seed export to European countries from Iran (Szabo *et al*., 2008). Until now, only Fom races 1 and 1.2, associated with VCG0134, have been reported in Iran (Banihashemi, 1968, 1982, 1989; Sarpeleh and Banihashemi, 2000; Mirtalebi *et al*., 2013). Initially, race 1.2 was reported in the provinces Fars, Isfahan and Kermanshah and race 1 was reported from Khorasan and Semnan provinces (Banihashemi, 1968, 1982, 1989; Sarpeleh and Banihashemi, 2000). In a recent study, only race 1.2 was reported from Khorasan, Fars, Isfahan, Markazi and Yazd provinces (Mirtalebi *et al*., 2013).

To obtain insight into the possible origin(s) and dispersal of Fom and the genetic relationship between the different Fom lineages, we characterized a collection of Fom strains representing the geographic and genetic diversity of Fom, including a large number of strains from Iran. We analysed the genome sequences of 43 Fom strains for the presence of candidate effector genes and identified nine Fom lineages based on core gene phylogeny as well as presence/absence patterns of candidate effector genes. Three of the nine lineages harbour race 1.2. Two of these three lineages are the only lineages found in Iran. We also identified a new sequence type of *AVRFom2* (*AVRFom2* ST2) and we show evidence that this sequence type is not recognized by the Fom2 resistance protein.

Results

Establishment of a Fom strain collection

To be able to investigate the number, the possible origin(s) and dispersal of Fom lineages, a variety of strains were selected from different geographical locations including Iran (Fig. 1), the United States, France, Spain, Israel, Mexico, New Zealand, the Netherlands, China, Japan and Belgium (see Table 1 for detailed information). For most of these strains, race determination and vegetative compatibility grouping have been performed in previous studies. Based on marker group (see below), reported clonal lineages (VCGs), geographical origin and race of the strains, we selected 33 Fom strains for whole-genome Illumina paired-end sequencing, including three strains of race 0, five strains of race 1, eight strains of race 2, 13 strains of race 1.2 and four strains of which we could not determine the race (Fom-Pathtah, Fom-Taip2a, Fom-Nasr1 and Fom020). We also sequenced the genome of five non-pathogenic strains isolated from melon and wild safflower plants in Khorasan and Fars provinces of Iran. Genome sequences of 10 additional Fom strains were already available (Schmidt *et al*., 2016), bringing the total number of Fom genome sequences analysed in this study to 43 (Table 1).

To determine or confirm the pathogenicity and race of the 38 strains of which we sequenced the genomes, we performed pathogenicity assays on differential melon cultivars. Fom strains which were pathogenic on Cha-T and Cha-*Fom1* (but not on Cha-*Fom2*) melon cultivars were classified as race 1, those that were able to cause disease on Cha-T and Cha-*Fom2* (but not on Cha-*Fom1*) cultivars were classified as race 2, strains that were only pathogenic on Cha-T were classified as race 0, those that were pathogenic on all three cultivars were classified as race 1.2 and those that failed to cause disease on any cultivar were classified as non-pathogenic (towards melon) (Fig. S1). Based on our pathogenicity-assay results, Fom025, Fom027, Fom017 and Fom-Bushehr2s were classified as race 0, 1, 2 and 1.2 respectively. Five strains were assigned a different race (or designated as non-pathogenic) compared with the published designation: Fom-Yazd2, Fom-KT2a, Fom-660A-1 and Fom-660A-17 strains belong to Fom race 1 and F-nonpath-Barmshour is non-pathogenic towards melon, while in a previous study Fom-Yazd2, Fom-KT2a and F-nonpath-Barmshour were classified as Fom race 1.2, and Fom-660A-1 and Fom-660A-17 were considered to be race 0 (Mirtalebi *et al*., 2013). Also, strains Fom-Pathtah and Fom-Taip2a were not able to cause disease in melon in our assays. For Fom-Pathtah this may be due to low spore concentration at inoculation (below 1×10^7 spores ml^{-1}), while Fom-Taip2a seems to have lost pathogenicity during cultivation. While earlier race 1.2 was the only race found among Iranian Fom strains (Mirtalebi *et al*., 2013), we here confirm the presence of both race 1 and 1.2 in Iran. We also determined the race of 12 strains not classified before: Fom025 strain is race 0, Fom019 and Fom027 are race 1, Fom017, Fom021, Fom024 and Fom026 are race 2 and Fom-Bushehr2s, Fom-Kavar-22, Fom-Mah9a, Fom007 and Fom023 are race 1.2. Table 1; Fig. S1A (plant fresh weight), and Fig. S1B (disease index score) give more detailed information.

Three main marker groups in Fom

A previous study, using nine candidate effector genes as markers, noted three marker groups (A, B and C) among Fom strains (Van Dam *et al*., 2018). These nine genes were also used in this study to classify our larger strain

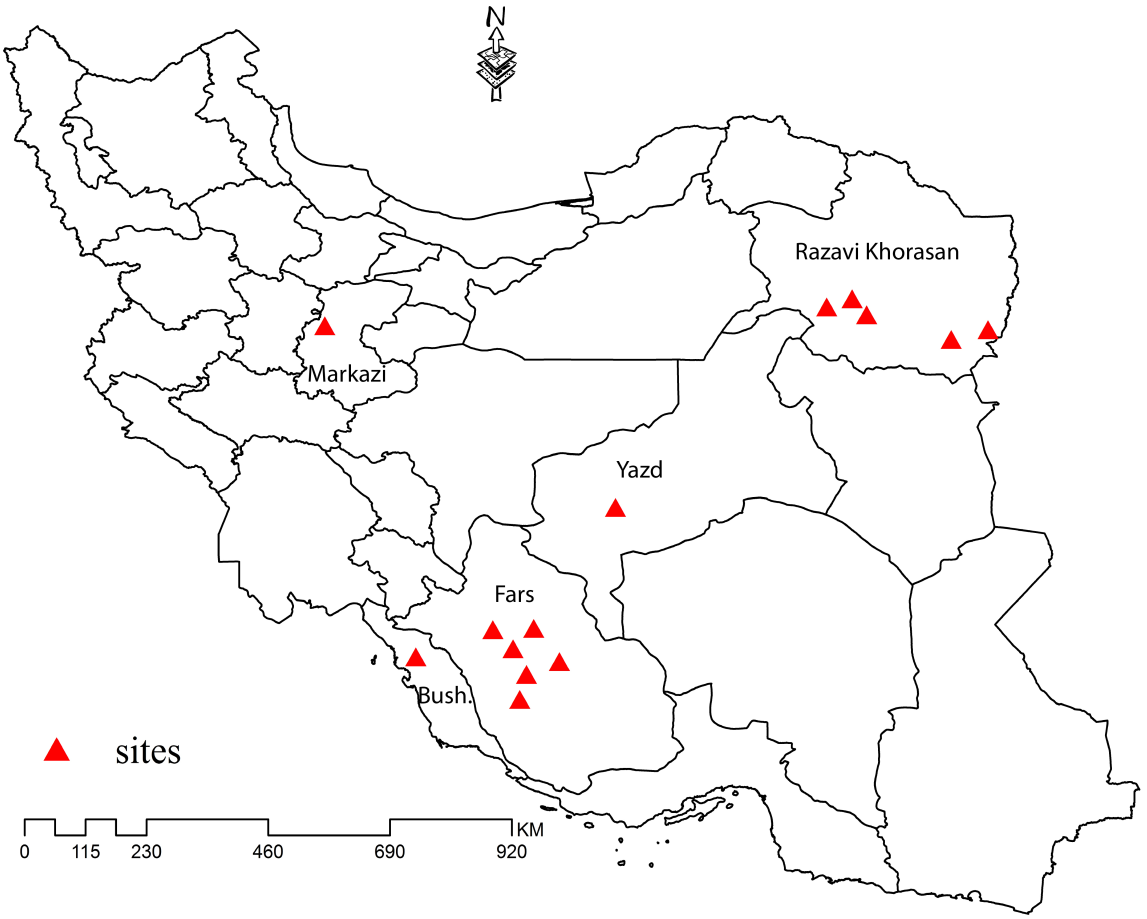

Fig 1. Map of Iran showing the location (triangles) where *Fusarium oxysporum* f. sp. *melonis* strains were isolated. [Color figure can be viewed at wileyonlinelibrary.com]

set to have extra information for selection of a representative set for Illumina sequencing. The same three main marker patterns were found. However, groups B and C were each divided into two subgroups, B1, B2, C1 and C2 (Table 2). The strains originating from Iran resided only in groups A and B1. The strains in subgroup B1 belong to races 1 and 1.2 while subgroup B2 only contains race 1.2 strains, subgroups C1 and C2 only race 2 strains. VCG0133 is associated with subgroup B2, VCG0132 with subgroup C1 and VCG0130 and 0131 with subgroup C2. Group A is more complex and includes strains of races 0, 1, 2 and 1.2 and VCG0134, 0135 and 0136.

Discovery of a second sequence type of AVRFom2

The identification of *AVRFom2* in race 0 and 1 has been reported by Schmidt *et al*. (2016). The same report showed the *AVRFom2* gene was absent from race 2 and 1.2 strains investigated, explaining why these strains escape recognition by Fom2. To investigate the presence and sequence of this avirulence gene in our strain set, *AVRFom2* primers were used to amplify and sequence this gene. The PCR results indicated that the *AVRFom2* gene is present in all of the race 0 and race 1 strains, as expected. Surprisingly, however, *AVRFom2* was also present in most race 2 strains of marker group C2 (except Fom021). The sequence of *AVRFom2* in all race 0 and race 1 strains, including the three race 1 strains from Iran (Fom-I-17, Fom-KT2a and Fom-Yazd2 all VCG0134), was identical to the one described previously (Schmidt *et al*., 2016). Hereafter, we will call this sequence *AVRFom2* sequence type 1 (ST1). However, the sequence of *AVRFom2* in the race 2 strains was different and was named *AVRFom2* ST2 (Table 2, Fig. S4A). We have thus identified a new sequence type of *AVRFom2*, which is only present in race 2 strains.

In planta *expression of* AVRFom2 *ST2 and genetic complementation suggest that AvrFom2 is not recognized by the Fom2 resistance protein*

Since the melon resistance gene *Fom2* is not effective against Fom race 2 strains, *AVRFom2* ST2 in race

Table 1. *Fusarium oxysporum* f. sp. *melonis* strains used in this study.

