A new eIF4E1 allele characterized by RNAseq data mining is associated with resistance to PVY in tomato albeit with a low durability

Caroline Lebaron, Aurélie Rosado, Christopher Sauvage, Camille Gauffier, Sylvie S. German-Retana, Benoît Moury, Jean-Luc J.-L. Gallois

Journal of General Virology

A new eIF4E1 allele characterized by RNAseq data mining is associated with resistance to PVY in tomato albeit with a low durability

--Manuscript Draft--

Manuscript Number:	JGV-D-16-00429R1
Full Title:	A new eIF4E1 allele characterized by RNAseq data mining is associated with resistance to PVY in tomato albeit with a low durability
Short Title:	RNAseq data mining for eIF4E resistance allele
Article Type:	Standard
Section/Category:	Plant - RNA Viruses
Corresponding Author:	Jean-Luc Gallois INRA UR-1052 Centre PACA Montfavet, FRANCE
First Author:	Caroline Lebaron
Order of Authors:	Caroline Lebaron
	Aurélie Rosado
	Christopher Sauvage
	Camille Gauffier
	Sylvie German-Retana
	Benoît Moury
	Jean-Luc Gallois
Abstract:	Allele mining on susceptibility factors offers opportunities to find new sources of resistance among crop wild relatives for breeding purposes. As a proof of concept, we used available RNAseq data to investigate polymorphisms among the four tomato genes encoding translation initiation factors (eIF4E1 and eIF4E2, eIFiso4E and the related gene New Cap-binding Protein -nCBP) to look for new potential resistance alleles to potyviruses. By analyzing polymorphism among RNAseq data obtained for 20 tomato accessions, ten belonging to the cultivated type Solanum lycopersicum and 10 belonging to the closest related wild species Solanum pimpinellifolium, we isolated one new eIF4E1 allele, in the Solanum pimpinellifolium LA0411 accession, which encodes a potential new resistance allele, mainly due to a polymorphism associated with an amino acid change within eIF4E1 region II. We confirmed that this new allele, pot1², is indeed associated with resistance to PVY, although with a restricted resistance spectrum and a very low durability potential. This suggests that mutations occurring in eIF4E region II only may not be sufficient to provide efficient and durable resistance in plants. However, our study emphasizes the opportunity brought by RNAseq data to mine for new resistance alleles. Moreover, this approach could be extended to seek for putative new resistance alleles by screening for variant forms of susceptibility genes encoding plant host proteins known to interact with viral proteins.

A new *eIF4E1* allele characterized by RNAseq data mining is associated with resistance to PVY in tomato albeit with a low durability

Caroline Lebaron[1], Aurélie Rosado[1], Christopher Sauvage[1], Camille Gauffier[1], Sylvie German-Retana[2], Benoît Moury[3], Jean-Luc Gallois[1]

[1] *GAFL, INRA, 84140, Montfavet, France*

[2] *BFP, INRA, Université Bordeaux, 33882, Villenave d'Ornon, France*

[3] *Pathologie Végétale, INRA, 84140, Montfavet, France*

Running title: RNAseq data mining for *eIF4E* resistance allele

Author for correspondence: Jean-Luc Gallois
tel: +00 33 (0)4 32 72 27 96
e-mail: jean-luc.gallois@inra.fr

Keywords: allele mining, RNAseq, lycopersicum spp., eIF4E, *Potato virus Y*, *Tobacco etch virus*

Total Word count: 5105, including abstract 224

1 Figure, 6 Tables, 2 sup Fig, 1 sup Table

The GenBank accession numbers for the haplotype sequences of eIF4E1 are KX855953 to KX855956.

Abstract

Allele mining on susceptibility factors offers opportunities to find new sources of resistance among crop wild relatives for breeding purposes. As a proof of concept, we used available RNAseq data to investigate polymorphisms among the four tomato genes encoding translation initiation factors (*eIF4E1 and eIF4E2, eIFiso4E* and the related gene *New Cap-binding Protein –nCBP*) to look for new potential resistance alleles to potyviruses. By analyzing polymorphism among RNAseq data obtained for 20 tomato accessions, ten belonging to the cultivated type *Solanum lycopersicum* and 10 belonging to the closest related wild species *Solanum pimpinellifolium*, we isolated one new *eIF4E1* allele, in the *Solanum pimpinellifolium* LA0411 accession, which encodes a potential new resistance allele, mainly due to a polymorphism associated with an amino acid change within eIF4E1 region II. We confirmed that this new allele, *pot1*^{2}, is indeed associated with resistance to PVY, although with a restricted resistance spectrum and a very low durability potential. This suggests that mutations occurring in eIF4E region II only may not be sufficient to provide efficient and durable resistance in plants. However, our study emphasizes the opportunity brought by RNAseq data to mine for new resistance alleles. Moreover, this approach could be extended to seek for putative new resistance alleles by screening for variant forms of susceptibility genes encoding plant host proteins known to interact with viral proteins.

Introduction

One of the main challenges of plant breeders is to identify new sources of resistance to pathogens. Crop wild relatives are the main source of such resistances that can be monitored through the direct phenotyping of collection of accessions with the pathogen of interest. However this approach is both time-consuming and costly. Allele mining, which entails looking for polymorphism within genes of interest, represents a complementary approach to plant phenotyping in order to isolate new sources of resistance to pathogens among the natural variation. Viruses are characterized by a particularly small genome and rely on host factors to infect their host (Fraser, 1990). Hence, polymorphism within those host factors can be associated with the plant resistance, a mechanism dubbed as loss-of-susceptibility (Pavan *et al.,* 2010; van Schie & Takken, 2014).

Such approach has been exemplified by the characterization of resistance to viruses based on the eukaryotic translation initiation factors eIF4E (Robaglia & Caranta, 2006; Wang & Krishnaswamy, 2012). Translation initiation factors 4E (*i.e.* eIF4E and the isoform eIFiso4E) are essential components of the eukaryotic cell that initiate translation by binding to the mRNA cap structure at the 5' end of most mRNAs. They are encoded by a small multigene family (Browning & Bailey-Serres, 2015). In the last decade, those factors have been shown to be associated with resistance to a broad range of single-stranded positive-sense RNA viruses and especially potyviruses, including *Potato virus Y* (PVY) and *Tobacco etch virus* (TEV) (Wang and Krishnaswamy, 2012). Natural resistance to potyviruses in lettuce (Nicaise *et al.*, 2003), pepper (Ruffel *et al.,* 2002), pea (Gao *et al.,* 2004) and tomato (Ruffel *et al.,* 2005) has been shown to rely on non-synonymous substitutions in the eIF4E coding sequences. Those substitutions are mainly located in two regions, named I and II, located near the cap binding pocket in the eIF4E (Robaglia & Caranta, 2006). However, the precise involvement of those two regions in resistance remains unclear. In pepper, a large set of *eIF4E1* resistance alleles has been characterized (see below). While mutations in region I of pepper *eIF4E1* are associated with resistance to PVY, the additional mutations in region II might be associated with enlargement of the

resistance spectrum to TEV (Charron *et al.,* 2008; Yeam *et al.,* 2007). However, no natural allele harboring only resistance-associated mutations in region II has been isolated so far among the pepper natural variation. Beside the need for new resistance alleles, it is also crucial to determine the minimal set of mutations within eIF4E allowing the broadest resistance spectrum possible, in the wake of new breeding technologies such as CRISPR/Cas9 (Andersen *et al.,* 2015). So far, the proof-of-concept of using this technology to develop resistance has been illustrated by knocking out eIF4E and eIFiso4E genes in *Cucumis sativus* and *Arabidopsis thaliana*, respectively (Chandrasekaran *et al.,* 2016; Pyott *et al.,* 2016). However, recent work in tomato suggests that, given the redundancy effect among eIF4E genes, CRISPR/Cas9 technology should rather be used to design functional alleles by introducing non-synonymous mutations in the eIF4E coding sequence, rather than using null alleles (Gauffier *et al.,* 2016).

