Published in final edited form as:

Theor Appl Genet. 2016 October ; 129(10): 1951–1959. doi:10.1007/s00122-016-2751-x.

Pentatricopeptide repeat 336 as the candidate gene for paternal sorting of mitochondria (*Psm*) in cucumber

A. R. Del Valle-Echevarria¹, W. Sanseverino², J. Garcia-Mas², and M. J. Havey³

M. J. Havey: mjhavey@wisc.edu

¹Department of Horticulture, University of Wisconsin-Madison, 1575 Linden Drive, Madison, USA

²IRTA, Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB, 08193 Barcelona, Spain

³USDA-ARS and Department of Horticulture, University of Wisconsin-Madison, 1575 Linden Drive, Madison, USA

Key message

Pentatricopeptide repeat (PPR) 336 was identified as the candidate gene for Paternal Sorting of Mitochondria (*Psm*), a nuclear locus that affects the predominant mitochondria transmitted to progenies.

Cucumber (*Cucumis sativus* L.) is a useful plant to study organellar-nuclear interactions because its organelles show differential transmission, maternal for chloroplasts and paternal for mitochondria. The mitochondrial DNA (mtDNA) of cucumber is relatively large due in part to accumulation of repetitive DNAs and recombination among these repetitive regions produces structurally polymorphic mtDNAs associated with paternally transmitted mosaic (MSC) phenotypes. The mitochondrial mutant MSC16 possesses an under-representation of ribosomal protein *S7* (*rps7*), a key component of the small ribosomal subunit in the mitochondrion. A nuclear locus, Paternal Sorting of Mitochondria (*Psm*), affects the predominant mitochondria transmitted to progenies generated from crosses with MSC16 as the male parent. Using single nucleotide polymorphisms, *Psm* was mapped to a 170 kb region on chromosome 3 of cucumber and pentatricopeptide repeat (PPR) 336 was identified as the likely candidate gene. PPR336 stabilizes mitochondrial ribosomes in *Arabidopsis thaliana* and because MSC16 shows reduced transcription of *rps7*, the cucumber homolog of PPR336 (CsPPR336) as the candidate for *Psm* is consistent with a nuclear effect on ribosome assembly or stability in the mitochondrion. We used polymorphisms in CsPPR336 to genotype progenies segregating at *Psm* and recovered only one *Psm*^{−/−} plant with the MSC phenotype, indicating that the combination of the *Psm*[−] allele with mitochondria from MSC16 is almost always lethal. This research illustrates the usefulness of the

Correspondence to: M. J. Havey, mjhavey@wisc.edu.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-016-2751-x) contains supplementary material, which is available to authorized users.

Author contribution statement ADVE: Helped to plan research and conducted all experiments, completed analyses, and manuscript writing; WS: Planned and completed bioinformatic analyses; JGM: Aided with allele-specific primer design and high-throughput genotyping; MJH: Planned and directed research, helped with analyses and writing of manuscript.

MSC mutants of cucumber to reveal and study unique interactions between the mitochondrion and nucleus.

Introduction

The mitochondrial DNA (mtDNA) of plants show great size differences, varying from relatively small at 208 kb in *Brassica hirta* to very large at 11,300 kb in *Silene conica* (Palmer and Herbon 1987; Sloan et al. 2012). Larger plant mtDNAs do not carry significantly more coding regions than smaller ones (Alverson et al. 2010), and possess 20–30 coding regions for sub-units of the complexes that form the electron-transport chain and genes coding for ribosomal proteins and transfer and ribosomal RNAs (Kitazaki and Kubo 2010). Repetitive regions in plant mtDNAs can recombine to produce structurally diverse molecules that can exist at low frequencies (sublimons) relative to the predominant molecule (Alverson et al. 2010, 2011; Kitazaki and Kubo 2010). The prevalence of specific mitochondrial sublimons can shift to produce unique phenotypes such as chlorotic sectoring in *Arabidopsis* (Mackenzie et al. 1994) or the non-chromosomal stripe of maize (Feiler and Newton 1987).

Plants in the family Cucurbitaceae have a wide range of mtDNA sizes (Ward et al. 1981), for example watermelon (*Citrullus lanatus* L.) at 379 kilobases (kb) and melon (*Cucumis melo* L.) at approximately 2740 kb (Alverson et al. 2010; Rodriguez-Moreno et al. 2011). Alverson et al. (2011) sequenced and assembled the 1685 kb mtDNA of the cucumber (*Cucumis sativus* L.) cultivar ‘Calypso’ and showed that it primarily exists as three chromosomes of approximately 1556, 84, and 45 kb. Pockets of repetitive DNAs present in the cucumber mtDNA allow for recombination to produce different conformations even among individual plants of the same cultivar or population (Lilly and Havey 2001; Lilly et al. 2001; Alverson et al. 2011). In addition to relatively large mtDNA sizes, cucumber and melon show paternal transmission of their mtDNAs (Havey 1997; Havey et al. 1998). This unique transmission mode allows for the separation of organellar effects by simple reciprocal crossing (Shen et al. 2015).