Isolate	Original designation	Previously determined race	Race designation in this study	VCG	Whole-genome sequenced	Origin of isolate	Source or Reference
F-nonpath-Barmshour	Barmshour	1.2	np		+	Iran, Barmshour-Fars	Mirtalebi *et al*. (2013)
Fo-nonpath-2 Ma4-5	2 Ma4-5	np	np		+	Iran, Khorasan	Mirtalebi *et al*. (2013)
Fom-Pathtah[c]	Pathtah	1.2	Unclear[a]	0134	+	Iran, Maharloo-Fars	Mirtalebi *et al*. (2013)
Fo-nonpath-TO1	TO1	np	np[b]		+	Iran, Khorasan	Mirtalebi *et al*. (2013)
Fo-nonpath2	nonpath2	np	np[b]		+	Iran, Fars	Mirtalebi *et al*. (2013)
Fo-nonpath5[c]	nonpath5	np	np[b]		+	Iran, Fars	Mirtalebi *et al*. (2013)
Fom-18L	18L	2	2	0130	+	USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-660A-1	660A/1	0	1[b]	0134		France	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-660A-17	660A/17	0	1[b]	0134	+	France	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Busherh-2s	Busherh-2 s		1.2		+	Iran, Bushehr	In this study
Fom-FomGol[c]	FomGol	1.2	1.2[b]	0134	+	Iran, Maharloo-Fars	Mirtalebi *et al*. (2013)
Fom-gh-30	gh-30	1.2	Untested	0134		Iran, Gharechay-Markazi	Mirtalebi *et al*. (2013)
Fom-gh2-5a	gh2-5a	1.2	Untested	0134		Iran, Gharechai-Markazi	Mirtalebi *et al*. (2013)
Fom-I-17	I-17	1	1	0134	+	Iran, Khorasan	Banihashemi (1968)
Fom-I1-1	I1/1	0	0[b]	0135	+	Israel	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-I1-5	I1/5	0	Untested	0135		Israel	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-K419-2	K419/2	1	Untested	0136		Mexico	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-K419-5	K419/5	1	Untested	0136		Mexico	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Kavar-22	Kavar 22		1.2		+	Iran, Kavar-Fars	In this study
Fom-Khaf1	Khaf1	1.2	1.2[b]		+	Iran, Khaf-Khorasan	Mirtalebi *et al*. (2013)
Fom-KT2a	KT2a	1.2	1	0134	+	Iran, Kashmar-Khorasan	Mirtalebi *et al*. (2013)
Fom-Mah8a	Mah8a	1.2	Untested			Iran, Maharloo-Fars	In this study
Fom-Mah9a	Mah9a		1.2[b]		+	Iran, Maharloo-Fars	In this study
Fom-Nasr1	Nasr1	1.2	Unclear		+	Iran, Nasrabad-Khorasan	Mirtalebi *et al*. (2013)
Fom-NYFom3	NYFom3	2	2[b]	0131	+	USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-NYFom59	NYFom59	1	Untested	0134		New York, USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-NYFom62	NYFom62	1	1[b]	0134	+	USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-NYFom67	NYFom67	2	Untested	0131		New York, USA	Gordon, T.R., Risser, G. and Zitter, T.Z.

(Continues)

Table 1. Continued

Isolate	Original designation	Previously determined race	Race designation in this study	VCG	Whole-genome sequenced	Origin of isolate	Source or Reference
Fom-P13	P13	1.2	1.2	0134	+	Iran, Poshtpar-Fars	Mirtalebi *et al*. (2013)
Fom-P2-1	P2/1	2	Untested	0130		California, USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Pt3-1	Pt3/1	2	Untested	0131		USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-R12-13	R12/13	1.2	1.2	0133	+	France	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-R12-14	R12/14	1.2	Untested	0133		France	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Sample 37	Sample 37	2	Untested	0131		USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Sample 65	Sample 65	1	Untested	0134		USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Sample 82	Sample 82	2	Untested	0130		USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Sample 88	Sample 88	1	Untested	0134		USA	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Seif3a	Seif3a	1.2	1.2[b]	0134	+	Iran, Seifabad-Fars	Mirtalebi *et al*. (2013)
Fom-Sh88	Sh88	1.2	Untested	SI		France	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Sh90	Sh90	1.2	Untested	SI		France	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-T61-1	T61/1	2	2[b]	0132	+	Japan	Gordon, T.R., Risser, G. and Zitter, T.Z.
Fom-Tai3	Tai3	1.2	1.2[b]		+	Iran, Taibad-Khorasan	Mirtalebi *et al*. (2013)
Fom-Taip2a	Taip2a	1.2	Unclear		+	Iran, Taibad-Khorasan	Mirtalebi *et al*. (2013)
Fom-Yazd2	Yazd2	1.2	1	0134	+	Iran, Yazd	Mirtalebi *et al*. (2013)
Fom001	NRRL26406	1	1[b]	0136	+	Mexico	Broad Institute
Fom002		np	Untested			Unknown	
Fom003			Untested			Unknown	
Fom004	Fom0122	0	0	0134	+	Spain	Schmidt *et al*. (2016)
Fom005	Fom0123	1	1[b]	0134	+	Spain	Schmidt *et al*. (2016)
Fom006	Fom0124	2	2	0134	+	Spain	Schmidt *et al*. (2016)
Fom007	Fom 0125		1.2[b]		+	Spain	Torres
Fom009		2	Untested	0135	+	Israel	Schmidt *et al*. (2016)
Fom010		1	Untested		+	Israel	Schmidt *et al*. (2016)
Fom011		0	Untested		+	Israel	Schmidt *et al*. (2016)
Fom012	ML2	0	Untested	0134	+	Unknown	Schmidt *et al*. (2016)
Fom013		2	Untested	0134	+	Spain	Schmidt *et al*. (2016)
Fom014	D'oleon8	1.2	1.2		+	France	Michel Pitrat
Fom015	TST	1.2	1.2[b]		+	France	Michel Pitrat
Fom016	Fom26	1	Untested	0134	+	Unknown	Schmidt *et al*. (2016)
Fom017	NRRL 22518		2		+	USA	O'Donnell *et al*. (1998)
Fom018		1	Untested			France	O'Donnell *et al*. (1998)
Fom019			1			USA	O'Donnell *et al*. (1998)
Fom020	NRRL 22521		Unclear		+	Belgium	O'Donnell *et al*. (1998)
Fom021	NRRL 26172		2		+	China	Gordon, T.
Fom022			Untested			China	
Fom023	NRRL 26174		1.2		+	China	Gordon, T.
Fom024	NRRL 26745		2[b]		+	Japan	Japan Collection of Microorganisms, RIKEN BioResource Center
Fom025	NRRL 26746		0		+	New Zealand	

(Continues)

Table 1. Continued

Isolate	Original designation	Previously determined race	Race designation in this study	VCG	Whole-genome sequenced	Origin of isolate	Source or Reference
Fom026	NRRL 38516		2[b]		+	New Zealand	Japan Collection of Microorganisms, RIKEN BioResource Center P. Johnston, International Collection of Microorganisms From Plants, Plant Disease Division, DSIR, Auckland
Fom027			1			New Zealand	
Fom041		1	Untested			Unknown	
Fom042		1.2	Untested			Unknown	
Fom043		2	2		+	Unknown	Diederik Smilde, NAKT
Fom044		0	Untested			Unknown	
Fom045			Untested			Netherlands	
Fom046		1	Untested			Netherlands	
Fom047		0	0		+	Netherlands	Diederik Smilde, NAKT
Fom048			Untested			Spain	
Fom049			Untested			Spain	
Fom050			Untested			Unknown	
Fom051			Untested			Unknown	

Abbreviations: SI, vegetatively self-incompatible; NAKT, NAKtuinbouw, Netherlands Inspection Service for Horticulture, Roelofarendsveen, Netherlands.

[a]Fom-Pathtah did not produce enough spores to inoculate with 1×10^7 spores ml^{-1}; Fom-Taip2a seems to have lost pathogenicity during cultivation.

[b]Strain was only tested once.

[c]Fom-Pathtah isolated from Alhagi camellorum; Fom-FomGol and Fo-nonpath5 isolated from wild safflower Fom-T61-1 produced less than 1×10^7 spores ml^{-1}.

2 strains is either not expressed, the AvrFom2 ST2 protein is not recognized by the Fom2 immune receptor, or recognition takes place but resistance is suppressed, for instance by another effector of Fom. To exclude the first possibility (*AVRFom2* ST2 is not expressed in race 2 strains) *in planta* expression of *AVRFom2* ST2 was analysed during infection. RT-PCR and qPCR experiments were carried out using *AVRFom2* primers, and *EF-1α* primers as a reference. Six strains, including one strain with *AVRFom2* ST1 (Fom001) as a positive control, four strains with *AVRFom2* ST2 (Fom026, Fom017, Fom-NYFom3, Fom-18L) and one strain without *AVRFom2* (Fom024) were used for the experiment. The results showed that *AVRFom2* ST2 is expressed during melon colonization (Fig. 2; Table 3). The expression level of *AVRFom2* in Fom026, Fom-NYFom3, Fom-18L and Fom017 was similar, while expression in Fom001 was higher compared with the other strains (Table 3). As *AVRFom2* ST2 is expressed during infection in race 2 strains, although at lower levels than ST1 in Fom001, it seems unlikely that *Fom2* is not effective against these strains due to lack of expression.

Now it remained possible that recognition of *AVRFom2* ST2 by Fom2 is suppressed in the background of the race 2 strains that produce it, for example by another effector. If this is the case, we reasoned that *AVRFom2* ST1 should not confer avirulence to these race 2 strains, because recognition of *AVRFom2* ST1 would also be suppressed in the background of these strains. To investigate this, we transformed the race 2 strains Fom006 (no *AVRFom2*) and Fom017 (*AVRFom2* ST2) with *AVRFom2* (ST1) (Schmidt *et al*., 2016). Presence of *AVRFom2*, including promotor and terminator region, in putative transformants was confirmed by PCR (Fig. S2A–G). Five independent transformed strains derived from Fom006 and Fom017 were randomly selected for bioassays.