Given their central role in resistance to RNA-viruses, eIF4E family members have also been a target of choice for proof-of-concept studies as well as for a large number of allele mining strategies. In pepper, direct sequencing of eIF4E1 cDNAs sampled on different accessions has characterized 9 *pvr2* recessive resistance alleles to PVY and TEV (Charron *et al.*, 2008; Ruffel *et al.*, 2002). Similar approaches using the EcoTILLING strategy or the High Resolution Melting (HRM) analysis have extended this repertoire up to 22 alleles (Ibiza *et al.,* 2010; Jeong *et al.,* 2011). Such studies have also been carried out in other crops such as melon (Nieto *et al.,* 2007), barley (Hofinger *et al.,* 2011) or pea (Konecna *et al.,* 2014), demonstrating the usefulness of the approach. In the last decade, the development of high-throughput sequencing, especially RNAseq, has provided large amounts of publicly available data and precise knowledge of the gene expression patterns in major crop plants, as well as giving access to the polymorphism associated with amino acid (AA) changes in the encoded proteins. Those data represent a largely untapped resource to mine for new alleles of interest (Barabaschi *et al.,* 2016). However, it is important to check whether the available data often collected to provide insight on plant evolution and analysis of speciation, in tomato for example (Koenig *et al.,* 2013; Pease *et al.,* 2016), are well suited for the specific purpose of allele mining.

Compared to the large amount of eIF4E resistance alleles discovered in *Capsicum* Spp., only one resistance allele of the *eIF4E1* tomato orthologous gene was discovered in the accession PI247087 (hereafter PI24). This allele is associated with a large resistance spectrum to several isolates of PVY and TEV (Legnani *et al.,* 1995; Legnani *et al.*, 1996). Because PI24 is an accession of the wild species *Solanum habrochaites*, introgression of the resistance into cultivated tomato is difficult (Bernacchi & Tanksley, 1997; Gauffier *et al.*, manuscript in preparation). In this study, we investigated polymorphisms among the four tomato genes encoding translation initiation factors (*eIF4E1 and eIF4E2, eIFiso4E* and the related gene *New Cap-binding Protein –nCBP*) to seek for new potential resistance alleles. To this aim, we looked for polymorphism among RNAseq data obtained for 20 tomato accessions, ten belonging to the cultivated type *Solanum lycopersicum* and 10 belonging to the closest related wild species *Solanum pimpinellifolium* (Sarah *et al*., 2016 and available at http://www.arcad-project.org/) and isolated one new *eIF4E1* allele, *pot1*^{2}, in the *Solanum pimpinellifolium* LA0411 accession. Our results show the high feasibility of using RNAseq data to mine for new resistance alleles and can be extended to other host susceptibility factors. Moreover, our data bring new insights into the minimal set of non-synonymous mutations in the *eIF4E1* genes that can be useful to design CRISPR/Cas9 variant alleles with a broad and durable resistance spectrum.

Results

Analysis of RNAseq data among 20 tomato accessions allows the characterization of a new eIF4E1 allele in the S pimpinellifolium LA0411 accession.

In plants, translation initiation factors are encoded by a small multigene family. In tomato, we previously identified three genes encoding initiation factors 4E, including one isoform eIFiso4E and two eIF4E homologs eIF4E1 and eIF4E2 that are involved in susceptibility to potyviruses (Piron *et al.,* 2010; Ruffel *et al.*, 2005). Verification of the *eIF4E2* annotation on the Solgenomics database (https://solgenomics.net/) revealed that the exon structure of the *eIF4E2* gene differs from the gene

model characterized earlier, as the fifth exon was not included. The annotation (Solyc02g021550) was revised accordingly. By analysing available Expressed Sequence Tags (EST) in Genbank database and using the reference tomato genome (The Tomato Genome Consortium, 2012), we further identified a tomato gene encoding a homolog to the *A. thaliana* novel Cap Binding Protein (nCBP), characterized by the substitution of two of the eight tryptophan residues conserved in eIF4E proteins by phenylalanine and tyrosine, respectively (Ruud *et al.,* 1998; Supplemental Figure S1).

Recently, transcriptome data of a set of 20 tomato accessions was obtained through the Arcad project (http://www.arcad-project.org/) using paired-end HiSeq2000 sequencing, aiming at deciphering the molecular footprints of domestication in crop species, such as tomato or rice. The experimental design was based on a comparative approach of 10 accessions belonging to cultivated tomato *Solanum lycopersicum* and *S.lycopersicum cerasiforme* and 10 accessions from the closest wild species *Solanum pimpinellifolium* (Supplemental Table 1). The cDNA sequences for the 4 genes encoding initiation factors, namely eIF4E1, eIF4E2, eIFiso4E and nCBP, were retrieved (Table 1) and polymorphisms called according to the approach detailed in Nabholz *et al.* (2014) were sought after.

Nucleotide sequences could be retrieved for the four genes in the 20 accessions, showing that those genes are well expressed in leaf tissues used to create the cDNA libraries. No non-synonymous amino-acid changing substitution was found among eIF4E2, eIFiso4E and nCBP coding sequences. Moreover, no nucleotide polymorphism at all could be found in either eIF4E2 or eIF4iso4E, whereas two silent substitutions were found in the nCBP coding sequence, including one common to all *S. pimpinellifolium* species sequenced. In comparison, four new haplotypes within the eIF4E1 coding sequence were characterized for tomato eIF4E1 and, as previously shown for the eIF4E1 homolog in pepper, all polymorphisms were associated with AA changes within the eIF4E1 coding sequence (Figure 1). For haplotypes 1 to 3, only one mutation was found, that was located outside the regions I and II where mutations associated with resistance are usually located (Robaglia & Caranta, 2006). More significantly, the LA0411 accession (haplotype 4) displayed two non-synonymous substitutions

within eIF4E1 sequence, including one mutation located at position 112 in the region II, where an aspartic acid (D) is replaced by a glycine (G) (D112G, Figure 1).

The presence of this polymorphism was confirmed in the LA0411 *eIF4E1* allele by Sanger sequencing. The rest of the study was focused on the LA0411 allele, as it encodes a protein with an AA change within the region II but not in region I. It was of particular interest to see whether the region II mutation by itself provides resistance to potyviruses.

The eIF4E1 allele of accession LA0411 is a new pot1^2 resistance allele to two PVY isolates

We first checked whether the *S. pimpinellifolium* LA0411 accession displayed any resistance to potyviruses. Thus, LA0411 plants were mechanically inoculated with three isolates of PVY (LYE90, N605 and SON41g) and three isolates of TEV (CAA10, S103 and HAT). The resistance phenotype was assayed by titrating viral accumulation by ELISA at 21 days post-inoculation, in the systemically-infected leaves. Susceptible *S. lycopersicum* plants M82, as well as the highly resistant *S. habrochaites* PI24 accession, which harbors the *eIF4E1-pot1* allele, were used as positive and negative controls, respectively. Our results showed that the LA0411 plants display resistance to the PVY isolates LYE90 and N605 and is susceptible to PVY SON41g as well as to all three TEV isolates tested (Table 2 and supplemental figure 2). Therefore, LA0411 is associated with a reduced but significant resistance spectrum to potyviruses.

The genetic basis of LA0411 resistance to potyviruses was further assessed using PVY N605 isolate. LA0411 plants were manually crossed with the susceptible *S. lycopersicum* M82 accession. F1 plants were crossed to each parent to produce resistant and susceptible backcrosses, respectively. F2 were produced by selfing F1 plants. All populations were tested for resistance or susceptibility to PVY N605 (Table 3). Both the F1 and the susceptible backcross progeny plants were found to be fully susceptible to PVY N605, suggesting the presence of a recessive resistance gene. In comparison,

resistant plants did segregate among the resistant backcross (BC1R) and the F2 progeny, in accordance with the presence of a single recessive gene.