The mosaic (MSC) mutants of cucumber have been isolated after passage of the highly inbred (>S₁₈) wild-type line B through cell cultures and regeneration of plants (Malepszy et al. 1996; Bartoszewski et al. 2004, 2007). These MSC mutants possess misshapen cotyledons and chlorotic leaves, and show lower seed germination, slower growth, and poor fertility relative to wild-type plants (Malepszy et al. 1996). When MSC plants are crossed as the male parent with wild-type female plants, progenies possess the MSC phenotype. Reciprocal crosses produce only wild-type progenies, demonstrating paternal transmission of the MSC phenotypes (Malepszy et al. 1996). Independent MSC lines have been isolated from different cell cultures, show paternal transmission of the MSC phenotypes, have distinctive MSC phenotypes on leaves and fruit, and possess unique rearrangements in their mtDNAs (Lilly et al. 2001; Łady y ski et al. 2002; Bartoszewski et al. 2004, 2007; Del Valle-Echevarria et al. 2015).

MSC16 was originally identified among plants regenerated from leaf-callus culture established using inbred B (Malepszy et al. 1996). Lilly et al. (2001) identified a deletion

(JLV5-DEL) in the mtDNA of MSC16; however, no genes or open reading frames (ORF) were found in the deleted region. Subsequent sequencing of the mtDNA of MSC16 revealed that 258 kb was either missing or significantly under-represented relative to the mtDNA of wild-type B (Del Valle-Echevarria et al. 2015). Ribosomal subunit S7 (*rps7*) is located in an under-represented region in the mtDNA of MSC16, resulting in significantly lower transcript abundance and the genetic basis of the MSC16 phenotype has been proposed to be impaired ribosome assembly and/or function (Del Valle-Echevarria et al. 2015).

In cucumber, there is a nuclear effect on the prevalence of specific mtDNAs transmitted via the male gametophyte to progeny. Havey et al. (2004) made crosses with MSC16 as the male parent to diverse plant introductions (PI) of cucumber and occasionally observed a high proportion of wild-type progenies. Using restriction fragment length polymorphisms (RFLPs), they demonstrated that these wild-type progenies possess the wild-type mtDNA that exists as a sublimon in MSC16; no evidence of maternal transmission of the mtDNA was observed (Havey et al. 2004). Lilly et al. (2001) demonstrated that MSC16 possesses wild-type sublimons that occasionally sort via the male gametophyte to produce the relatively rare wild-type progenies. The nuclear locus Paternal Sorting of Mitochondria (*Psm*) controls the predominance of MSC versus wild-type mitochondria in progenies from MSC16 as the male parent (Havey et al. 2004). When the seed parent is homozygous for the most common (+) allele at *Psm* and is crossed with MSC16 as the male, essentially all progenies are MSC. When the seed parent is homozygous for the rarer (–) allele at *Psm* and crossed with MSC16 as the male, wild-type progenies are produced that possess the wild-type mitochondrial sub-limon from MSC16. When the seed parent is heterozygous (+/–) at *Psm*, approximately equal numbers of wild-type and MSC progenies are observed (Havey et al. 2004). *Psm* has been mapped to a 459-kb region on chromosome 3 of cucumber (Al-Faifi et al. 2008; Calderón et al. 2012). In this research, we used next-generation sequencing and high-throughput genotyping to produce a finer map of *Psm* and identified a likely candidate gene controlling sorting of paternally transmitted mitochondria in cucumber.

Materials and methods

Identification and genotyping of single nucleotide polymorphisms (SNPs)

The mapping population was previously described by Calderón et al. (2012). A single plant (P3A) from USDA PI 401734 homozygous for the *Psm*– allele was crossed as the female with a single plant of *Cucumis sativus* var. *hardwickii* (*Csh*) homozygous for the *Psm*+ allele. A single hybrid plant was propagated by cuttings and each propagule self-pollinated to produce a large F₂ family. Each F₂ plant was self-pollinated and crossed as the female with MSC16 as the male. Numbers of wild-type and MSC progenies were scored for each family. F₃ progenies from F₂ plants homozygous for the *Psm*+ or *Psm*– allele were crossed as the female with MSC3, MSC 12, or MSC16 (Bartoszewski et al. 2004) as males, and progenies were scored as MSC or wild-type.

Total plant DNA from leaf tissue was isolated from the parents P3A and *Csh* (Calderón et al. 2012) for sequencing using the Illumina platform and manufacturer's protocols (Illumina, La Jolla CA). Single-end reads from P3A and *Csh* were trimmed using the program “cutadapt” to remove adapter sequences (Martin 2011), and the FastQC program was used to

select reads more than 40 basepairs (bp) in length and Phred value of at least 28. Sequences were aligned to scaffolds 02047 and 03356 of the Gy14 reference sequence (Yang et al. 2012) which span the region where *Psm* mapped (Calderón et al. 2012). The pipeline Simply Unified Pair-End Re-sequencing Workflow (SUPERW) was used to reveal SNPs between sequences of P3A, *Csh*, and the Gy14 reference (Sanseverino et al. 2015). Reads were aligned with Burrows-Wheeler Aligner (Li and Durbin 2009), followed by SNP calling with Samtools MpileUp and PileUp (Li et al. 2009). SNPs identified by both programs were chosen, and SNPs with greater than 97 % of reads in one parent and less than 2.5 % for the alternative allele with five reads or greater were chosen for further evaluation. The Integrative Genomics Viewer (Robinson et al. 2011; Thorvaldsdóttir et al. 2012) was used to capture 50 bp on both sides of the SNP. Forty-five SNPs were chosen that had no other SNPs in the flanking sequences and primer sequences were designed for genotyping (Supplemental File 1). DNAs from the segregating family from P3A by *Csh* were genotyped using the Fluidigm platform (South San Francisco, CA) and polymorphisms scored using the Fluidigm SNP Genotyping Analysis Software v3.