Susceptible (Cha-T) and resistant (Cha-*Fom2*) melon cultivars were inoculated with a race 1 strain (Fom-I-17), the two race 2 strains (Fom006, Fom017) and the 10 strains genetically transformed with *AVRFom2*. The results of disease severity assessment for the *Fom2* melon plants clearly showed that *AVRFom2* ST1 can confer avirulence to a race 2 strain containing *AVRFom2* ST2 (Fom017): fresh weight (Fig. 3A), disease index (Fig. 3B) and overall phenotype (Fig. 3C) of infected *Fom2* melon plants were similar to the race 1 control (Fom-I-17), while the untransformed race 2 strains (Fom006, Fom017) did show severe disease symptoms on Cha-*Fom2*. Transformation did not impair pathogenicity, as all strains were pathogenic on Cha-T plants (Fig. 3A–C). These results indicate that there is no factor

Table 2. Effector patterns of 69 Fom strains determined by PCR. [Color table can be viewed at wileyonlinelibrary.com]

Name	Origin	Race determined previously	Race	VCG	Fem	66 Focuc	21 focuc	1 Fom	20 Fom	18 Fom + Foniv	99 Foniv	100 Foniv	1 Fomom +Folog	99 Folag	AVRFom2	AVRFom2 ST	Marker group
Fom001	Mexico	1	1	0136	+			+	+	+					+	ST1	A
Fom002	Unknown	np			+			+	+	+							A
Fom004	Spain	0	0	0134	+			+	+	+					+	ST1	A
Fom005	Spain	1	1	0134	+			+	+	+					+	ST1	A
Fom006	Spain	2	2	0134	+			+	+	+							A
Fom009	Israel	2		0135	+			+	+	+							A
Fom011	Israel	0			+			+	+	+							A
Fom012	Unknown	0		0134	+			+	+	+							A
Fom013	Spain	2		0134	+			+	+	+							A
Fom015	France	1.2	1.2		+			+	+	+							A
Fom016	Unknown	1		0134	+			+	+	+							A
Fom018	France	1			+			+	+	+					+	ST1	A
Fom019	USA		1		+			+	+	+					+	ST1	A
Fom020	Belgium		Unclear		+			+	+	+					+	ST1	A
Fom025	New Zealand		0		+			+	+	+					+	ST1	A
Fom027	New Zealand		1		+			+	+	+					+	ST1	A
Fom041	Unknown	1			+			+	+	+					+	ST1	A
Fom042	Unknown	1.2			+			+	+	+							A
Fom043	Unknown	2	2		+			+	+	+							A
Fom044	Unknown	0			+			+	+	+					+	ST1	A
Fom045	Netherlands				+			+	+	+							A
Fom046	Netherlands	1			+			+	+	+					+	ST1	A
Fom047	Netherlands	0	0		+			+	+	+					+	ST1	A
Fom050	Unknown				+			+	+	+							A
Fom-660A-17	France	0	1	0134	+			+	+	+					+	ST1	A
Fom-Busherh-2 s	Iran		1.2		+			+	+	+							A
Fom-FomGol	Iran	1.2		0134	+			+	+	+							A
Fom-gh2-5a	Iran	1.2		0134	+			+	+	+							A
Fom-gh-30	Iran	1.2		0134	+			+	+	+							A
Fom-I1-1	Israel	0	0	0135	+			+	+	+					+	ST1	A
Fom-I-17	Iran	1	1	0134	+			+	+	+					+	ST1	A
Fom-K419-2	Mexico	1		0136	+			+	+	+					+	ST1	A
Fom-K419-5	Mexico	1		0136	+			+	+	+					+	ST1	A
Fom-Kavar-22	Iran		1.2		+			+	+	+							A
Fom-KT2a	Iran	1.2	1	0134	+			+	+	+					+	ST1	A
Fom-Mah9a	Iran		1.2		+			+	+	+							A
Fom-NYFom 62	USA	1	1	0134	+			+	+	+					+	ST1	A
Fom-NYFom59	USA	1		0134	+			+	+	+					+	ST1	A
Fom-P13	Iran	1.2	1.2	0134	+			+	+	+							A
Fom-Pathtah	Iran	1.2	Unclear	0134	+			+	+	+					+	ST1	A
Fom-Sample 65	USA	1		0134	+			+	+	+					+	ST1	A
Fom-Sample 88	USA	1		0134	+			+	+	+					+	ST1	A
Fom-Seif3a	Iran	1.2	1.2	0134	+			+	+	+							A
Fom-Yazd2	Iran	1.2	1	0134	+			+	+	+					+	ST1	A
Fom003	Unknown				+				+	+							A

(Continues)

Table 2. Continued

Name	Origin	Race determined previously	Race	VCG	Fem	66 Focuc	21 focuc	1 Fom	20 Fom	18 Fom + Foniv	99 Foniv	100 Foniv	1 Fomom +Folog	99 Folag	AVRFom2	AVRFom2 ST	Marker group
Fom007	Spain		1.2		+			+	+								A
Fom-660A-1	France	0	1	0134	+			+	+						+	ST1	A
Fom-I1-5	Israel	0		0135	+			+	+						+	ST1	A
Fom010	Israel	1			+			+	+	+	+	+		+	+	ST1	B1
Fom-Khaf1	Iran	1.2	1.2		+			+	+	+	+	+		+			B1
Fom-Mah8a	Iran				+			+	+	+	+	+		+			B1
Fom-Nasr1	Iran	1.2	Unclear		+			+	+	+	+	+		+			B1
Fom-Tai3	Iran	1.2	1.2		+			+	+	+	+	+		+			B1
Fom-Taip2a	Iran	1.2	Unclear		+			+	+	+	+	+		+			B1
Fom014	France	1.2	1.2		+				+	+	+	+		+			B2
Fom023	China		1.2		+				+	+	+	+		+			B2
Fom-R12-13	France	1.2	1.2	0133	+				+	+	+	+		+			B2
Fom022	China				+	+	+			+			+				C1
Fom024	Japan		2		+	+	+			+			+				C1
Fom-T61-1	Japan	2	2	0132	+	+	+			+			+				C1
Fom017	USA		2		+	+	+						+		+	ST2	C2
Fom021	China		2		+	+	+						+				C2
Fom026	New Zealand		2		+	+	+						+		+	ST2	C2
Fom-18 l	USA	2	2	0130	+	+	+						+		+	ST2	C2
Fom-NYFom3	USA	2	2	0131	+	+	+						+		+	ST2	C2
Fom-NYFom67	USA	2		0131	+	+	+						+		+	ST2	C2
Fom-P2-1	USA	2		0130	+	+	+						+		+	ST2	C2
Fom-Pt3-1	USA	2		0131	+	+	+						+		+	ST2	C2
Fom-Sample 37	USA	2		0131	+	+	+						+		+	ST2	C2
Fom-sample 82	USA	2		0130	+	+	+						+		+	ST2	C2
F-nonpath-Barmshour	Iran	1.2	np		+					+							np
Fom048	Spain				+												np
Fom049	Spain				+												np
Fom051	Unknown				+												np
Fom-R12-14	France	1.2		0133	+												np
Fom-Sh88	France	1.2		SI	+												np
Fom-Sh90	France	1.2		SI	+					+							np
Fo-nonpath2	Iran	np	np		+												np
Fo-nonpath-2 Ma4-5	Iran	np	np		+												np
Fo-nonpath5	Iran	np	np		+												np
Fo-nonpath-TO1	Iran	np	np		+												np

Tested in paper van Dam *et al*. (2018).

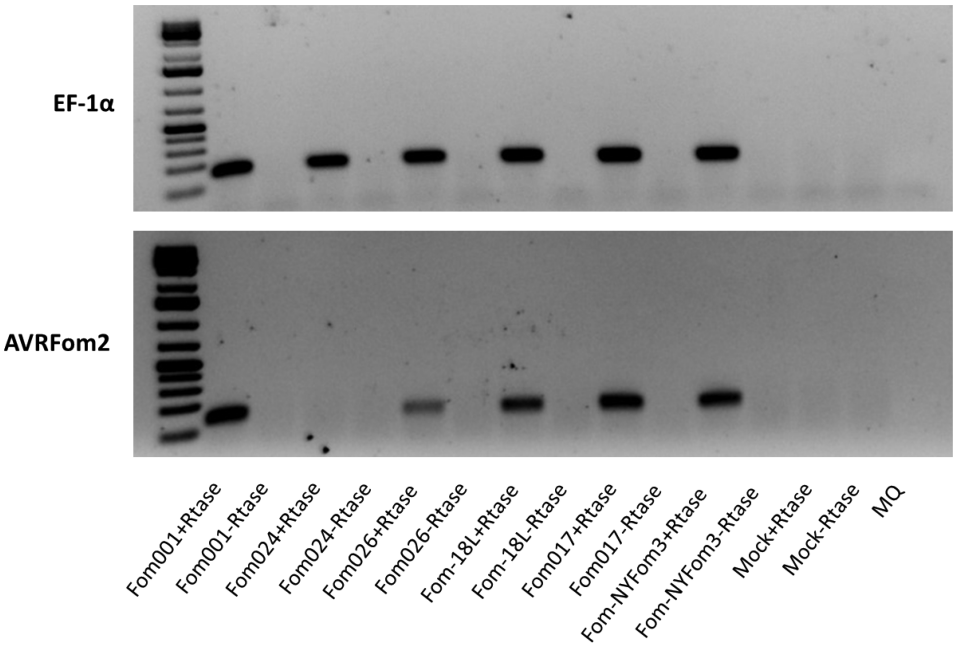

Fig 2. *AVRFom2* ST2 is expressed during colonization of melon. RT-PCR analysis of *AVRFom2* and *EF-1α* using RNA isolated from roots and hypocotyls of 10-day-old melon seedlings inoculated with race 1 strain Fom001 (*AVRFom2* ST1), and race 2 strains Fom024 (no *AVRFom2*), Fom026 (*AVRFom2* ST2), Fom-18L (*AVRFom2* ST2), Fom017 (*AVRFom2* ST2) and Fom-NYFom3 (*AVRFom2* ST2). Fom001-inoculated seedlings and the *EF-1α* gene were included as positive controls, while mock-inoculated seedlings and water (MQ) were used as negative controls. [Color figure can be viewed at wileyonlinelibrary.com]

Table 3. *AVRFom2* ST2 is expressed by *F. oxysporum* f. sp. *melonis* during melon infection.