Finally, an allelism test was carried out between the broad resistance allele *eIF4E1-pot1* that originates from the wild accession PI24 (Parrella *et al.,* 2002) and the LA0411 *eIF4E1* allele. Wild species belonging to *S. habrochaites*, such as PI24, are notably difficult to cross with cultivated tomato (Bernacchi & Tanksley, 1997). Therefore, near-isogenic lines that have been generated by introgressing the *eIF4E1-pot1* allele into the elite Mospomorist (*S. lycopersicum*) cultivar were used, hereafter named NIL-pot1 (Gauffier *et al.*, in preparation). F1 plants issued from the cross between NIL-pot1 and LA0411 were all resistant to PVY N605 (n=20 plants) suggesting that the monogenic recessive resistance in LA0411 is allelic to *eIF4E1-pot1*, and that the resistance in the *S. pimpinellifolium* LA0411 is caused by mutations in its *eIF4E1* allele. Consequently, we propose to name this allele as *eIF4E1-pot1*^{2}. To confirm this allelism test, F1 plants generated between NIL-pot1 and LA0411 were self-crossed and F2 plants were tested for their resistance to PVY-N605. Although the F2 plants were mostly resistant to PVY-N605 as expected given the allelism between *eIF4E1-pot1* and *eIF4E1-pot1*^{2} (38 resistant plants out of 40 plants inoculated with PVY-N605), two plants (representing 5%) were found to be fully susceptible to the virus. This low occurrence of susceptibility could be caused by resistance breakdown of the *eIF4E1* resistance allele. Consequently, we further investigated the durability of *eIF4E1-pot1*^{2} mediated resistance to PVY-N605.

<u>The resistance associated with eIF4E1-pot1^2 is easily overcome by PVY N605-derived variants</u>

Partial breakdown of the resistance harbored by the *S. pimpinellifolium* accession LA0411 could have been missed due to the small numbers of plants assayed. Therefore, two more resistance tests were carried out, by assaying at least 50 plants per genotype. All M82 plants were found to be susceptible while all PI24 plants were resistant to the virus. In comparison, LA0411 plants showed a surprising high level of susceptibility ranging from 43 % to 56% of the plants (Table 4).

Resistance breakdown by PVY in pepper and tomato has been associated with non-synonymous substitutions occurring within the viral VPg (Viral protein genome-linked) cistron (Charron *et al.*, 2008; Moury *et al.,* 2004). Sequencing of the RT-PCR-amplified VPg-cistron of the PVY N605 isolate after its propagation in three independent susceptible M82 plants did not reveal any mutation in comparison with the original inoculum. Then, 25 independent LA0411 plants infected by PVY-N605 were sampled and the corresponding VPg cistrons sequenced to identify potential mutations associated with the gain of virulence. All sequences showed polymorphisms associated with AA changes within the VPg coding region. These mutations were a substitution of a leucine by a serine at position 115 in all of the sequenced progenies (L115S) and an additional substitution of an isoleucine by a valine at position 139 (I139V) in 9 of the sequenced progenies. This result confirms that resistance-breaking variants derived from PVY N605 were detected after a first passage on LA0411 plants. Notably, the mutation affecting L115 had previously been characterized as a mutation consistently associated with resistance-breaking in pepper for the related, PVY SON41p isolate (Ayme *et al.,* 2006).

To further confirm that LA0411 susceptibility is caused by resistance-breaking PVY N605 variants that emerged in the progeny, back-inoculations were carried out. Three kinds of viral inoculum were used: a PVY N605 viral progeny propagated on susceptible M82 plants, compared with two PVY N605 variant progenies propagated in LA0411 plants and carrying the L115S and L115S+I139V substitutions, respectively. All inoculated M82 plants were susceptible to the three isolates, whereas all PI24 inoculated plants were resistant (Table 5). The later result shows that the mutations acquired during the first passage on LA0411 did not confer to the PVY N605 variants the ability to overcome *eIF4E1-pot1* mediated-resistance in the broadly resistant line PI24. When the evolved viral isolates present in the LA0411 symptomatic plants after a first passage were back-inoculated onto LA0411 plants, 93% of the plants were infected. Therefore, the second passage allowed the increase in the number of infected plants (ranging from 62% to 93%) confirming that the L115S and L115S+I139V variants were breaking the resistance associated with *eIF4E1-pot1*[2]. The remaining non-infected

plants (accounting for 7% of the inoculated plants) may be due to technical variation in mechanical inoculation.

Altogether, these results show that the *eIF4E1-pot1²* allele is associated with a very low durability to PVY-N605.

The PVY-N605 resistance-breaking isolates use both eIF4E1 and eIF4E2 in tomato

We showed previously that although the natural allele *eIF4E1-pot1* was associated with resistance to most PVY and TEV isolates, a null TILLING allele knocking out (KO) *eIF4E1* was associated with a very narrow resistance spectrum (Piron *et al.*, 2010). We further showed that resistance could be restored, including to PVY-N605, by combining null mutations affecting both *eIF4E1* and *eIF4E2*, uncovering a redundancy effect between *eIF4E1* and *eIF4E2* (Gauffier *et al.*, 2016). To investigate how the Resistance Breaking (RB) PVY-N605 variants are able to overcome the *eIF4E1-pot1²* mediated resistance, we looked at the resistance status of previously characterized tomato TILLING loss-of-function mutants (Gauffier *et al.*, 2016; Piron *et al.*, 2010) towards those RB variants. The reference isolate PVY-N605 and the two RB isolates, respectively harboring L115S and L115+I139V substitutions in the VPg, were propagated on susceptible M82 plants and inoculated on single KO mutant affecting eIF4E1 and eIF4E2 (hereafter Δ4E1 and Δ4E2 plants, respectively) and on plants combining both mutations Δ4E1 Δ4E2 (Table 6). Δ4E1 and Δ4E2 plants were susceptible to all three PVY isolates tested, showing that PVY-N605 isolate and its derived *eIF4E1-pot1²* resistance breaking PVY-N605 L115S and L115+I139V variants, are able to recruit either eIF4E1 or eIF4E2 to infect tomato. In comparison, plants combining mutations in both *eIF4E1* and *eIF4E2,* Δ4E1XΔ4E2, in addition to being resistant to PVY-N605, were also fully resistant to the RB variants PVY-N605 L115S and L115+I139V. As suggested for the pepper/ PVY and pepper/ TEV pathosystems, these results suggest that the mutations in the VPg of the two evolved PVY-N605 isolates allow the virus to recruit the mutated form of eIF4E1-pot1² rather than hijacking new host factors in the plant. Combination of

the eIF4E1-pot1² allele with a null eIF4E2 TILLING allele will be carried out to confirm this hypothesis: the resulting genotype is expected to be susceptible to the RB PVY-N605 isolates.

Discussion

In the present study, by using RNAseq data collected from twenty tomato accessions, including 10 from the wild related species *Solanum pimpinellifolium*, we could assess for the first time the sequence variability among all eukaryotic translation initiation factors 4E, including the isoform eIFiso4E and the atypical related protein nCBP that had not been described yet in tomato. This new set of data is crucial in the light of the comparison with data already put forward in pepper (*Capsicum spp.*) for eIF4E1 (Cavatorta *et al.,* 2008; Charron *et al.,* 2008; Ibiza *et al.*, 2010). Indeed, both pepper and tomato are close relatives among the Solanaceae family and both are host of PVY and TEV potyviruses. In both *Capsicum* and *Solanum* species, recessive resistance to these potyviruses has been characterized as relying on natural variants of eIF4E1. Finally, both plant species encode the same set of eukaryotic translation initiation factors 4E: eIF4E1, a closely related eIF4E2 gene, the eIFiso4E isoform and the nCBP (Gauffier *et al.*, 2016; this study; Gallois and Caranta, unpublished results).