Genetic and statistical analyses

The R/qtl package (Broman et al. 2003) in the R software version 3.1.2 (R Core Team 2014) was used for mapping and trait analyses. For mapping, an initial screening of data quality was assessed based on the suggestions of Broman et al. (2003). A total of 45 SNPs and 340 F₂ individuals were selected after eliminating individuals with fewer than 25 (out of 45) SNPs scored, as well as markers that were genotyped on fewer than 100 (out of 349) individuals. Markers were evaluated for goodness-to-fit to the expected 1:2:1 ratio and those with poor fits were eliminated. After filtering, a total of 40 markers (Supplemental File 2) were used to construct the genetic map using the ripple command with a four-marker window, following the likelihood method (error probability = 0.005). After the genetic map was generated, percent wild-type progenies in crosses of segregating progenies with MSC16 as the male were used with composite interval mapping (CIM). The most significant peaks were evaluated for statistical significance using fitqtl with a one and two QTL models. A total of 10,000 permutations were performed to establish a threshold at probability of 0.05.

Gene annotation and sequence analyses

Sequences of parents (P3A and *Csh*) and Gy14 (Yang et al. 2012) between significant markers from the QTL analysis were extracted using the “extractseq” command from EMBOSS software and were loaded into GENES-CAN Web server (hosted by Massachusetts Institute of Technology, USA) under the following conditions: organism = *Arabidopsis*, suboptimal exon cutoff = 1.00 and print = CDS (coding sequences) and peptides. Genes in the significant region were searched using BLAST to identify putative functions. Protein sequences were compared to identify polymorphisms between P3A versus Gy14 and TMG1 (Sequence Read Archive SRX1091638) using Clustal Omega (version 1.2.1, <http://www.ebi.ac.uk/Tools/msa/clustalo/>). Genomic sequences of genes were surveyed for restriction-site polymorphisms using NEBcutter V2.0 (Vincze et al. 2003).

Expression analyses of candidate CsPPR336

Expression of CsPPR336 in leaves and flowers was compared using P3A (*Psm* $-/-$) versus GY14, Straight 8 (ST8), and inbred B (all *Psm* $+/+$). Plant growth, tissue sampling, RNA extraction, cDNA synthesis, and reverse-transcriptase quantitative (RT-q) PCR using primers (5'-TTCCCCTTTCCCTTCACTTC and 5'-CGGATG GACTGGTGGTAGAT) and *actin* as the control were performed as described by Del Valle-Echevarria et al. (2015).

Segregation of CsPPR336 and the MSC phenotype

Progenies from self-pollination of plant P3A (*Psm* $-/-$) were crossed as the seed parent with MSC16 (*Psm* $+/+$) and progenies were scored as wild-type or MSC. Three wild-type hybrid progenies were crossed as the female with three of the MSC hybrid progenies to determine segregations at CsPPR336. Progenies from these sib matings were scored as wild-type or MSC, and DNAs were extracted using NucleoSpin Plant II kit (Macherey–Nagel, Düren Germany). Candidate gene CsPPR336 was amplified using primers 5'-CAGAACTCCAAGACGCCTGT and 5'-GCC CTTTGAGCCTCTTCTTT. Amplicons were digested by *Bsa*WI followed the manufacturer's (New England Bio Labs) instructions and reactions were size fractionated through 1 %-agarose gels. Two simple sequence repeats [SSRs 13436 and 20338 (Yang et al. 2012)] flanking CsPPR336 were genotyped as previously described (Calderón et al. 2012).

Results

SNP discovery and genotyping

A total of 63,780,631 and 92,921,066 sequence reads from P3A and Csh remained after cleaning, and were deposited into the Sequence Read Archives as accessions SRR2096392 and SRR2096458, respectively. Reads from each parent were aligned against scaffolds 02047 and 03356 of the GY14 reference (Yang et al. 2012), yielding approximately 10 $\times$ coverage. After assembling to scaffolds, 828 SNPs were identified between P3A and *Csh* that met the initial quality parameters, of which 613 were chosen based on number of reads and support for the polymorphism. Of these 613 SNPs, 45 SNPs evenly spaced across the two scaffolds were selected (Supplemental File 1) and genotyped using 340 F₂ DNAs.