Experiment 1			Mean CT ± stdev			Relative expression
Isolate	AVRFom2	Marker group	AVRFom2	EF-1α	TUB	AVRFom2/EF-1α
Fom001	ST1	A	30.764 ± 0.456	29.896 ± 0.072	31.847 ± 0.271	0.548
Fom026	ST2	C	32.894 ± 0.234	29.915 ± 0.053	30.136 ± 0.018	0.127
Fom-18L	ST2	C	31.240 ± 0.263	28.767 ± 0.079	30.385 ± 0.077	0.180
Fom-NYFom3	ST2	C	30.443 ± 0.282	28.512 ± 0.075	30.013 ± 0.054	0.262
Fom017	ST2	C	30.499 ± 0.247	27.544 ± 0.053	31.474 ± 0.071	0.129
Fom024	–	C	–	29.577 ± 0.097	30.227 ± 0.194	0.000
mock	–	–	39.187 ± 0.503	–	26.721 ± 0.108	0.000
mQ	–	–	–	–	–	0.000
Experiment 2			Mean CT ± stdev			Relative expression
Isolate	AVRFom2	Marker group	AVRFom2	EF-1α	TUB	AVRFom2/EF-1α
Fom001	ST1	A	31.177 ± 0.312	30 0.216 ± 0 0.022	30.361 ± 0.029	0.514
Fom026	ST2	C	32.966 ± 0.251	28 0.165 ± 0 0.054	31.037 ± 1.511	0.036
Fom-18L	ST2	C	31.269 ± 0.240	27 0.319 ± 0 0.076	30.279 ± 0.927	0.065
Fom-NYFom3	ST2	C	31.144 ± 0.184	27 0.245 ± 0 0.048	29.933 ± 1.331	0.067
Fom017	ST2	C	32.118 ± 0.294	26 0.364 ± 0 0.082	33.055 ± 1.417	0.019
Fom024	–	C	–	29 0.386 ± 0 0.122	27.159 ± 0.180	0.000
mock	–	–	38.702 ± 0.512	–	25.300 ± 0.108	0.000
mQ	–	–	–	–	–	0.000

AVRFom2 expression by Fom *in planta* was analysed by quantitative real-time PCR (qPCR). Four strains with *AVRFom2* ST2 were selected. One Fom strain with *AVRFom2* ST1 (Fom001) was used as a positive control. Fom024 (no *AVRFom2*) and water-inoculated plants were used as negative controls. RNA was extracted from roots and hypocotyls of 10-day-old melon seedlings (material of five plants was pooled per sample) infected by these strains and cDNA was synthesized. *EF-1α* was used as a reference gene and *AVRFom2* as a target gene. To be able to compare the amount of plant versus fungal material the musk melon *Tubulin* (TUB) gene was also targeted. This experiment was repeated twice with similar results.

in the background of *AVRFom2* ST2-containing race 2 strains that are capable of suppressing the Fom2 resistance response. As we have ruled out the other two possibilities, these experiments suggest that AvrFom ST2 is not recognized by Fom2.

Identification of new candidate effectors

The paired-end Illumina sequencing method was applied to sequence the genome of the 38 selected strains (Supporting information Table S1, trimmed and cleaned Illumina data). To identify new candidate effectors we first made *de novo* genome assemblies using CLC genomics workbench and then ran an effector prediction pipeline. The number of scaffolds resulting from *de novo* assembly ranged widely from 1675 (Fom-18L) to 4472 (Fom-R12-13). The smallest and largest *de novo* assembly were 51.0 (Fom024) and 56.5 (F-nonpath-Barmshour) megabase pairs (Mbp) respectively. An indication of the quality of the assembly was determined by calculating the N50, which ranged from 99 974 bp (Fom-R12-13) to 527 362 bp (Fo-nonpath-TO1).

To predict candidate effectors in the 38 selected strains, the known association of effector genes with mimps in Fo was exploited. First, the presence of mimp TIRs was determined in each genome separately. Subsequently, the method developed by Van Dam *et al*. (2016) was used for candidate effector identification. By applying this pipeline to our selected genomes, 1871 ORFs encoding predicted secreted proteins were identified within 2000 bp from a mimp TIR in all 38 genome sequences. To reduce redundancy, all ORFs were grouped into gene families using a self-BLAST survey, which led to a set of 124 candidate effectors of which 50 were already described by Van Dam *et al*. (2016). We then removed 34 ORFs that were either incomplete, coded for a mature protein below 35 amino acids or for which the signal peptide (SP) did not adhere to additional scrutiny (see Experimental procedures). Thus, our search resulted in 40 new candidate effectors in Fom genomes (Supporting information Table S2). To obtain information about the possible function of these 40 candidate effector proteins a blastP search was performed against the non-redundant nucleotide database maintained by NCBI. The result showed that 23 of the predicted protein sequences have high similarity to hypothetical or uncharacterized proteins (mainly from Fo isolates), two proteins showed high similarity to enzymes (an endopolygalacturonase and a metallo-hydrolase/oxidoreductase) and no significant hits were found for the remaining 15 proteins.

Fom strains are present in nine genetic lineages, which correspond to effector groups

In an earlier study, it was shown that strains infecting melon (f. sp. *melonis*), watermelon (f.sp. *niveum*), cucumber (f. sp. *cucumerinum*) or all three plant species (f. sp. *radicis-cucumerinum*) separate according to their host(s) in hierarchical clustering of presence/absence of candidate effector genes (van Dam *et al*., 2016). To assess whether this holds true for the larger and more diverse set of 43 Fom strains analysed here, hierarchical clustering was performed of presence/absence of 144 candidate effector genes in 87 Fo genomes, including the cucurbit-infecting ff. spp. mentioned above as well as, for comparison, f. sp. *lycopersici* (Fig. 4; Fig. S3). This resulted in a pattern of clustering of Fom strains (coloured yellow in Fig. 4 and Fig. S3) that corresponds to the marker groups based on the eight candidate effectors described above (A, B1, B2, C1, C2, see Table 2), except that this more detailed analysis allowed splitting group A into four subgroups (A1–4). In addition, the two strains in group C1 (Fom-T61-1 and Fom024) are separated from each other in the hierarchical clustering. The candidate effector patterns of these two strains are highly similar to C2 Fom strains and - remarkably - a group of Foc strains (Fig. 4; Foc strains are indicated green). Most of the eight 'effector groups', as we will call them now, (A1-4, B1-2, C1-2) correspond to VCGs: group A1 corresponds to VCG0134, group A2 to VCG0136, group A4 to VCG0135, group B1 to a previously identified, as yet unnumbered VCG (see below), group B2 to VCG0133, group C1 to VCG0132 and group C2 to VCG0130 and VCG0131. Strain Fom025 (group A3) represents a separate lineage with no VCG assigned.

Effector genes are known to predominantly reside on accessory chromosomes that are amenable to horizontal transfer and therefore can have a different evolutionary history from the strains themselves, as represented by their core genomes (van Dam *et al*., 2016). To assess the evolutionary history of the 43 Fom strains, a phylogenetic tree was constructed based on the alignment of 422 core genes of the same strain set used for the effector pattern analysis described above. This resulted in three major clades, as commonly observed for the FOSC (O'Donnell *et al*., 1998) (Fig. 5). Fom strains are present in nine clades in the phylogenetic tree. These clades correspond to VCGs and have extremely little nucleotide diversity (represented by vertical lines in Fig. 5) and are therefore considered to be clonal lineages. Eight of these lineages are in clade 2 and one (Fom001) in clade 3. Each lineage corresponds to an effector group, except that two separate lineages (VCG0130 and VCG0131) both belong to effector group C2. *AVRFom2* ST2 is only present in the C2 effector group, while *AVRFom2* ST1 is present in groups A1–4 and B1. Remarkably, the two

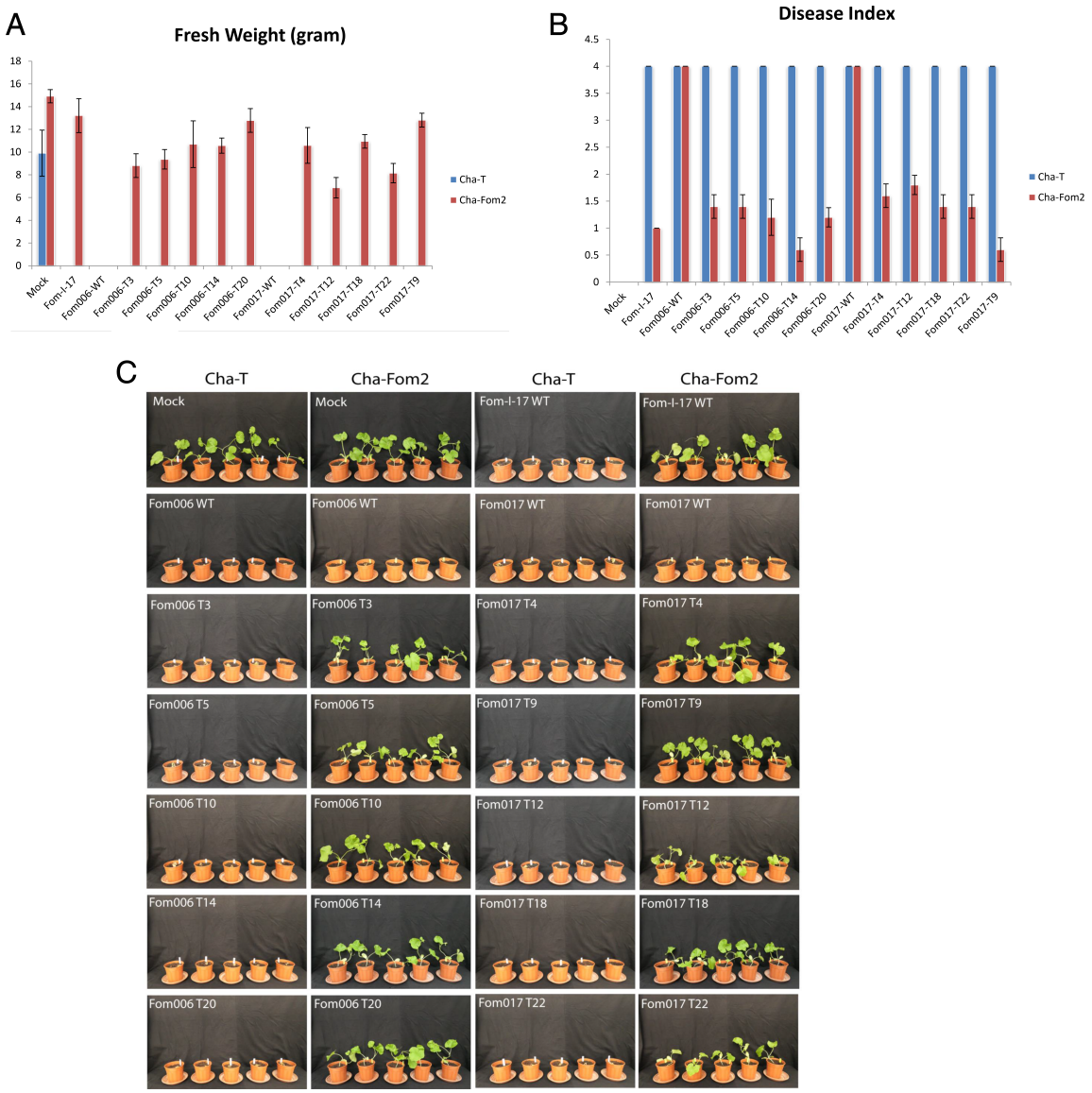

Fig 3. *AVRFom2* ST1 can confer avirulence to Fom strains carrying *AVRFom2* ST2 on melon plants harbouring the *Fom2* resistance gene. Genetic complementation of Fom006 (no *AVRFom2*) and Fom017 (*AVRFom2* ST2) with *AVRFom2* ST1 was performed by Agrobacterium-mediated transformation. Susceptible (Cha-T) and resistant (Cha-*Fom2*) 10-day-old melon seedling were inoculated with conidia of two wild type race 2 strains (Fom006, Fom017), 10 independent strains transformed with *AVRFom2* ST1 (Fom006 T3, T5, T10, T14, T20 and Fom017 T4, T12, T18, T22, T9) and a race 1 strain (Fom-I-17). Two weeks after inoculation, we determined the plant fresh weight (A) and disease index (B). Error bars indicated ± SD (n = 5). Representative plants 2 weeks after inoculation are shown in (C). The experiment was repeated twice with similar results. [Color figure can be viewed at wileyonlinelibrary.com]

strains of the C1 effector group and *F. oxysporum* f. sp. *cucumerinum* strain Foc035 belong to the same lineage (Fig. 5) and have similar effector patterns, while infecting different hosts (Fig. 4).