Our results suggest contrasted evolution patterns within the eukaryotic translation initiation factor 4E gene family. As shown previously in pepper, the eIF4E1 gene displayed the largest number of nucleotide polymorphisms associated with AA changes. Overall, the screening of the 20 tomato accessions from the ARCAD data set allowed us to find polymorphism within 7 accessions and identify 4 new haplotypes coding for 4 distinct new eIF4E1 proteins. Those polymorphisms affected five positions within the eIF4E1 coding sequence and were all associated with AA changes. These results are in accordance with the analysis of polymorphism of the Capsicum eIF4E1 (Cavatorta *et al.*, 2008; Charron *et al.,* 2008; Ibiza *et al.,* 2010) and suggest that, as in pepper, positive selection may act on eIF4E1 to fix new alleles. In pepper, this positive selection has been associated with

coevolution with potyviruses (Charron *et al.*, 2008), and, consequently, a similar process could occur in tomato. In contrast with the high level of polymorphism among eIF4E1, no polymorphism was detected in eIF4E2 or in eIFiso4E sequences. The results on eIFiso4E nucleotide variability in tomato differ from the relatively high diversity of polymorphism discovered among *Capsicum* spp. accessions, again associated with AA changes (Ibiza *et al.*, 2010). Interestingly, several pepper accessions harbor a natural knock-out of the *eIFiso4E* gene, due to a deletion occurring within the gene. This loss-of-function, when combined with a *eIF4E1*-mediated *pvr2* resistance allele is associated with resistance to several isolates of *Pepper veinal mottle virus* and *Chilli veinal mottle virus*, as well as an increase in *pvr2* resistance overcoming (Hwang *et al.,* 2009; Quenouille *et al.,* 2016; Ruffel *et al.,* 2006). Therefore, this nucleotide variability in the pepper *eIFiso4E* could be driven by an involvement in susceptibility to potyviruses, and on the contrary, eIFiso4E would not be involved in susceptibility to potyviruses in tomato. Indeed all *eIFiso4E* loss-of-function tomato mutants studied so far remain susceptible to potyvirus infection (Piron *et al.*, 2010). Finally, we report for the first time an analysis of nucleotide diversity within the eIF4E2 coding sequence. No polymorphism could be detected among the eIF4E2 coding sequence for any of the 20 tomato accessions tested. eIF4E2 has been shown to be a susceptibility factor for potyviruses, but only when the *eIF4E1* gene was knocked out, through RNAi or EMS-induced mutation (Gauffier *et al.*, 2016; Mazier *et al.,* 2011). More precisely, comparing natural functional eIF4E1 variants and non-functional KO TILLING mutants in tomato allowed us to show that the presence of a *eIF4E1* natural resistance allele makes the eIF4E2 factor unavailable for the potyvirus. Consequently, the stability of *eIF4E2* coding sequence is fully consistent with an absence of selection pressure from the viruses on *eIF4E2.* Similar results were obtained on the pepper eIF4E2 sequence variability (Charron, 2007), suggesting that a similar mechanism of redundancy between eIF4E1 and eIF4E2 could be found in this related species.

This study allowed us to characterize in *Solanum pimpinellifolium* accession LA0411 one new *eIF4E1* allele whose sequence suggests that it could provide a new resistance allele, due to the presence of

an AA change within region II. Both genetic studies and resistance assays showed that this new allele is indeed a recessive *eIF4E1* resistance allele to PVY, hence named *pot1*^{2}. The eIF4E1 protein encoded by this allele is characterized by two AA changes. One of these, V54K, is located at the N-terminal part of the protein, outside regions I and II where mutations associated with resistance to potyviruses are consistently located in eIF4E factors in different plant species (Robaglia & Caranta, 2006). This mutation is also present on its own in two cultivated accessions from our set (LA1420 and LA0409). Both LA1420 and LA0409 plants were found to be fully susceptible to PVY N605 (10 susceptible plants out of 10) suggesting that this AA change does not play a role in resistance (data not shown). The second mutation, D112G, is located in region II, a short region predicted to locate near the cap-binding pocket in the eIF4E 3D structure and involves a substitution of an acidic residue by a non-polar one. Because it originates from a close wild relative to cultivated tomato, this allele can be easily introgressed into the cultivated tomato. However, compared with the *eIF4E1-pot1* resistance allele, which originates from the PI24 *Solanum habrochaites* accession and confers resistance to PVY and TEV, *eIF4E1-pot1*^{2} is associated with a narrow resistance spectrum as it confers PVY resistance with limited durability and no resistance to TEV. These results suggest also that efficient resistance to potyviruses in *Solanaceae* may require simultaneous AA changes in two distant regions of the eIF4E1 protein (regions I and II): these results must be taken into account for strategies involving new breeding techniques such as CRISPR/Cas9 to engineer resistance alleles that may require two rounds of CRISPR/Cas9 gene modifications. Alternatively, one may consider combining the use of natural variation and genome modification by using the *eIF4E1-pot1*^{2} allele as a template to introduce mutations in region I by CRISPR/Cas9.

The development and evolution of high-throughput sequencing technologies, such as RNAseq, have a tremendous impact on plant breeding by allowing access to the genome-wide genetic diversity, molecular marker development and genome-wide association studies at a reasonable cost (Barabaschi *et al.*, 2016; Goodwin *et al.*, 2016). Our study demonstrates also the straightforward use of RNAseq data to mine for alleles of interest, and can be implemented using new sets of data, on a

larger set of accessions, that were released (Pease *et al.*, 2016). However, the limitations of next generation sequencing data should be kept in mind, including the presence of false negative single nucleotide polymorphisms (*i.e.* the lack of detection of actual polymorphisms), a problem already acknowledged for clinical genetic diagnostic (Huang *et al.,* 2015; Park *et al.,* 2015). Such false negative rates have been assessed as high as 4% by comparing RNAseq data with Sanger sequences in *Brassica rapa* and could account for missing potentially interesting polymorphisms (Devisetty *et al.,* 2014). Therefore, the lack of detected polymorphism does not necessarily reflect a wild-type allele. For example, Sanger sequencing confirmed that the Cervil cultivated tomato accession, included in our panel, contains an eIF4E1 variant (D112N) that was not detected in our RNAseq data screening.

Overall, the use of RNAseq data potentially allows the simultaneous analysis of all members from the same gene family. That is of particular interest for plant translation initiation factors 4E, given the large redundancy among the genes encoding those factors in plant genome, as well as the specific recruitment of the different eIF4E isoforms by various species of potyviruses as exemplified in *Arabidopsis thaliana* (Duprat *et al.,* 2002; Nicaise *et al.,* 2007; Sato *et al.,* 2005). Furthermore, such studies could be extended by investigating a large number of susceptibility candidate genes, such as the *chloroplastic phosphoglycerate kinase 2* (cPGK2) that has been shown to be associated with resistance to the potyvirus *Watermelon mosaic virus* in *Arabidopsis thaliana*, or the *protein disulfide isomerase Like 5-1 (*HvPDIL5-1), a gene involved with resistance to bymoviruses in barley (Ouibrahim *et al.,* 2014; Yang *et al.,* 2014). Because of their limited genome, viruses recruit host factors to complete all the steps of their infectious cycle and are therefore engaged in many interactions with host proteins. Several studies have allowed the characterization of plant proteins interacting with viral proteins, mainly using yeast two hybrid protein-protein interaction systems (for review see Elena & Rodrigo 2012) or affinity purification (Martinez *et al.,* 2016). Those host genes are potentially encoding susceptibility factors, and variants of those genes could possibly encode new resistance factors, mirroring the example of eIF4E proteins in potyvirus resistance. This reverse genetics

approach could allow extending the portfolio of resistance genes based on loss-of-susceptibility and could define potentially resistant accessions that could then be used as sources for transfer to crops.

Methods.