Genetic analysis

Numbers of wild-type and MSC progenies were scored for each family from crosses of F₂ plants as females with MSC16 as the male (Fig. 1). Composite Interval Mapping (CIM) using 40 filtered SNPs (Supplemental File 2) revealed two significant QTL across the region reported by Calderón et al. (2012). A major QTL at position 119 cM (LOD = 40) and a minor QTL at position 131 cM (LOD = 3.1) were above the threshold LOD of 2.6 (Fig. 2). The most significant QTL was located between SNP27 and SNP33, separated by 170 kb on GY14 scaffold 03356 (Fig. 2). The minor QTL was placed between SNP3 and SNP9 at a distance of 173 kb apart on the same scaffold, and was in repulsion phase with the major QTL. We analyzed one and two-QTL models to determine if both QTLs should be considered for gene discovery. An ANOVA performed on the major QTL was significantly ($p < 0.0001$) associated with percentage of wild-type progeny, explaining 69 % of the

variance. The ANOVA based on a two-QTL model increased the explained variation by only 3 % and no significant interaction between the two QTL was observed. Therefore, we focused on the region flanked by SNP27 and SNP33 (Fig. 2) to identify potential candidates for *Psm*.

In the 170-kb genomic region between SNP27 and SNP33, 18 coding sequences were identified by GENES-CAN and used for BLAST searches at the website of the International Cucurbit Genomics Initiative (version 2; <http://www.icugi.org/cgi-bin/ICuGI/genome/home.cgi?organism=cucumber&ver=2&cultivar=Chinese-long>). Six genes in this region were annotated to have organellar functions, two mitochondrial (Csa3M824840.1 and Csa3M824970.2) and four chloroplast (Csa3M824870.1, Csa3M825000.1, Csa3M825040.1 and Csa5M146950.1) (Supplemental File 3). We disregarded chloroplast-targeted genes because *Psm* is associated with paternally transmitted mitochondria. Of the two mitochondrially targeted genes, Csa3M824970.2 encodes a dihydroorotase that functions in pyrimidine synthesis (Kennedy 1974) and Csa3M824840.1 (Genbank accession XP_004136798.1) encodes a pentatricopeptide repeat (PPR) protein with 53 % similarity to PPR336 of *Arabidopsis thaliana* (At1G61870). We focused on Csa3M824840.1 as a potential candidate gene for *Psm* because *Arabidopsis* PPR336 has been proposed to confer stability to mitochondrial ribosomes (Uyttewaal et al. 2008); we previously reported that the mtDNA of MSC16 has fewer copies and transcripts of *rps7*, a key subunit in mitochondrial ribosomes (Del Valle-Echevarria et al. 2015).

PPR336 expression and sequence analyses

Expression of the cucumber homolog of PPR336 (CsPPR336) was assessed by RT-qPCR and transcripts were detected in leaves and flowers of P3A (*Psm* $-/-$) versus ST8, GY14, and inbred B (all *Psm* $+/+$). However, there were no significant ($p < 0.05$) differences for transcript levels among these cucumbers, revealing similar expression levels of CsPPR336 in leaves and flowers of plants homozygous for either the *Psm* $+$ or *Psm* $-$ allele.

Sequences of CsPPR336 were compared to identify amino-acid and restriction-site differences between Gy14 and TMG1 (both *Psm* $+/+$) versus P3A (*Psm* $-/-$). Two amino-acid differences (P168S and S220R) distinguished CsPPR336 from P3A versus Gy14 and TMG1 (Fig. 3). The P168S polymorphism is located before the first PPR motif and the S220R polymorphism is located in the second PPR motif, according to the annotation of CsPPR336 (Genbank XP_004136798.1).

Segregation analysis of restriction-site polymorphism in CsPPR336

The S220R polymorphism in CsPPR336 results in the loss of a *Bsa*WI restriction site in P3A (Fig. 4), putatively the *Psm* $-$ allele. We determined segregations for this *Bsa*WI polymorphism and the phenotypes (MSC versus wild-type) of progenies. The cross of P3A (*Psm* $-/-$) as the female with MSC16 (*Psm* $+/+$) as the male produced 98 % (638 of 654) wild-type and 2 % (16 of 654) MSC progenies. We then crossed three wild-type hybrid progenies as the female with three of the relatively rare MSC progenies as the male (all plants should be heterozygous at *Psm*), and progenies from this cross produced the expected 1:1 ratio of wild-type versus MSC progeny ($p = 0.880$). For 96 randomly selected wild-type

progenies, the presence versus absence of the *Bsa*WI site segregated 54:42, fitting the expected 1:1 ratio ($p = 0.221$). For 95 of 96 randomly selected MSC progenies, the *Bsa*WI site in CsPPR336 (putatively the *Psm*⁺ allele) was present, indicating that the combination of MSC mitochondria with the homozygous *Psm*[−] genotype is deleterious. For the one MSC plant that lacked the *Bsa*WI site in CsPPR336 (putatively *Psm*^{−/−}), we were not able to amplify mitochondrial *rps7* from its genomic DNA, which is consistent with our visual classification of the phenotype. Two SSRs (13436 and 20338) flanking CsPPR336 were polymorphic between P3A and MSC16 (Calderón et al. 2012), and were genotyped using DNA from this single plant putatively MSC and *Psm*^{−/−}. Both SSR 13436 and the *Bsa*WI site in CsPPR336 carried the P3A allele; however, SSR 20338 was heterozygous indicating that a cross-over had occurred between SSR 20338 and CsPPR336, a region that spans approximately 278 kb (Yang et al. 2012). The minor QTL affecting the proportion of wild-type progenies mapped between SNP3 and SNP9 (Fig. 2), a region located approximately 457–293 kb from SSR 20338 in the direction away from CsPPR336. This indicates that the cross-over between SSR 20338 and CsPPR336 revealed the effect of this minor QTL and explains the recovery of one progeny with the MSC phenotype and homozygous for the *Psm*[−] allele.