Iranian strains are present in only two of the nine lineages: A1/VCG0134 and B1. Lineage B1 contains strains from Iran and a single strain from Israel. The Israelian strain is race 1 and all Iranian strains in B1 belong to race 1.2. The B1 strains from Iran belong to an unassigned VCG described by Mirtalebi *et al*. (2013), which has not yet been numbered. The A1 effector group (VCG0134) is cosmopolitan: it contains Iranian strains of races 1 and 1.2, as well as strains from outside of Iran including all races. This implies that Iranian race 1.2 isolates in this lineage have either evolved from race 1 in Iran or have been imported into Iran.

Discussion

To create a comprehensive overview of the genomic diversity of Fom, and clues to the possible origin and relatedness of clonal lineages of this notorious melon

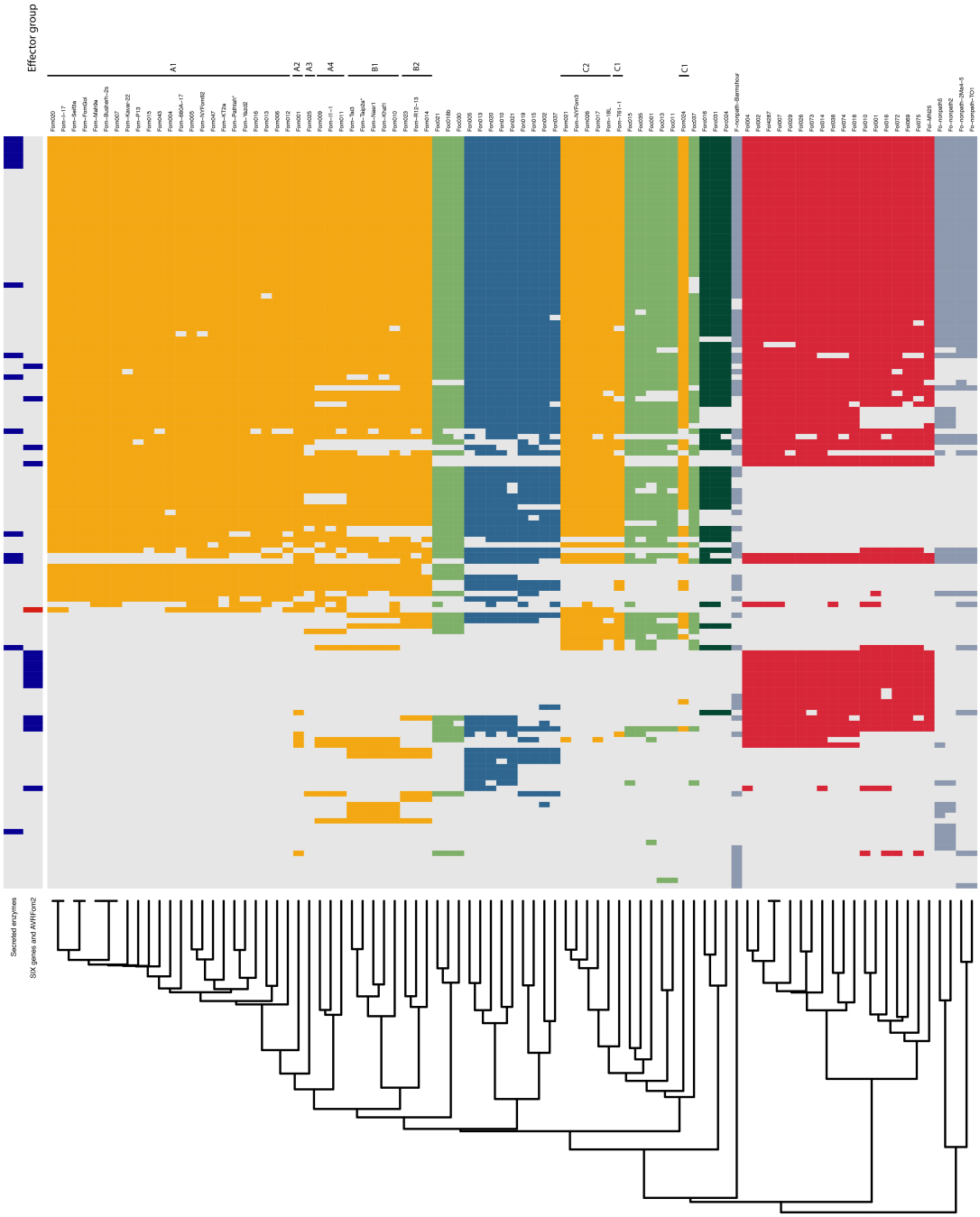

Fig 4. Clustering of *Fusarium oxysporum* strains based on presence/absence of 144 candidate effector genes in 87 genomes using megaBLAST. A colour indicates presence, grey indicates absence. Top: secreted enzymes and SIX genes are marked with blue, *AVRFom2* with red. Right: Effector groups of Fom strains: A (includes four subgroups, A1–A4); B (includes two subgroups, B1 and B2) and C (includes two subgroups, C1 and C2). F: *Fusarium*; Fom: *F. oxysporum* f. sp. *melonis*; Foc: *F. oxysporum* f. sp. *cucumerinum*; Fon: *F. oxysporum* f. sp. *niveum*; Forc: *F. oxysporum* f. sp. *radicis-cucumerinum*; Fol: *F. oxysporum* f. sp. *lycopersici*; Fo-nonpath: *F. oxysporum* non-pathogenic strains. [Color figure can be viewed at wileyonlinelibrary.com]

pathogen, we collected a diverse set of Fom strains from different sites in Iran and other parts of the world. We performed bio-assay experiments on differential cultivars of melon to confirm pathogenicity and race of 38 strains selected for genome sequencing.

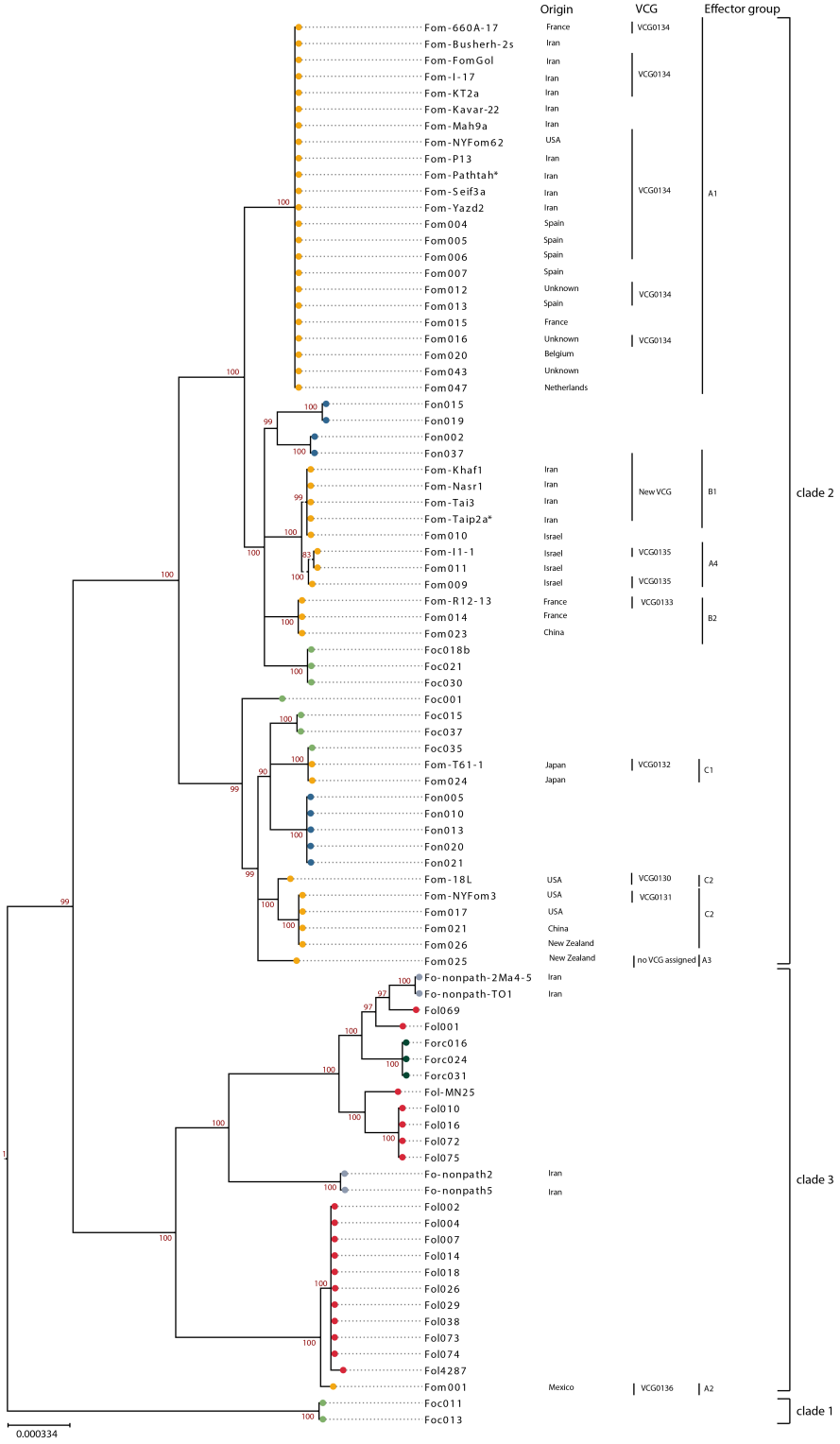

Fig 5. Phylogenetic relationships between Fom strains based on a concatenated alignment of 422 conserved genes. A concatenated sequence alignment was generated using 'strictplus' and a core phylogenetic tree with 100 bootstraps was generated using RAxML v8.2.12. *F-nonpath-Barmshour* was chosen as an outgroup (outgroup not shown). Strains that are coloured yellow belong to *F. oxysporum* f. sp. *melonis*, green belongs to *F. oxysporum* f. sp. *cucumerinum*, blue belongs to *F. oxysporum* f. sp. *niveum*, red belongs to *F. oxysporum* f. sp. *lycopersici* and those that are grey are non-pathogenic strains. The geographic origin, VCG (0130–0136), and effector group (A1-A4, B1-2, C1-2) of Fom strains are shown on the right. All strains are grouped into three major clades, as observed before (O'Donnell *et al*., 1998). [Color figure can be viewed at wileyonlinelibrary.com]