Plant materials

M82 (*S. lycopersicum*) was used as the susceptible control. The resistant accession *S. habrochaites* PI247087 (PI24) has been described previously in Parrella *et al.* (2002). LA0411 accession seeds were provided by INRA CRB (Centre de Ressources Biologiques, http://www6.paca.inra.fr/gafl). The TILLING lines KO for *eIF4E1* (ΔE1) and *eIF4E2* (ΔE2) were previously obtained in a *S. lycopersicum* M82 background (Piron *et al.,* 2010) and the double mutant (ΔE1, ΔE2) has been previously described (Gauffier *et al.,* 2016).

eIF4E sequence analysis

The RNAseq data were obtained in the framework of the ARCAD project aiming at deciphering the molecular footprints of domestication through a comparative genomic approach. Towards this objective, the experimental design relied on 10 crop and 10 wild accessions from 2 species (*S. lycopersicum* and *S. pimpinellifolium*) chosen on the basis of their molecular diversity. A standardized protocol was followed for each of these species in terms of RNA extraction, RNA sequencing and bioinformatic analyses as described in Nabholz *et al.* (2014). Among these species, crop tomato accessions and wild accessions were grown in the greenhouse under classical conditions. Briefly RNA was extracted from young leaves, fruits and flowers that were pooled in a 65/20/15 proportion following their quantification on an Agilent Bioanalyser. Once quantified, RNAseq libraries were prepared using TruSeq RNA sample Preparation v2 kit (Illumina Inc., USA) according to the manufacturer's protocol and individually tagged (6bp). These paired-end libraries (2x101bp) were sequenced and the raw reads were checked and cleaned using FastQC (0.10.2,

http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Cutadapt (v1.8.5, Martin, 2011, with the options –q 20 -e 0.1 -O 7 –m 20) tools. Then, these reads were mapped onto the reference sets of ORF corresponding to the v2.4 of the tomato reference genome annotation (n=34,727 ORFs), using BWA (v0.7.3, Li and Durbin, 2009, using the command line 'aln -e -1 -O 11 -n 2 -m 2000000 -d 10 -E 4 -l 32 -M 3 -k 2 -q 0 -R 30 -t 4 -i 5 -o 1') by allowing no more than two mismatches per read, no multiple hits and the local realignment option. Finally GATK (McKenna *et al.,* 2010) called simultaneously the SNP. The VCFtools (v0.1.10, Danacek *et al.,* 2011) filtered the output variant calling file (vcf) to retain sites showing a minimal coverage of 8x per individual and a MAF of 0.05. Raw reads are hosted online on the Southgreen platform (http://arcad-bioinformatics.southgreen.fr).

For sequence confirmation, Total RNA was isolated from tomato leaves using TRI-reagent (Sigma-Aldrich, Saint Louis, MO, USA). *eIF4E1* cDNAs from LA0411 were obtained by RT-PCR with primer pairs F 5'-ATGGCAGCAGCTGAAATGGAGAGA-3'/R 5'-CTATACGGTGTAACGATTCTTGGC-3'. To compare polymorphisms between the accessions, amplification products were sequenced by Genoscreen (Lille, France) and the corresponding protein sequences were aligned using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo; website) and BoxShade (http://www.ch.embnet.org/software/BOX_form.html; website).

Virus isolates and infection assays

The PVY isolates LYE90v (Moury *et al.*, 2004), N605 (Parrella *et al.*, 2002) and SON41g (Charron *et al.,* 2008) and the TEV isolates CAA10 (Charron *et al.*, 2008), S103 (Ruffel *et al.,* 2005) and HAT (Schaad *et al.,* 2000) were propagated on *Nicotiana tabacum* cv Xanthi before inoculating 14-day-old tomato plants. The accumulation of PVY and TEV viruses in non-inoculated upper leaves was then assayed 21 days post inoculation by DAS-ELISA using respectively anti-PVY (Sediag, France) and anti-TEV (Sediag, France) antisera and detection sets. Non-inoculated plants were used as healthy controls. Mean and standard errors of absorbance values at 405nm of samples from at least 6 independent plants per

parental genotype and 12 independent plants per progeny were calculated. The susceptibility threshold was set as three times the mean value of healthy controls.

Additional files

Figure S1. Protein sequence alignment of nCBP encoded by *Arabidopsis thaliana* and *Solanum lycopersicum* M82 accession

Figure S2. LA0411 is associated with a narrow resistance spectrum to PVY.

Table S1. List of tomato accessions sequenced by RNAseq in the ARCAD project

Acknowledgements

The RNAseq data production was initially supported by Agropolis Foundation under the ARCAD project No 0900-001 (http://www.arcad-project.org/). We thank Renaud Duboscq for additionnal RNAseq data analysis as well as André Moretti and Carole Caranta for providing the NIL-pot1 plants. We thank the INRA Avignon Centre de Ressources Biologiques (CRB) for providing the LA0411 seeds.

References

Andersen, M.M., Landes, X., Xiang, W., Anyshchenko, A., Falhof, J., Østerberg, J.T., Olsen, L.I., Edenbrandt, A.K., Vedel, S.E., Thorsen, B.J., Sandøe, P., Gamborg, C., Kappel, K. and Palmgren, M.G. (2015) Feasibility of new breeding techniques for organic farming. *Trends in Plant Science*, **20**, 426-434.

Ayme, V., Souche, S., Caranta, C., Jacquemond, M., Chadoeuf, J., Palloix, A. and Moury, B. (2006) Different mutations in the genome-linked protein VPg of potato virus Y confer virulence on the pvr2(3) resistance in pepper. *Molecular plant-microbe interactions: MPMI*, **19**, 557-563.

Barabaschi, D., Tondelli, A., Desiderio, F., Volante, A., Vaccino, P., Vale, G. and Cattivelli, L. (2016) Next generation breeding. *Plant Sci*, **242**, 3-13.

Bernacchi, D. and Tanksley, S.D. (1997) An interspecific backcross of Lycopersicon esculentum × L. hirsutum: linkage analysis and a QTL study of sexual compatibility factors and floral traits. *Genetics*, **147**, 861-877.

Browning, K.S. and Bailey-Serres, J. (2015) Mechanism of cytoplasmic mRNA translation. *The Arabidopsis Book*, e0176.

Cavatorta, J.R., Savage, A.E., Yeam, I., Gray, S.M. and Jahn, M.M. (2008) Positive Darwinian selection at single amino acid sites conferring plant virus resistance. *J Mol Evol*, **67**, 551-559.

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., Sherman, A., Arazi, T. and Gal-On, A. (2016) Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Mol Plant Pathol*. , 17: 1140–1153.

Charron, C. (2007) *Caractérisation fonctionnelle et évolution moléculaire des gènes codant pour les facteurs d'initiation de la traduction eIF4E : des facteurs clés dans la résistance des plantes aux potyvirus*. PhD thesis Aix Marseille 2 University.

Charron, C., Nicolaï, M., Gallois, J.-L., Robaglia, C., Moury, B., Palloix, A. and Caranta, C. (2008) Natural variation and functional analyses provide evidence for co-evolution between plant eIF4E and potyviral VPg. *The Plant Journal*, **54**, 56–68.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R. and and 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156-2158.

Devisetty, U.K., Covington, M.F., Tat, A.V., Lekkala, S. and Maloof, J.N. (2014) Polymorphism identification and improved genome annotation of Brassica rapa through Deep RNA sequencing. *G3 (Bethesda)*, **4**, 2065-2078.

Duprat, A., Caranta, C., Revers, F., Menand, B., Browning, K.S. and Robaglia, C. (2002) The Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. *Plant Journal*, **32**, 927-934.

Elena, S.F. and Rodrigo, G. (2012) Towards an integrated molecular model of plant-virus interactions. *Curr Opin Virol*, **2**, 719-724.

Fraser, R.S.S. (1990) The genetics of resistance to plant viruses. *Annual Review of Phytopathology*, **28**, 179-200.

Gao, Z., Johansen, E., Eyers, S., Thomas, C.L., Noel Ellis, T.H. and Maule, A.J. (2004) The potyvirus recessive resistance gene, sbm1, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. *Plant J.* **40**, 376-385.

Gauffier, C., Lebaron, C., Moretti, A., Constant, C., Moquet, F., Bonnet, G., Caranta, C. and Gallois, J.L. (2016) A TILLING approach to generate broad-spectrum resistance to potyviruses in tomato is hampered by eIF4E gene redundancy. *Plant J*. 85: 717–729.