Segregation analysis of *Psm* homozygotes crossed with different MSC lines

MSC3, MSC 12, and MSC 16 are mitochondrial mutants selected from different cell cultures, all established using highly inbred B (Bartoszewski et al. 2004). MSC3 has an under-representation of the polycistronic *nad5ex4-atp4-nad5ex5* mitochondrial region; both MSC12 and 16 possess an under-representation of the *rps7* coding region (Del Valle-Echevarria et al. 2015). F₃ families from the self pollination of F₂ plants from P3A by Csh that were either homozygous for the *Psm*⁺ or *Psm*[−] alleles were chosen, and F₃ progenies were crossed as females with MSC3, MSC12 or MSC16. *Psm*^{+/+} plants produced almost exclusively MSC progenies from crosses with all three MSC lines as males (Table 1). Crosses of *Psm*^{−/−} plants with MSC12 and MSC16 as males produced almost exclusively wild-type progenies, revealing similar results for two independently produced MSC lines both carrying an under-representation of the same mitochondrial region (Del Valle-Echevarria et al. 2015). However, crosses of *Psm*[−] homozygotes with MSC3 males produced wild-type to MSC progenies that fit a 1:1 ratio for all of four independent families ($p = 0.317$ – 0.785). This indicates that another nuclear locus independent of *Psm* may interact with the MSC3 mitochondrial mutant to yield wild-type progenies.

Discussion

The MSC mutants of cucumber are useful research tools for mitochondrial genetics (Lilly et al. 2001; Bartoszewski et al. 2004, 2007; Del Valle-Echevarria et al. 2015). Sequencing of the mtDNA of MSC16 revealed missing and under-represented regions relative to its progenitor wild-type inbred B (Lilly et al. 2001; Del Valle-Echevarria et al. 2015), and that MSC16 has significantly fewer copies of the *rps7* gene (Del Valle-Echevarria et al. 2015). The prokaryotic homolog (*rpsG* of *E. coli*) of *rps7* functions as a first-hierarchy protein in the assembly of the small subunit of the ribosome (Shajani et al. 2011). We hypothesized that compromised ribosomes may be responsible for the MSC16 phenotype; however,

proteomic analysis did not show significant differences in protein production between MSC16 and wild-type B (Del Valle-Echevarria et al. 2015). It is possible that translation of certain mitochondrial transcripts is dependent on the *RPS7* protein (Fargo et al. 2001) and results in the MSC phenotype (Del Valle-Echevarria et al. 2015).

Psm is a nuclear locus that sorts paternally transmitted, wild-type mitochondria in progenies from crosses of MSC16 as the male (Havey et al. 2004). Genetic mapping placed *Psm* in a 459-kb region on chromosome 3 of cucumber (Al-Faifi et al. 2008; Calderón et al. 2012). Genes involved with surveillance of mtDNA structure in other organisms, such as *Msh1* and *RecA* homologs in *A. thaliana* (Shedge et al. 2007), did not map to this region suggesting that a novel gene may regulate sorting for paternally transmitted, wild-type mitochondria in cucumber (Al-Faifi et al. 2008; Calderón et al. 2012). Using next-generation sequencing and high-throughput genotyping, we mapped SNPs using the segregating family described by Calderón et al. (2012) and detected a major QTL controlling paternal sorting of mitochondria across a 170-kb region on chromosome 3 (Fig. 2). We focused on Csa3M824840.1 as a likely candidate for *Psm* because it is a PPR-containing gene annotated as PPR336 in *Arabidopsis* (At1G61870). This gene is associated with stability of mitochondrial polysomes (groups of ribosomes translating the same mRNA) in *Arabidopsis* (Uyttewaal et al. 2008). Mutants of PPR336 in *Arabidopsis* were not phenotypically different from the wild-type plants, but mitochondrial polysomes were of lower molecular weight suggesting that this mitochondrially targeted protein plays a role in mitochondrial translation (Uyttewaal et al. 2008). Williams and Barkan (2003) reported that the PPR2 protein of maize is associated with ribosome accumulation in plastids and a mutant (*ppr2*) lacked plastid-translation products. Since MSC16 shows reduced transcription of *rps7*, a key component of ribosome assembly (Shajani et al. 2011), CsPPR336 as the candidate for *Psm* is consistent with a nuclear effect on ribosome assembly or stability in the MSC16 mitochondrial mutant. Although transformation of cucumber remains challenging, proof that CsPPR336 is indeed the *Psm* locus will require transformation of the *Psm*⁺ allele into a *Psm*^{-/-} plant, followed by crossing of transformed plants with MSC16 as the male and scoring of progenies as MSC versus wild-type. Since there were no expression-level differences between plants homozygous for the *Psm*⁺ and *Psm*⁻ alleles, the two amino-acid changes may be responsible for the different phenotypes, and we expect that families from crosses of transformed plants with MSC16 as the male will show greater frequencies of MSC progenies.