Based on a presumed gene-for-gene relationship, we expected that *AVRFom2*, which was identified by Schmidt *et al*. (2016), should be present in all strains of race 0 and race 1, but absent in the race 2 and race 1.2 strains. Surprisingly, however, we found that *AVRFom2* is also present in some race 2 strains as a new sequence type (*AVRFom2* ST2). This sequence type was found only in some race 2 strains in lineage C2, which were not analysed by Schmidt *et al*. (2016). *AVRFom2* ST2 contains 17 single nucleotide polymorphisms compared with the previously published sequence of *AVRFom2* (*AVRFom2* ST1) (Fig. S4A), which lead to 12 amino acid differences between the respective proteins (Fig. S4B) (Schmidt *et al*., 2016). The strains that contain *AVRFom2* ST2 are all virulent on *Fom2*-containing plants, unlike strains containing *AVRFom2* ST1. The result of *in planta* expression of *AVRFom2* ST2 and genetic complementation with *AVRFom2* ST1 in the background of *AVRFom2*-containing race 2 strains suggest that AvrFom2 is not recognized by the Fom2 resistance protein. A significant role in virulence seems unlikely for AvrFom2, since many highly virulent Fom strains (race 2 and 1.2) do not contain either ST1 or ST2 of *AVRFom2* (Schmidt *et al*., 2016 and Table 2).

Phylogenetic analysis revealed nine Fom lineages, two of which, A1/VCG0134 and B1, were found in Iran. In itself, the geographic distribution of Fom lineages and races cannot determine their origin or dispersal. However, this information does delimit a number of potential scenario's. Lineage A1/VCG0134 is the most cosmopolitan Fom lineage and contains all four races, of which race 1 and 1.2 were found in Iran. Since melon has been cultivated in Iran for millenia, and there is evidence for export from but not import of melon seeds into Iran (see Introduction), it is plausible that this lineage originated in Iran, with race 1.2 arising in this lineage from race 1 in Iran.

The second lineage containing Iranian strains, B1, is more limited in distribution: in Iran, it has been found in Khorasan province, and outside Iran only in Israel. The Iranian strains of lineage B1 belong to a newly defined VCG (Mirtalebi *et al*., 2013). The only Israelian strain of lineage B1 included in this study, Fom010, has no VCG assigned. Only the presence of Fom strains belonging to VCG0135 and VCG0138 has been described in Israel (Katan *et al*., 1994). VCG0138 was first identified in Israel and found to contain strains belonging to races 0, 1 and 1.2. Compatibility of the Iranian strains with VCG0138 was not tested by Mirtalebi *et al*. due to the lack of a tester strain for this VCG. Therefore, it is possible that Fom010 and the Iranian isolates belonging to lineage B1 correspond to VCG0138. If this is the case, this lineage may have been introduced into Iran because more races are present in Israel (0, 1 and 1.2) than in Iran (only 1.2). It was indeed suggested that the three races in VCG0138 may have originated in Israel (Katan *et al*., 1994).

Our results are in general accordance with Mirtalebi *et al*. (2013) who grouped Iranian strains into two different clades based on IGS sequences. However, according to our results, the novel, unassigned VCG group (possibly corresponding to VCG0138) and VCG0135 represent two distinct lineages (B1 and A4 respectively), while these two lineages were classified into a single group by Mirtalebi *et al*. (2013). In this case, IGS sequencing is not able to distinguish these lineages, while phylogeny based on whole-genome sequences can. Similarly, while a close genetic relationship was reported between the strains of the unassigned VCG and non-pathogenic strains (Fo-nonpath-2 Ma4-5 and Fo-nonpath-To1) based on IGS sequences (Mirtalebi *et al*., 2013), our results revealed that these strains are quite distinct based on whole-genome sequences.

By applying the methods described by Van Dam *et al*. (2016), 40 new candidate effectors in 38 newly sequenced genomes were identified in this work. Clustering based on presence/absence candidate effectors and marker PCR separated Fom strains into eight groups, A1–4, B1–2 and C1–2, which correspond to the nine lineages defined by core gene phylogeny, but with the C2 pattern being present in two relatively closely related lineages (VCG0130 and VCG0131, Fig. 5). However, effector patterns between some strains are much more similar than host-specificity (melon or cucumber) or genetic relationship (of Fom lineages) suggest.

A close similarity of accessory genetic material (candidate effector genes) between Fom017, Fom021, Fom024, Fom026 strains on the one hand and *F. oxysporum* f. sp. *cucumerinum* strains on the other was observed before (Van Dam *et al*., 2018). Strains Fom017, Fom021 and Fom026 belong to group C2/VCG0131 while, according to our phylogenetic analysis, the strains of group C1/VCG0132, Fom024 and Fom-T61-1 (which do not have *AVRFom2* ST2), belong to a separate lineage with a *F. oxysporum* f. sp. *cucumerinum* strain (Foc035). The strains in this Fom024/Fom-T61-1/Foc035 lineage have similar effector patterns (Fig. 4) and may represent a relatively recent host-shift, between cucumber and melon.

We also observed high similarity between effector patterns and effector sequences between some distant Fom lineages, such as between A1/VCG0134 and A3/Fom025 and between A1/VCG0134 and A2/Fom001. This suggests that horizontal chromosome transfer (HCT) has occurred between these lineages, and supports the notion that (HCT) has an important role in increasing genetic diversity in Fusarium and may underlie the generation of new (pathogenic) clonal lines (Ma *et al*., 2013;

Kang *et al*., 2014). Indeed, we have recently shown that Fom, like *F. oxysporum* f. sp. *lycopersici* (Fol) and *F. oxysporum* f. sp. *radicis-cucumerinum* (Forc), contains pathogenicity-conferring chromosomes that can be horizontally transferred between strains (Li *et al*., 2020). In a scenario in which the major, cosmopolitan lineage, A1/VCG0134, originated from Iran, HCT may have given rise to additional lineages outside Iran, such as A2/Fom001 and A3/Fom025.

Experimental procedures

Plant lines and fungal strains

The following muskmelon (*Cucumis melo* L.) cultivars were used: Cha-T, Cha-*Fom1* and Cha-*Fom2* (a kind gift from ENZA Zaden). While Cha-T does not include any known resistance gene, Cha-*Fom1* and Cha-*Fom2* contain the dominant resistance gene *Fom1* and *Fom2* respectively (Risser *et al*., 1976).

Fom strains were selected based on diversity of race, VCG and geographical origin. Detailed information about these strains is listed in Table 1.

Marker detection

Genomic DNA was isolated from 7- to 10-day-old mycelium scraped off a Czapek Dox Agar plate. The tissue was disrupted by vortexing it for 2 min in the presence of 400 µl of Tris-EDTA pH 8.0 (TE), 300 µl of phenol-chloroform (1:1) and glass beads. After this, a chloroform extraction was performed on the upper phase. The presence of nine candidate effector genes (66 Focuc, 21 Focuc, 1 Fomln, 20 Fomln, 18 Fomln + Foniv, 99 Foniv, 100 Foniv, 1 Fomom + Folag, 99 Folag) was determined with PCR using primers described previously in van Dam *et al*. (2018). To determine the presence of the *AVRFom2* gene, *AVRFom2* primers (fp8657 and fp8658) were designed using Primer3 software (http://primer3.ut.ee/) (Table S4). PCR mix and program were used as described in van Dam *et al*. (2018). *FEM1* primers were used as a positive control for gDNA quality and sterile Milli-Q was used as a negative control for each of the primer combinations instead of template DNA.

Whole-genome sequencing and de novo *assembly*

Strains were grown in liquid NO_3-medium (0.17% yeast nitrogen base, 3% sucrose and 100 mM KNO_3) for 7 days, mycelium was harvested and freeze-dried. Genomic DNA was isolated by shaking mycelial powder in a TissueLyser (Qiagen) for 2 min at 30 Hz in the presence of 800 µl of extraction buffer (100 mM Tris pH 8.0, 50 mM EDTA, 1 M NaCl and 3% SDS). After 30 min of incubation at 65°C, 800 µl buffer saturated phenol:chloroform: isoamyl alcohol (25:24:1) was added, mixed and centrifuged at maximum speed at 4°C for 15 min. The aqueous phase was transferred to a fresh tube and an equal volume of buffer saturated phenol:chloroform:isoamyl alcohol (25:24:1) was added, mixed and centrifuged at maximum speed at 4°C for 15 min. DNA was precipitated from the aqueous phase with 0.1 volume 5 M NaCl and 2 volumes 96% EtOH, incubation at −20°C for 15 min and centrifugation at maximum speed at 4°C for 15 min. The pellet was dissolved in 250 µl R2 buffer from the Purelink plant total DNA purification kit (Invitrogen) by vortexing. The DNA was then further purified with the purification kit (Invitrogen), according to the manufacturer's instructions. Library preparation was either performed by the Hartwig Medical Foundation using either the TruSeq Nano DNA Library Prep Kit for NeoPrep (NP-101-1001, Illumina) (Fomln017) or the TruSeq Nano DNA Low Throughput Library Prep Kit (20015964, Illumina) (Fomln023) in combination with TruSeq LT single-index adapters or by the RNA Biology and Applied Bioinformatics department at the University of Amsterdam using a NEBNext Ultra DNA Library Prep Kit for Illumina (E7370L, New England Biolabs) (all other genomes) in combination with NEBNext Multiplex Oligos for Illumina (96 Unique Dual Index Primer Pairs) (E6440S, New England Biolabs). Whole-genome sequencing was then performed at the Hartwig Medical Foundation on a HiSeq Illumina Xten system. Genome sequences were adapter and quality trimmed with Trimmomatic (v0.39) (Bolger *et al*., 2014) using the options ILLUMINACLIP: TruSeq3-PE-2.fa:2:20:8:4:false and SLIDINGWINDOW:4:20 MINLEN:100. *De novo* genome assemblies were made using CLC genomics workbench v 8.5. Default CLC settings were used, except that contigs below 500 bp were discarded. FastQ Screen (v0.14.0) and FastQC (v0.11.3) were used for quality control of both the raw and trimmed reads. Furthermore, the results of the Contamination Screen of NCBI were used to remove any contamination in the initial assemblies using a custom python script. If a contig became smaller than 500 bp after removal of the contamination it was discarded from the assembly. Both the FastQScreen and the Contamination Screen identified contamination with DNA from other species in Fom017 (primates), Fom023 (primates) and Fom047 (*Stenotrophomonas maltophilia*). For the first two strains the contamination was restricted to only a few contigs, which were (partially) removed from the final assembly. For Fom047 contaminated reads mapping only to the *Stenotrophomonas maltophilia* genome were first removed using bbsplit (BBMap v38.76) (https://sourceforge.net/projects/bbmap/) and then a new assembly was made. Any remaining

contamination was removed based on the results of the contamination screen. Strains of which genomes were sequenced in this study are listed in Table 1 and NCBI genome accession numbers of the genomes of other Fo strains used in this study are listed in Table S3.