Goodwin, S., Mcpherson, J.D. & McCombie, W.R. (2016) Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* **17**, 333–351 (2016).

Hofinger, B.J., Russell, J.R., Bass, C.G., Baldwin, T., dos Reis, M., Hedley, P.E., Li, Y., Macaulay, M., Waugh, R., Hammond-Kosack, K.E. and Kanyuka, K. (2011) An exceptionally high nucleotide and haplotype diversity and a signature of positive selection for the eIF4E resistance gene in barley are revealed by allele mining and phylogenetic analyses of natural populations. *Mol Ecol*, **20**, 3653-3668.

Huang, X.F., Wu, J., Lv, J.N., Zhang, X. and Jin, Z.B. (2015) Identification of false-negative mutations missed by next-generation sequencing in retinitis pigmentosa patients: a complementary approach to clinical genetic diagnostic testing. *Genet Med*, **17**, 307-311.

Hwang, J., Li, J., Liu, W.-Y., An, S.-J., Cho, H., Her, N.H., Yeam, I., Kim, D. and Kang, B.-C. (2009) Double mutations in eIF4E and eIFiso4E confer recessive resistance to Chilli veinal mottle virus in pepper. *Molecules and Cells*, **27**, 329-336.

Ibiza, V.P., Cañizares, J. and Nuez, F. (2010) EcoTILLING in Capsicum species: searching for new virus resistances. *BMC Genomics*, **11**, 631.

Jeong, H.-J., Kwon, J.-K., Pandeya, D., Hwang, J., Hoang, N.H., Bae, J.-H. and Kang, B.-C. (2011) A survey of natural and ethyl methane sulfonate-induced variations of eIF4E using high-resolution melting analysis in Capsicum. *Molecular Breeding*, **29**, 349-360.

Koenig, D., Jimenez-Gomez, J.M., Kimura, S., Fulop, D., Chitwood, D.H., Headland, L.R., Kumar, R., Covington, M.F., Devisetty, U.K., Tat, A.V., Tohge, T., Bolger, A., Schneeberger, K., Ossowski, S., Lanz, C., Xiong, G., Taylor-Teeples, M., Brady, S.M., Pauly, M., Weigel, D., Usadel, B., Fernie, A.R., Peng, J., Sinha, N.R. and Maloof, J.N. (2013) Comparative

transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc Natl Acad Sci U S A*, **110**, E2655-2662.

Konecna, E., Safarova, D., Navratil, M., Hanacek, P., Coyne, C., Flavell, A., Vishnyakova, M., Ambrose, M., Redden, R. and Smykal, P. (2014) Geographical gradient of the eIF4E alleles conferring resistance to potyviruses in pea (Pisum) germplasm. *PLoS One*, **9**, e90394.

Legnani, R., Gognalons, P., Moretti, A., Marchoux, G., Selassie, K.G. and Laterrot, H. (1996) Identification and characterization of resistance to *tobacco etch virus* in *Lycopersicon* species. *Plant Disease*, **80**, 306.

Legnani, R., Selassie, K.G., Womdim, R.N., Gognalons, P., Moretti, A., Laterrot, H. and Marchoux, G. (1995) Evaluation and inheritance of the Lycopersicon hirsutum resistance against potato virus Y. *Euphytica*, **86**, 219-226.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and DePristo, M.A. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*, **20**, 1297-1303.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10-12.

Martinez, F., Rodrigo, G., Aragones, V., Ruiz, M., Lodewijk, I., Fernandez, U., Elena, S.F. and Daros, J.A. (2016) Interaction network of tobacco etch potyvirus NIa protein with the host proteome during infection. *BMC Genomics*, **17**, 87.

Mazier, M., Flamain, F., Nicolaï, M., Sarnette, V. and Caranta, C. (2011) Knock-down of both eIF4E1 and eIF4E2 genes confers broad-spectrum resistance against potyviruses in tomato. *PLoS ONE*, **6**, e29595.

Moury, B., Morel, C., Johansen, E., Guilbaud, L., Souche, S., Ayme, V., Caranta, C., Palloix, A. and Jacquemond, M. (2004) Mutations in Potato virus Y genome-linked protein determine virulence toward recessive resistances in Capsicum annuum and Lycopersicon hirsutum. *Molecular Plant-Microbe Interactions*, **17**, 322–329.

Nabholz, B., Sarah, G., Sabot, F., Ruiz, M., Adam, H., Nidelet, S., Ghesquiere, A., Santoni, S., David, J. and Glemin, S. (2014) Transcriptome population genomics reveals severe bottleneck and domestication cost in the African rice (Oryza glaberrima). *Mol Ecol*, **23**, 2210-2227.

Nicaise, V., German-Retana, S., Sanjuan, R., Dubrana, M.P., Mazier, M., Maisonneuve, B., Candresse, T., Caranta, C. and LeGall, O. (2003) The eukaryotic translation initiation factor 4E controls lettuce susceptibility to the Potyvirus Lettuce mosaic virus. *Plant Physiol*, **132**, 1272-1282.

Nicaise, V., Gallois, J.-L., Chafiai, F., Allen, L.M., Schurdi-Levraud, V., Browning, K.S., Candresse, T., Caranta, C., Le Gall, O. and German-Retana, S. (2007) Coordinated and selective recruitment of eIF4E and eIF4G factors for potyvirus infection in Arabidopsis thaliana. *FEBS Letters*, **581**, 1041-1046.

Nieto, C., Piron, F., Dalmais, M., Marco, C.F., Moriones, E., Gomez-Guillamon, M.L., Truniger, V., Gomez, P., Garcia-Mas, J., Aranda, M.A. and Bendahmane, A. (2007) EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility. *BMC Plant Biol*, **7**, 34.

Ouibrahim, L., Mazier, M., Estevan, J., Pagny, G., Decroocq, V., Desbiez, C., Moretti, A., Gallois, J.-L., Caranta, C. (2014) Cloning of the Arabidopsis *rwm1* gene for resistance to *Watermelon mosaic virus* points to a new function for natural virus resistance genes. Plant J. 79, 705–716.

Park, J.Y., Clark, P., Londin, E., Sponziello, M., Kricka, L.J. and Fortina, P. (2015) Clinical exome performance for reporting secondary genetic findings. *Clin Chem*, **61**, 213-220.

Parrella, G., Ruffel, S., Moretti, A., Morel, C., Palloix, A. and Caranta, C. (2002) Recessive resistance genes against potyviruses are localized in colinear genomic regions of the tomato (Lycopersicon spp.) and pepper (Capsicum spp.) genomes. *TAG Theoretical and Applied Genetics*, **105**, 855-861.

Pavan, S., Jacobsen, E., Visser, R.G.F. and Bai, Y. (2010) Loss of susceptibility as a novel breeding strategy for durable and broad-spectrum resistance. *Molecular Breeding*, **25**, 1-12.

Pease, J.B., Haak, D.C., Hahn, M.W. and Moyle, L.C. (2016) Phylogenomics reveals threesources of adaptive variation during a rapid radiation. *PLoS Biol*, **14**, e1002379.

Piron, F., Nicolaï, M., Minoïa, S., Piednoir, E., Moretti, A., Salgues, A., Zamir, D., Caranta, C. and Bendahmane, A. (2010) An induced mutation in tomato eIF4E leads to immunity to two potyviruses. *PLoS ONE*, **5**, e11313.

Pyott, D.E., Sheehan, E. and Molnar, A. (2016) Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants. *Mol Plant Pathol*. , 17: 1276–1288.

Quenouille, J., Saint-Felix, L., Moury, B. and Palloix, A. (2016) Diversity of genetic backgrounds modulating the durability of a major resistance gene. Analysis of a core collection of pepper landraces resistant to Potato virus Y. *Mol Plant Pathol*, **17**, 296-302.

Robaglia, C. and Caranta, C. (2006) Translation initiation factors: a weak link in plant RNA virus infection. *Trends in Plant Science*, **11**, 40-45.