Segregation of the *Bsa*WI polymorphism in CsPPR336 indicated that the absence of the restriction site (putatively the *Psm*⁻ allele) together with MSC mitochondria may be lethal in the male gametophyte. Wild-type progenies segregated 1:1 for the presence versus absence of the *Bsa*WI site ($p = 0.884$); however, 95 of 96 MSC progenies possessed the *Bsa*WI site (putatively the *Psm*⁺ allele) and only one MSC plant lacked this restriction-enzyme site. Assuming that absence of the *Bsa*WI site corresponds to the *Psm*⁻ allele, these observations reveal strong selection against progenies that are homozygous for the *Psm*⁻ allele in the presence of MSC mitochondria. We previously reported that crosses of *Psm*^{-/-} plants with MSC16 as the male produce rare MSC progenies (Al-Faifi et al. 2008; Calderón et al. 2012). Our recovery of one MSC plant with the *Psm*^{-/-} genotype is consistent with these previous

observations; however, this plant resulted from a gamete carrying a cross-over between CsPPR336 and the minor QTL on chromosome 3 and this may account for its survival.

The effect of CsPPR336 on the predominance of wild-type or MSC mitochondria in progenies could be an active or passive process. Passive selection against MSC mitochondria could occur if CsPPR336 encoded by the *Psm*⁺ allele stabilized mitochondrial ribosomes, while the protein from *Psm*[−] allele does not and selects against MSC mitochondria allowing relatively rare wild-type mitochondria or mitochondrial sublimons to predominate to produce wild-type progenies. Active selection could occur if CsPPR336 from the *Psm*[−] allele with compromised ribosomes from MSC16 provided a signal for active degradation or elimination of MSC mitochondria, again allowing wild-type sublimons or mitochondria to prevail.

This research illustrates the usefulness of cucumber for studying important nuclear-organellar interactions. The unique transmission modes of the cucumber organelles [maternal for chloroplast and paternal for mitochondria (Havey 1997)] allow for the separation of organellar effects by reciprocal crossing. Independently derived MSC mutants of cucumber possess unique phenotypes and different rearrangements and under-represented regions in their mtDNAs (Bartoszewski et al. 2004; Del Valle-Echevarria et al. 2015). The *Psm* locus specifically affects the proportion of wild-type progenies when MSC12 and MSC16 are crossed as the male parent (Table 1), and both of these independently selected MSC lines possess an under-representation of the mitochondrial region carrying the *rps7* gene. However, *Psm* does not show this same effect when MSC3 is used as the male parent (Table 1), revealing a specific nuclear-mitochondrial interaction between alleles at *Psm* with the mitochondria of MSC12 and MSC16. Crosses of different MSC mutants with diverse accessions of cucumber should reveal nuclear genes such as *Psm*, specifically interacting with the MSC mitochondria to recover wild-type plants, and lead towards a better understanding of important nuclear-mitochondrial interactions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the Agricultural Research Service of the United States Department of Agriculture, a Science and Medicine Graduate Research Scholars Fellowship, an Advance Opportunity Fellowship of the Graduate School of the University of Wisconsin-Madison, and the National Institutes of Health (NIH) Biotechnology Predoctoral Training Program (BTP) of the University of Wisconsin-Madison (NIH Grant 5 T32 GM08349).

References

- Al-Faifi S, Meyer JDF, Garcia-Mas J, Montforte AJ, Havey MJ. Exploiting synteny in *Cucumis* for mapping of *Psm*: a unique locus controlling paternal mitochondrial sorting. *Theor Appl Genet*. 2008; 117:523–529. [PubMed: 18521565]
- Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol*. 2010; 27:1436–1448. [PubMed: 20118192]