Phylogenetic analysis of Fom strains

To generate a core phylogeny we searched for homologues of 440 conserved Fol4287 genes (van Dam *et al*., 2018) in 86 selected Fo genomes using megaBLAST with default parameters. We continued the analysis with 422 genes that had a single hit (e-value 0.001) that overlapped with at least 70% of the query sequence and showed at least 80% identity to the query in each genome. An alignment was made for each gene (including the query sequence) using the program MUSCLE (Edgar, 2004). A custom python script was then used to concatenate these alignments (https://github.com/marads/conserved_gene_tree_scripts). The concatenated alignment was subsequently trimmed using trimAl using the option 'strictplus' (Capella-Gutiérrez *et al*., 2009). A core phylogenetic tree with 100 bootstraps was generated using RAxML v8.2.12 with GTRGAMMA as substitution model (Stamatakis, 2014) and the trimmed, concatenated alignment of 422 conserved genes from which any columns with only undetermined data were removed as input. To visualize the tree we used ETE3 v3.1.1 (Huerta-Cepas *et al*., 2016) and a custom python script (https://github.com/pvdam3/ete3_plottree). Branches supported with bootstrap values below 80 were collapsed. Strain 'F-nonpath-Barmshour' was removed from the final tree, because it turned out not to be *F. oxysporum*.

Data deposition

The Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under BioProject PRJNA596396 and the raw reads have been deposited at the Sequence Read Archive under Project number PRJNA596396. See Table S3 for the individual accession numbers.

Identification of candidate effector genes

Prediction of candidate effector genes in Fom genome sequences was carried out as described previously (Van Dam *et al*., 2016) using the FoEC.py script (https://github.com/pvdam3/FoEC) with the following dependencies: NCBI blast (v2.9.0), signalP (v4.1) and Augustus (v3.3.3). Running this script led to the identification of 124 candidate effectors of which 74 showed no significant similarity to the previously identified 104 candidate effector genes described in van Dam *et al*. (2016). We continued further analysis with 40 of these 74 candidate effector genes, excluding genes missing a stop codon, with a mature protein length (=protein length without SP) below 35 amino acids or containing a SP that did not adhere to the following rules: (i) length SP between 15 and 30 amino acids, (ii) the SP contains a processing site: the last three residues are a small amino acid (G, A, V, S, T, C), followed by any amino acid, followed by another small amino acid, (iii) the SP contains a hydrophobic stretch of 9–20 consecutive hydrophobic residues (A, V, I, L, M, F, W, C, G, P) that starts at least after the first and at most after the eight residues, ends at most two residues before the last three residues, contains no hydrophilic residues (R, K, D, E, Q, N, H) and contains a maximum of three hydroxylated residues (S, T, Y). We made a final candidate effector list containing 144 genes: the 104 genes described in Van Dam *et al*. (2016) and the 40 new genes identified in this article.

Disease assays

Disease assays were performed using the root-dip method (Wellman, 1939). Conidia were isolated from 5-day-old NO_3 cultures by filtering through two layers of miracloth (Merck; pore size of 22–25 μm) and resuspending them in sterile MilliQ water. Spore concentration was estimated using a haemocytometer and then diluted to a final concentration of 1×10^7 spores ml^{-1}. Ten-day-old seedlings were up-rooted, rinsed with water, inoculated for ~5 min and then re-planted in individual pots. For race determination plants were planted in soil, while for expression analysis plants were planted in vermiculite. Plants were then grown at 25°C and 65% relative humidity in the greenhouse. To determine the virulence and race of Fom strains, plants were harvested 2 weeks after treatment. Each plant was cut at cotyledon level, plant weight was determined and a disease index was given from 0 to 4 (0: healthy plant; 1: small brown lesion belowground and/or slight root rot symptoms only at the tip of the main root; 2: root rot symptoms and/or stem lesions visible above ground; 3: very clear root rot symptoms of the entire root system and large lesion extending above the cotyledons; 4: plant completely dead or very small and wilted). At least four plants of each melon cultivar were used per fungal strain and each bioassay was performed at least twice with similar results.

AVRFom2 in planta *expression analysis*

For *in planta* expression analysis, 10-day old melon seedlings were root-inoculated with spore suspensions as described above. Hypocotyls and roots were harvested 10 days after treatment. Hypocotyl and root material of five plants/treatment was pooled. Total RNA

extraction and cDNA synthesis was performed as described previously (Schmidt *et al*., 2013). Primers for quantitative real-time PCR (qPCR) were designed using Primer3 software (http://primer3.ut.ee/) and are listed in Table S4. qPCR was carried out on a QuantStudio 3 Real-Time PCR system (Thermo Fisher Scientific) with the following two-step program: 95°C for 2 min; 40 cycles (95°C for 15 s, 62°C for 1 min) and a melting curve step at the end. All reactions were performed in a total volume of 10 μl containing 1 μl of 10× Super *Taq* buffer, 0.2 μl of dNTPs (10 mM each), 1 μl of each primer (3 pmol μl^{-1}), 0.05 μl of Super *Taq* polymerase (SphaeroQ), 0.02 μl of ROX reference dye (Jena Bioscience), 0.2 μl of Evagreen fluorescent DNA stain (Jena Bioscience), 5.53 μl sterile Milli-Q and 1 μl of 16× diluted template DNA. Three technical and two biological replicates were performed for each treatment to confirm the reproducibility of the results. Sterile Milli-Q instead of DNA template was used as a negative control for each of the primer combinations.

Transformation of Fo with AVRFom2

T-DNA plasmids carrying *AVRFom2* (ST1) were previously generated by Schmidt *et al*. (2016) by amplifying the Fom001 *AVRFom2* ORF and flanking sequences with primers fp5225 and fp5226 (Table S4) and placing the PCR products in pRWh2. Genetic complementation was achieved by agrobacterium-mediated transformation of Fom006 (no *AVRFom2*) and Fom017 (*AVRFom2* ST2) with this construct as described earlier (Takken *et al*., 2004). Putative transformants were selected on cefotaxime (0.2 g ml^{-1}) and hygromycin (40 mg ml^{-1}) and monospored before further analysis. Complete transfer of the promotor, ORF and terminator of *AVRFom2* ST1 in hygromycin-resistant colonies was confirmed by PCR using primers designed with Primer3 software (http://primer3.ut.ee/). Primers are listed in Table S4. Then, as described above, the fungal spore suspensions were used to inoculate five 10-day-old melon seedlings. Two weeks after inoculation, the fresh plant weight and a disease index score were used to determine disease severity.

Acknowledgements

We are thankful to the ENZA Zaden for providing the melon Cha-T, Cha-Fom1 and Cha-Fom2 seeds. We are grateful to the Ministry of Science, Research Technology of Iran for financial support, and the Department of Molecular Plant Pathology at the University of Amsterdam for the warm hospitality.

References

Armstrong, G.M., and Armstrong, J.K. (1978) *Formae speciales* and races of *Fusarium oxysporum* causing wilts of cucurbitaceae. *Phytopathology* **68**: 19–28.

Banihashemi, Z. (1968) The existence of fusarium wilt of melon in Iran. *Proc First Nat Cong Plant Med*: 47–48.

Banihashemi, Z. (1982) A new physiological race of *Fusarium oxysporum* f. sp melonis in Iran Iranian. *J Plant Pathol* **18**: 1–6.

Banihashemi, Z. (1989) The existence of race 1 of *Fusarium oxysporum* f. sp. *melonis* on longmelon in Garmsar and its virulence to different cultivars of *Cucumis melo*. Proceedings of the 9th Plant Protection Congress (p. 91).

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Brotman, Y., Normantovich, M., Goldenberg, Z., Zvirin, Z., Kovalski, I., Stovbun, N., *et al*. (2013) Dual resistance of melon to *Fusarium oxysporum* races 0 and 2 and to *papaya ring-spot virus* is controlled by a pair of head-to-head-oriented NB-LRR genes of unusual architecture. *Mol Plant* **6**: 235–238.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

Hammond-Kosack, K.E., and Jones, J.D.G. (1997) Plant disease resistance genes. *Annu Rev Plant Physiol Plant Mol Biol* **48**: 575–607.

Huerta-Cepas, J., Serra, F., and Bork, P. (2016) ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol Biol Evol* **33**: 1635–1638. https://doi.org/10.1093/molbev/msw046.

Jacobson, D.J., and Gordon, T.R. (1988) Vegetative compatibility and self-incompatibility within *Fusarium oxysporum* f. sp *melonis*. *Phytopathology* **78**: 668–672.

Jacobson, D.J., and Gordon, T.R. (1990) Further investigation of vegetative compatibility with *Fusarium oxysporum* f. sp. *melonis*. *Can J Bot* **68**: 1245–1248.

Jacobson, D.J., and Gordon, T.R. (1991) *Fusarium oxysporum* f. sp. *melonis*: a case study of diversity within a *forma specialis*. *Phytopathology* **81**: 1064–1067.

Joobeur, T., King, J.J., Nolin, S.J., Thomas, C.E., and Dean, R.A. (2004) The fusarium wilt resistance locus Fom-2 of melon contains a single resistance gene with complex features. *Plant J* **39**: 283–297.