Ruffel, S., Dussault, M.-H., Palloix, A., Moury, B., Bendahmane, A., Robaglia, C. and Caranta, C. (2002) A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). *The Plant Journal*, **32**, 1067–1075.

Ruffel, S., Gallois, J., Moury, B., Robaglia, C., Palloix, A. and Caranta, C. (2006) Simultaneous mutations in translation initiation factors eIF4E and eIF(iso)4E are required to prevent pepper veinal mottle virus infection of pepper. *Journal of General Virology*, **87**, 2089-2098.

Ruffel, S., Gallois, J.L., Lesage, M.L. and Caranta, C. (2005) The recessive potyvirus resistance gene pot-1 is the tomato orthologue of the pepper pvr2-eIF4E gene. *Mol Genet Genomics*, **274**, 346-353.

Ruud, K.A., Kuhlow, C., Goss, D.J. and Browning, K.S. (1998) Identification and Characterization of a Novel Cap-binding Protein from Arabidopsis thaliana. *Journal of Biological Chemistry*, **273**, 10325-10330.

Sarah, G., Homa, F., Pointet, S., Contreras, S., Sabot, F., Nabholz, B., Santoni, S., Sauné, L., Ardisson, M., Chantret, N., Sauvage, C., Tregear, J., Jourda, C., Pot, D., Vigouroux, Y., Chair, H., Scarcelli, N., Billot, C., Yahiaoui, N., Bacilieri, R., Khadari, B., Boccara, M., Barnaud, A., Péros, J.-P., Labouisse, J.-P., Pham, J.-L., David, J., Glémin, S. and Ruiz, M. (2016) A large set of 26 new reference transcriptomes dedicated to comparative population genomics in crops and wild relatives. *Molecular Ecology Resources*, DOI 10.1111/1755-0998.12587

Sato, M., Nakahara, K., Yoshii, M., Ishikawa, M. and Uyeda, I. (2005) Selective involvement of members of the eukaryotic initiation factor 4E family in the infection of Arabidopsis thaliana by potyviruses. *FEBS Letters*, **579**, 1167-1171.

Schaad, M.C., Anderberg, R.J. and Carrington, J.C. (2000) Strain-specific interaction of the tobacco etch virus NIa protein with the translation initiation factor eIF4E in the yeast two-hybrid system. *Virology*, **273**, 300-306.

The Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485, 635-641.

van Schie, C.C. and Takken, F.L. (2014) Susceptibility genes 101: how to be a good host. *Annu Rev Phytopathol*, **52**, 551-581.

Wang, A. and Krishnaswamy, S. (2012) Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement: eIF4E-mediated resistance to plant viruses. *Molecular Plant Pathology*, **13**, 795-803.

Yang, P., Lupken, T., Habekuss, A., Hensel, G., Steuernagel, B., Kilian, B., Ariyadasa, R., Himmelbach, A., Kumlehn, J., Scholz, U., Ordon, F. and Stein, N. (2014) PROTEIN DISULFIDE ISOMERASE LIKE 5-1 is a susceptibility factor to plant viruses. *Proc Natl Acad Sci U S A*, **111**, 2104-2109.

Yeam, I., Cavatorta, J.R., Ripoll, D.R., Kang, B.C. and Jahn, M.M. (2007) Functional dissection of naturally occurring amino acid substitutions in eIF4E that confers recessive potyvirus resistance in plants. *Plant Cell*, **19**, 2913-2928.

Figure legend:

Figure 1. New eIF4E1 haplotypes characterized in this study. (a) Sequence alignment of eIF4E1 proteins encoded by *S. lycopersicum* M82 and *S. habrochaites* PI24 and PI13 alleles (Ruffel *et al*., 2005) as well as those encoded by the four new haplotypes characterized in this study. Regions I and II (defined by Robaglia & Caranta 2006) are represented in red and green, respectively. (b) Table of the accession where the four new haplotypes have been characterized.

Tables:

Table 1. Haplotypes of the tomato translation initiation factors 4E found in the RNAseq ARCAD dataset.

Gene	Identifier	Chromosome	CDS length	Haplotypes	New haplotypes with AA changes
eIF4E1	Solyc03g005870	3	696	5	4
eIF4E2	Solyc02g021550	2	663	1	0
eIFiso4E	Solyc09g090580	9	603	1	0
nCBP	Solyc10g080660	10	672	3	0

CDS: Coding DNA Sequence. AA: Amino Acid. The haplotypes references correspond to the sequences from *Solanum lycopersicum* M82 accession .

Table 2. LA0411 is associated with a narrow resistance spectrum to PVY

		PVY			TEV		
species	accession	LYE90	N605	SON41	CAA10	S103	HAT
S. lycopersicum	M82	S	S	S	S	S	S
S. pimpinellifolium	LA0411	R	R	S	S	S	S
S. habrochaites	PI24	R	R	R	R	R	R

Resistance (R) or Susceptibility (S) was deduced following DAS-ELISA on at least 8 inoculated plants per genotype and per viral strain.

Table 3. Inheritance of resistance to PVY N605 in *Solanum pimpinellifolium* LA0411

Genotype	Phenotype		Expected ratio	CHI^2	P
	S	R			
S. lycopersicum M82	7	0	1:0	-	-
S. pimpinellifolium LA0411	0	7	0:1	-	-
(M82 X LA0411)F1	20	0	1:0	-	-
BC1S = (M82 X LA0411)F1 X M82	20	0	1:0	-	-
BC1R = (M82 X LA0411)F1 X LA0411	20	20	1:1	-	-
(M82 X LA0411)F2	26	14	3:1	2.13	0.14

Table 4. LA0411 shows lower durability to PVY N605 than PI24

	Accession					
	M82		PI24		LA0411	
	R	n	R	n	R	n
Assay 1	0%	57	100%	57	43%	58
Assay 2	0%	51	100%	60	56%	59

Results from DAS-ELISA from two independent assays. R represents the percentage of resistant plants. n = number of plants assayed. ELISA were performed at 21 dpi.

Table 5. Back inoculation of PVY-N605 evolved isolates on LA0411 plants is associated with a reduction of resistance

	Accession					
	M82		PI24		LA0411	
	R	n	R	n	R	n
PVY N605	0%	10	100%	9	38%	50
L115S	0%	9	100%	10	7%	55
L115S + I139V	0%	10	100%	9	7%	57

Results from DAS-ELISA. R represents the percentage of resistant plants. n = number of plants assayed.

Table 6. Resistance of tomato 4E TLLING mutants to PVY N605 and two RB variants

	M82 Accession		
	Δ4E1	Δ4E2	Δ4E1 Δ4E2
PVY N605	0 R / 8	0 R / 8	8 R / 8
L115S	1 R / 7	0 R / 7	7 R / 7
L115S + I139V	0 R / 8	0 R / 7	8 R / 8

Resistance (R) to PVY was characterized by DAS-ELISA. For each category, the number of resistant plants out of the total number of plant assayed is given.