- Alverson AJ, Rice DW, Dickinson S, Barry K, Palmer JD. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell*. 2011; 23:2499–2513. [PubMed: 21742987]
- Bartoszewski G, Malepszy S, Havey MJ. Mosaic (MSC) cucumbers regenerated from independent cell cultures possess different mitochondrial rearrangements. *Curr Genet*. 2004; 45:45–53. [PubMed: 14586555]
- Bartoszewski G, Havey MJ, Ziółkowska A, Długosz M, Malepszy S. The selection of mosaic (MSC) phenotype after passage of cucumber (*Cucumis sativus* L.) through cell culture—a method to obtain plant mitochondrial mutants. *J Appl Genet*. 2007; 1:1–9.
- Broman K, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. *Bioinformatics*. 2003; 19:889–890. [PubMed: 12724300]
- Calderón CI, Yandell BS, Havey MJ. Genetic mapping of paternal sorting of mitochondria in cucumber. *Theor Appl Genet*. 2012; 125:11–18. [PubMed: 22350175]
- Del Valle-Echevarria AR, Kielkowska A, Bartoszewski G, Havey MJ. The mosaic mutants of cucumber: a method to produce knock-downs of mitochondrial transcripts. *Genes, Genomes, Genetics*. 2015; G3(5):1211–1221.
- Fargo DC, Boynton JE, Gillham NW. Chloroplast ribosomal protein S7 of *Chlamydomonas* binds to chloroplast mRNA leader sequences and may be involved in translation initiation. *Plant Cell*. 2001; 13:207–218. [PubMed: 11158540]
- Feiler HS, Newton KJ. Altered mitochondrial gene expression in the nonchromosomal stripe 2 mutant in maize. *EMBO*. 1987; 6:1535–1539.
- Havey MJ. Predominant paternal transmission of the mitochondrial genome cucumber. *J Hered*. 1997; 88:232–235.
- Havey MJ, McCreight JD, Rhodes B, Taurick G. Differential transmission of the *Cucumis* organellar genomes. *Theor Appl Genet*. 1998; 97:122–128.
- Havey MJ, Park YH, Bartoszewski G. The *Psm* locus controls paternal sorting of the cucumber mitochondrial genome. *J Hered*. 2004; 95:492–497. [PubMed: 15475394]
- Kennedy J. Dihydroorotase from rat liver: purification, properties and regulatory role in pyrimidine biosynthesis. *Arch Biochem Biophys*. 1974; 160(2):358–365. [PubMed: 4831620]
- Kitazaki K, Kubo T. Cost of having the largest mitochondrial genome: evolutionary mechanism of plant mitochondrial genome. *J Bot*. 2010; doi: 10.1155/2010/620137
- Łady y ski M, Burza W, Malepszy S. Relationship between somaclonal variation and type of culture in cucumber. *Euphytica*. 2002; 125:349–356.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*. 2009; 25:1754–1760. [PubMed: 19451168]
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing Subgroup (2009) (2009) the sequence alignment/map format and SAMtools. *Bioinformatics*. 2010; 25:2078–2079.
- Lilly JW, Havey MJ. Small, repetitive DNAs contribute significantly to the expanded mitochondrial genome of cucumber. *Genetics*. 2001; 159:317–328. [PubMed: 11560907]
- Lilly JW, Bartoszewski G, Malepszy S, Havey MJ. A major deletion in the cucumber mitochondrial genome sorts with the MSC phenotype. *Curr Genet*. 2001; 40:144–151. [PubMed: 11680824]
- Mackenzie S, Hue S, Lyznik A. The elusive plant mitochondria as a genetic system. *Plant Physiol*. 1994; 105:775–780. [PubMed: 12232241]
- Malepszy S, Burza W, Ćmiech M. Characterization of a cucumber (*Cucumis sativus* L.) somaclonal variant with paternal inheritance. *J Appl Genet*. 1996; 37:65–78.
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011; 17:10–12.
- Palmer JD, Herbon LA. Uncircular structure of the *Brassica hirta* mitochondrial genome. *Curr Genet*. 1987; 11:565–570. [PubMed: 2836082]
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotech*. 2011; 29:24–26.

- Rodriguez-Moreno L, Gonzalez VM, Benjak A, Marti MC, Puigdomenech P, Aranda MA, Garcia-Mas J. Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having nuclear origin. *BMC Genom.* 2011; 12:424.
- Sanseverino W, Hénaff E, Vives C, Pinosio S, Burgos-Paz W, Morgante M, Ramos-Onsins SE, Garcia-Mas J, Casacuberta JM. Transposon insertion, structural variations and SNPs contribute to the evolution of the melon genome. *Mol Biol Evol.* 2015; 32:2760–2774. [PubMed: 26174143]
- Shajani Z, Sykes MT, Williamson JR. Assembly of bacterial ribosomes. *Annu Rev Biochem.* 2011; 80:501–526. [PubMed: 21529161]
- Shedge V, Arrieta-Montiel M, Christenses AC, Mackenzie SA. Plant mitochondrial recombination surveillance requires unusual *RecA* and *MutS* homologs. *Plant Cell.* 2007; 19:1251–1264. [PubMed: 17468263]
- Shen J, Dirks R, Havey MJ. Diallel crossing among doubled haploids of cucumber reveals significant reciprocal-cross differences. *J Amer Soc Hort Sci.* 2015; 140:178–182.
- Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* 2012; 10(1):e1001241. [PubMed: 22272183]
- R Core Team. R: A language and environment for statistical computing. 2014. <http://www.R-project.org/>
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 2012; 14:178–192. [PubMed: 22517427]
- Uyttewaal M, Mireau H, Rurek M, Hammani K, Arnal N, Quadrado M, Giegé P. PPR336 is associated with polysomes in plant mitochondria. *J Mol Biol.* 2008; 375:626–636. [PubMed: 18054044]
- Vincze T, Posfai J, Roberts RJ. NEBcutter: a program to cleave DNA with restriction enzymes. *Nucl Acids Res.* 2003; 31:3688–3691. [PubMed: 12824395]
- Ward BL, Anderson RS, Bendich AJ. The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). *Cell.* 1981; 25:793–803. [PubMed: 6269758]
- Williams PM, Barkan A. A chloroplast-localized PPR protein required for plastid ribosome accumulation. *Plant J.* 2003; 36:675–686. [PubMed: 14617068]
- Yang L, Koo D, Li Y, Zhang X, Luan F, Havey MJ, Jiang J, Weng Y. Chromosome rearrangements during domestication of cucumber as revealed from high-density genetic mapping and draft genome assembly. *Plant J.* 2012; 71:895–906. [PubMed: 22487099]

Fig. 1. Families of wild-type (WT), mosaic (MSC), and segregating (WT + MSC) progenies from crosses of plants segregating for Paternal Sorting of Mitochondria (*Psm*) with MSC16 as the male parent. Each row is seed from one cross and the percentage of wild-type progenies was used for mapping of *Psm*

Fig. 2.