Kang, S., Demers, J., del Mar Jimenez-Gasco, M., and Rep, M. (2014) Fusarium oxysporum. In *Genomics of Plant-Associated Fungi and Oomycetes: Dicot Pathogens*, Dean, R.A., Lichens-Park, A., and Kole, C. (eds). Berlin, Germany: Springer, pp. 99–119.

Katan, T., Katan, J., Gordon, T.R., and Pozniak, D. (1994) Physiologic races and vegetative compatibility groups of *Fusarium oxysporum* f. sp. *melonis* in Israel. *Phytopthology* **84**: 153–157.

Kim, D.H., Martyn, R.D., and Magill, C.W. (1993) Mitochondrial DNA (mtDNA)-relatedness among formae speciales

of *Fusarium oxysporum* in the Cucurbitaceae. *Phytopathology* **83**: 91–97. https://doi.org/10.1094/Phyto-83-91.

Li, J., Fokkens, L., Conneely, L.J., and Rep, M. (2020) A partial pathogenicity chromosome in *Fusarium oxysporum* is sufficient to cause disease and can be horizontally transferred. *Mol Plant Pathol* **21**: 761–776. https://doi.org/10.1111/mpp.12927.

Ma, L.J., Geiser, D.M., Proctor, R.H., Rooney, A.P., O'Donnell, K., Trail, F., *et al*. (2013) Fusarium pathogenomics. *Annu Rev Microbiol* **67**: 399–416.

Martin, G.B., Bogdanove, A.J., and Sessa, G. (2003) Understanding the functions of plant disease resistance proteins. *Annu Rev Plant Biol* **54**: 23–61.

Mirtalebi, M., Banihashemi, Z., and Linde, C.C. (2013) Phylogenetic relationships of *Fusarium oxysporum* f. sp. *melonis* in Iran. *Eur J Plant Pathol* **136**: 749–762.

Namiki, F., Shiomi, T., Kayamura, T., and Tsuge, T. (1998) Pathogenic and genetic variation in the Japanese strains of *Fusarium oxysporum* f. sp. *melonis*. *Phytopathology* **88**: 804–810.

O'Donnell, K., Kistler, H.C., Cigelnik, E., and Ploetz, R.C. (1998) Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. *Appl Bio Sci* **95**: 2044–2049.

Oumouloud, A., El-Otmani, M., Chikh-Rouhou, H., Claver, A. G., Torres, R.G., Perl-Treves, R., and Alvarez, J.M. (2013) Breeding melon for resistance to *Fusarium wilt*: recent developments. *Euphytica* **192**: 155–169.

Risser, G., Banihashemi, Z., and Davis, D.W. (1976) Proposed nomenclature of *Fusarium oxysporum* f. sp. melonis races and resistance genes in *Cucumis melo*. *Phytopathology* **66**: 1105.

Sarpeleh, A., and Banihashemi, Z. (2000) Vegetative compatibility groups within races of *Fusarium oxysporum* f. sp. *melonis* in Iran and *Fusarium oxysporum* from weeds in Maharloo region of Fars province of Iran. *Iran J Plant Pathol* **36**: 9–15.

Schmidt, S.M., Houterman, P.M., Schreiver, I., Ma, L., Amyotte, S., Chellappan, B., *et al*. (2013) MITEs in the promoters of effector genes allow prediction of novel virulence genes in *Fusarium oxysporum*. *BMC Genomics* **14**: 119.

Schmidt, S.M., Lukasiewicz, J., Farrer, R., van Dam, P., Bertoldo, C., and Rep, M. (2016) Comparative genomics of *Fusarium oxysporum* f. sp. *melonis* reveals the secreted protein recognized by the Fom-2 resistance gene in melon. *New Phytol* **209**: 307–318.

Schreuder, W., Lamprecht, S.C., and Holz, G. (2000) Race determination and vegetative compatibility grouping of *Fusarium oxysporum* f. sp. *melonis* from South Africa. *Plant Dis* **84**: 231–234.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Szabo, Z., Gyulai, G., Toth, Z., and Heszky, L. (2008) Morphological and molecular diversity of 47 melon (*Cucumis melo*) cultivars compared to an extinct landrace excavated from the 15th century. Cucurbitaceae 2008, Proceedings of the IXth Eucarpia meeting on genetics and breeding of Cucurbitaceae (pp. 313–321). INRA, Avignon, France.

Takken, F.L.W., Wijk, R. van, Michielse, C.B., Houterman, P.M., Ram, A.F.J., and Cornelissen, B.J.C. (2004). A one-step method to convert vectors into binary vectors suited for Agrobacterium-mediated transformation. Cur. *Genet*, **45**, 242–248.

van Dam, P., de Sain, M., ter Horst, A., van der Gragt, M., and Rep, M. (2018) Use of comparative genomics-based markers for discrimination of host specificity in *Fusarium oxysporum*. *Appl Environ Microbiol* **84**: e01868-17.

van Dam, P., Fokkens, L., Schmidt, S.M., Linmans, J.H.J., Kistler, H.C., Ma, L.J., and Rep, M. (2016) Effector profiles distinguish formae speciales of *Fusarium oxysporum*. *Environ Microbiol* **18**: 4087–4102.

Wellman, F.L. (1939) A technique for studying host resistance and pathogenicity in tomato fusarium wilt. *Phytopathology* **29**: 945–956.

Zohary, D., and Hopf, M. (2000) *Domestication of Plants in the Old World*, 3rd ed. New York: Oxford University Press, p. 316.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Statistics of de novo assembled genomes.

Table S2. Candidate effector genes that were identified in 38 Fom strains. Previously 104 candidate effectors were identified by van Dam *et al*. A search for genes with a signal peptide in the proximity of mimp TIRs in 38 Fom genomes sequenced in this study resulted in the identification of 74 new candidate effectors, which were numbered 105–178. The first 40 were used for further analysis and are indicated in black, the rest was not used for further analysis and is therefore indicated in grey. Columns B-F show genomic DNA, transcript, protein, signal peptide and mature protein sequence respectively. Column G-I show protein, signal peptide and mature protein length. Column J and K indicate the amount of cysteines in the entire or the mature (without signal peptide) protein sequence. To assign possible functions to the candidate effectors a blastP search was performed against the non-redundant nucleotide database (NCBI). Column L-K shows the ID, the description, the percent identity the percent query coverage, the E-value and the bitscore of the best blastP hit found. Column R indicates why a candidate effector was not used for further analysis.

Table S3. *Fusarium oxysporum* strains used in this study and their GenBank assembly accession numbers.

Table S4. Primers used in this study.

Fig. S1. Determination of pathogenicity towards melon and race of Fom strains. Disease assays of Fom strains were performed on ten-day-old melon seedlings using the root-dip method. To determine race, three differential hosts including Cha-T (no resistance gene), Cha-*Fom1* (*Fom1* resistance gene) and Cha-*Fom2* (*Fom2* resistance gene) were used. At least four plants of each melon cultivar were inoculated with 1×10^7 spores ml^{-1} of each fungal strain. Two weeks after inoculation, the plant fresh weight (a) and a disease index score (b) were determined as described in experimental procedures. Mock-inoculated seedlings were used as a negative control. Error bars indicate ± SD (n = 4 for Cha-T

inoculated with Fom-Pathtah, Fom-Bushehr-2 s, Fom-KT2a, Fom-R12-13, Fom024, Fom027 or n = 5 for the other strains). The experiment was repeated at least twice with similar results unless indicated otherwise in Table 1.

Fig. S2. Confirmation of presence of *AVRFom2* in putative transformants by PCR. To investigate the possibility that recognition of *AVRFom2* ST2 by *Fom2* is suppressed in the background of the race 2 strains that produce it, we transformed the race 2 strains Fom006 (no *AVRFom2*) and Fom017 (*AVRFom2* ST2) with *AVRFom2* (ST1) (Schmidt *et al*., 2016). We selected 19 putative transformants of Fom006 and 20 putative transformants of Fom017. The presence of *AVRFom2* (ST1) in Fom006 was confirmed by primes fp2620/fp8914, fp8657/fp8915 and fp8916/fp659 (a, b, c). These three primers were also used for confirmation of the presence of *AVRFom2* (ST1) in Fom017 (d, e, f). The primers fp8657/fp8917 were used to check if any of the transformants were in-locus: replacing AVRFom2 ST2 in Fom017 with AVRFom2 ST1 (g). The sequences of these primers and target genes are listed in Table S4. Genomic DNA (gDNA) of Fom017, Fom-I-17, Fom006 and water were used as negative controls for primers fp2620/8914 and fp8916/659. gDNA of Fom017 was used as a positive control and gDNA of Fom-I-17, Fom006 and water were used as negative controls for primers fp8657/8917. gDNA of Fom-I-17 was used as a positive control and gDNA of Fom017, Fom006 and water were used as negative controls for primers fp8657/8915.

Fig. S3. Clustering of *Fusarium oxysporum* strains based on presence/absence of 144 candidate effector genes in 87 genomes using megaBLAST. A colour indicates presence, grey indicates absence. Top: secreted enzymes and SIX genes are marked with blue, *AVRFom2* with red. Right: Effector groups of Fom strains: A (includes four subgroups, A1-A4); B (includes two subgroups, B1 and B2) and C (includes two subgroups, C1 and C2). F: *Fusarium*; Fom: *F. oxysporum* f. sp. *melonis*; Foc: *F. oxysporum* f. sp. *cucumerinum*; Fon: *F. oxysporum* f. sp. *niveum*; Forc: *F. oxysporum* f. sp. *radicis-cucumerinum*; Fol: *F. oxysporum* f. sp. *lycopersici*; Fo-nonpath: *F. oxysporum* nonpathogenic strains. This is the same figure as Fig. 4, except that here the annotation of the candidate effectors is indicated below.

Fig. S4. Sequence alignment of sequence type 1 (ST1) and sequence type 2 (ST2) of *AVRFom2*. The upstream flank of *AVRFom2*, the *AVRFom2 ORF*, and the downstream flank *AVRFom2* are indicated with blue, black, and red lines respectively. Between *AVRFom2* ST1 and ST2 are 17 single nucleotide polymorphisms (SNPs) (a). These 17 SNPs lead to 12 amino acid differences between *AVRFom2* ST1 and ST2 (b). Sequence differences between *AVRFom2* ST1 and ST2 are indicated with a grey background colour. The ST1-specific and ST2-specific primers used in this study are shown by yellow boxes (ST1: fp8657/fp8915, ST2: fp8657/fp8917).