Figure 1

Click here to download Figure Figure 1 LEBARON without legenda.docx

(a)

```
M82           1 MAAAEMERTMSFDAAEKLKAADGGGGEVDDELEEGEIVEESNDTASYLGKEITVKHPLEH
haplotype1    1 .....................................................K......
haplotype2    1 .........................................................V..
haplotype3    1 ............................................................
haplotype4    1 .....................................................K......
PI13          1 ............................................................
PI24          1 ...............................................F............

M82          61 SWTFWFDNPTTKSRQTAWGSSLRNVYTFSTVEDFWGAYNNIHHPSKLIMGADFHCFKHKI
haplotype1   61 ............................................................
haplotype2   61 ............................................................
haplotype3   61 ............................................................
haplotype4   61 ...................................................G........
PI13         61 ........S...............L...................................
PI24         61 .......KS.......D.......L.......................I...........

M82         121 EPKWEDPVCANGGTWKMSFSKGKSDTSWLYTLLAMIGHQFDHGDEICGAVVSVRAKGEKI
haplotype1  121 ............................................................
haplotype2  121 ............................................................
haplotype3  121 ..............................M.............................
haplotype4  121 ............................................................
PI13        121 ..Q.........................................................
PI24        121 ..Q.........................................................

M82         181 ALWTKNAANETAQVSIGKQWKQFLDYSDSVGFIFHDDAKRLDRNAKNRYTV
haplotype1  181 ...................................................
haplotype2  181 ...................................................
haplotype3  181 ...................................................
haplotype4  181 ...................................................
PI13        181 ...........................................S.......
PI24        181 ...........................................S.......
```

(b)

new haplotype	mutation	cultivated accessions	*pimpinellifolium* accessions
haplotype 1	V54K	LA1420, LA0409	
haplotype 2	L58V		LA1478, LA1582, LA1593
haplotype 3	T151M		LA1245
haplotype 4	V54K, D112G		LA0411

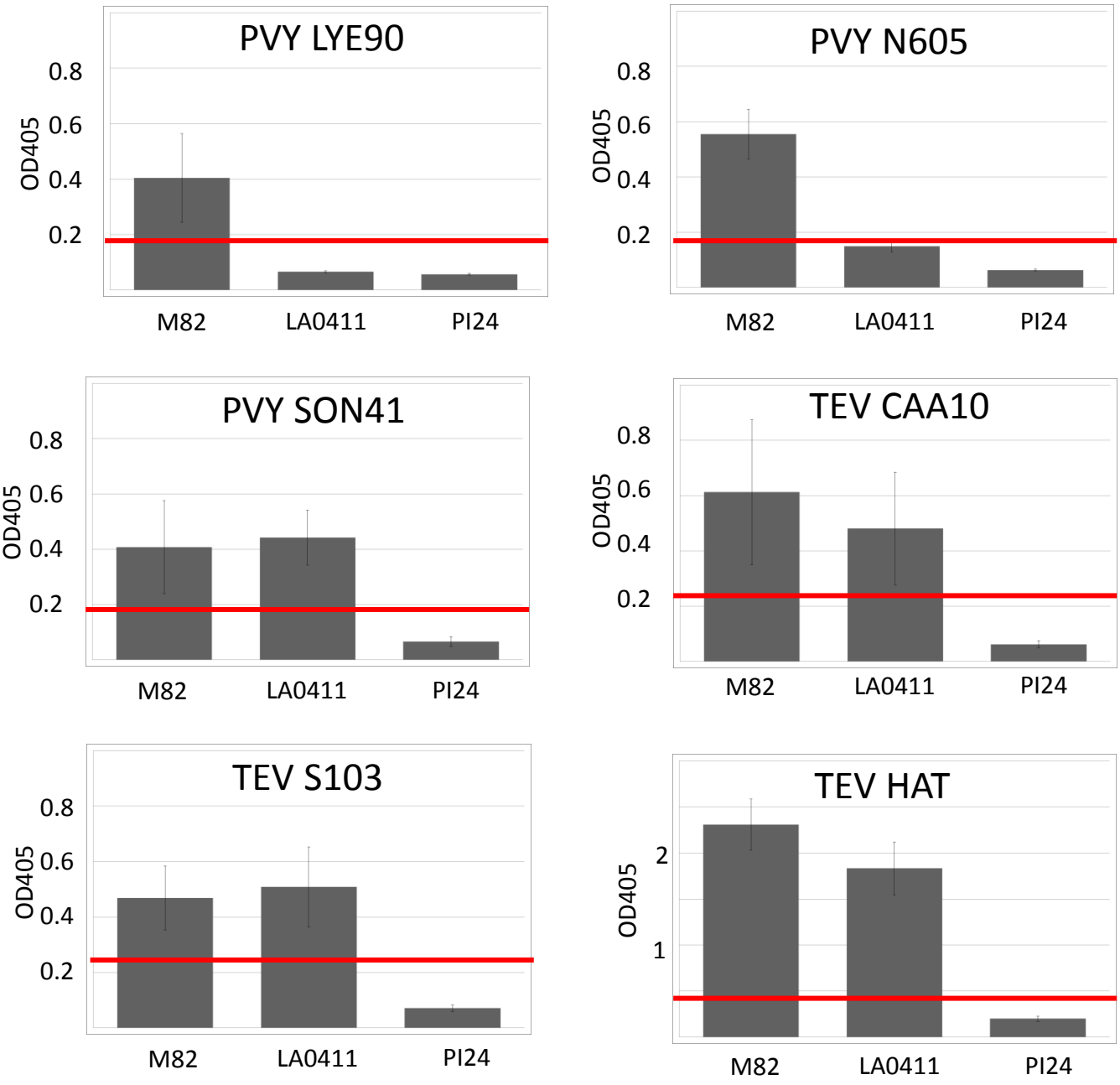
```
                                                                    *1  *2
AtnCBP        1 MEVLDRRDDEIRDSGNMDSIKSHYVTDSVSEERRSR...ELKDGDHPLRYKFSIWYTRRT
LsnCBP        1 ...TSEKKE..L.NK.TN.TQ.LIIDP.IAA.D.E.IAVD..A.L...KN..VF......

                        *3                  *4                          *5
AtnCBP       58 PGVRNQ.SYEDNIKKMVEFSTVEGFWACYCHLARSSLLPSPTDLHFFKDGIRPLWEDGAN
LsnCBP       59 ....T.T........I.D........V.......P.A........L.RE........A..

                     *6            *7                                 *8
AtnCBP      117 CNGGKWIIRFSKVVSARFWEDLLLALVGDQLDDADNICGAVLSVRFNEDIISVWNRNASD
LsnCBP      119 .H........K.A..G......V.........YG.........I......L.........


AtnCBP      177 HQAVMGLRDSIKRHLKLPHAYVMEYKPHDASLRDNSSYRNTWLRG
LsnCBP      179 Q....A............GG.I....A..............I.S
```

Supplementary Figure 1. Protein sequence alignement of nCBP encoded by *Arabidopsis thaliana* and *Solanum lycopersicum* M82 accession. The respective gene references are At5g18110 (*Arabidopsis thaliana*) and Solyc10g080660 (*Solanum lycopersicum*). Localisation of the 8 Trp residues usually conserved among eIF4E protein are numbered and noted with stars, including the italicized residues 1 and 3 that are substituted in nCBP.



Supplementary Figure 2. LA0411 is associated with a narrow resistance spectrum to PVY. Tomato plants were assayed for viral coat protein accumulation by ELISA at 21 days post-inoculation. At least 8 independent plants were assayed for each condition. The horizontal red lines on each graph represent three times the mean value of healthy control plants.

Table S1. List of tomato accessions sequenced by RNAseq in the ARCAD project.

ARCAD reference	Species	Genotype
SC1	*S. lycopersicum*	Levovil
SC2	*S. lycopersicum*	Stupicke Polni Rane
SC3	*S. lycopersicum* var *cerasiforme*	Plovdiv 24A
SC4	*S. lycopersicum* var *cerasiforme*	LA1420
SC5	*S. lycopersicum* var *cerasiforme*	Criollo
SC6	*S. lycopersicum*	LA0147
SC7	*S. lycopersicum* var *cerasiforme*	Cervil
SC8	*S. lycopersicum*	FERUM
SC9	*S. lycopersicum*	LA0767
SC10	*S. lycopersicum*	LA0409
SS1	*S. pimpinellifolium*	LA1589
SS2	*S. pimpinellifolium*	LA1478
SS3	*S. pimpinellifolium*	LA1582
SS4	*S. pimpinellifolium*	LA1593
SS5	*S. pimpinellifolium*	LA1602
SS6	*S. pimpinellifolium*	LA1729
SS7	*S. pimpinellifolium*	L.pimpi.site10(F300044)
SS8	*S. pimpinellifolium*	732292
SS9	*S. pimpinellifolium*	LA0411
SS10	*S. pimpinellifolium*	LA1245