Composite interval mapping of percentage of wild-type progenies from crosses of F_2 plants segregating for Paternal Sorting of Mitochondria (*Psm*) with MSC16 as the male parent. Single nucleotide polymorphisms (SNPs) along the x -axis are listed in Supplemental File 1. *Numbers in parentheses* following SNPs indicate kilobase positions. *Horizontal line* shows threshold LOD score from permutation analysis. Scaffolds from the sequence assembly of GY14 (Yang et al. 2012) are shown below SNPs

PPR336_Gy14	MAAALPRTPRRLFLISRLHSFSYSTTPPLQPTSDSPFSLRAAKSAILSQSDPKLAQSF
PPR336_TM1	MAAALPRTPRRLFLISRLHSFSYSTTPPLQPTSDSPFSLRAAKSAILSQSDPKLAQSF
PPR336_P3A	MAAALPRTPRRLFLISRLHSFSYSTTPPLQPTSDSPFSLRAAKSAILSQSDPKLAQSF
PPR336_Gy14	IQASTLPSCRYRPIYHQSIKRLARAQRFDLIDVLIQSHHKSPSATSEGFWIRLIMLYSS
PPR336_TM1	IQASTLPSCRYRPIYHQSIKRLARAQRFDLIDVLIQSHHKSPSATSEGFWIRLIMLYSS
PPR336_P3A	IQASTLPSCRYRPIYHQSIKRLARAQRFDLIDVLIQSHHKSPSATSEGFWIRLIMLYSS
PPR336_Gy14	VGMVNQALYILDQAILHKSCNLSEKSLCAILSVFLDNSMPEKVHEMFRSIPKIGVTPTA
PPR336_TM1	VGMVNQALYILDQAILHKSCNLSEKSLCAILSVFLDNSMPEKVHEMFRSIPKIGVTPTA
PPR336_P3A	VGMVNQALYILDQAILHKSCNLSEKSLCAILSVFLDNSMPEKVHEMFRSIPKIGVTPTA
PPR336_Gy14	VSHNLVLKAFVRQNDLPSARNWIDELCKDDAKVIPNIDSFTILLGAYWSNGDMIGFDEIE
PPR336_TM1	VSHNLVLKAFVRQNDLPSARNWIDELCKDDAKVIPNIDSFTILLGAYWSNGDMIGFDEIE
PPR336_P3A	VSHNLVLKAFVRQNDLPSARNWIDELCKDDAKVIPNIDSFTILLGAYWSNGDMIGFDEIE
PPR336_Gy14	KEISKRGLEFNLATYNYRISRLCKNKECARAKKILDEMISGKVPNSSSYDSIIHGCDV
PPR336_TM1	KEISKRGLEFNLATYNYRISRLCKNKECARAKKILDEMISGKVPNSSSYDSIIHGCDV
PPR336_P3A	KEISKRGLEFNLATYNYRISRLCKNKECARAKKILDEMISGKVPNSSSYDSIIHGCDV
PPR336_Gy14	GDIESAMKILKGILEDGHVSPTSRIYYRLIRSMVKEGEFEMALETCRETIKRRWVPPEA
PPR336_TM1	GDIESAMKILKGILEDGHVSPTSRIYYRLIRSMVKEGEFEMALETCRETIKRRWVPPEA
PPR336_P3A	GDIESAMKILKGILEDGHVSPTSRIYYRLIRSMVKEGEFEMALETCRETIKRRWVPPEA
PPR336_Gy14	MEALVRGLVAMSKVEEAKVVEKMKRRLKGPVDSWRKIEAALPL
PPR336_TM1	MEALVRGLVAMSKVEEAKVVEKMKRRLKGPVDSWRKIEAALPL
PPR336_P3A	MEALVRGLVAMSKVEEAKVVEKMKRRLKGPVDSWRKIEAALPL

Fig. 3.
Amino-acid alignment between predicted protein sequences of CsPPR336 from Gy14 and TMG1 (both *Psm* +/+) and 401734-3A (P3A; *Psm* -/-)

Fig. 4.
*Bsa*WI digestion of genomic amplicons of CsPPR336 from 401734-3A (P3A; *Psm* -/-), inbred B (*Psm* +/+), and an artificial heterozygote (AH) from mixture of equal amounts of DNAs from P3A and inbred B

Table 1

Genotypes of female parents and numbers of wild-type (WT) and mosaic (MSC) progenies from crosses with mitochondrial mutants MSC3, MSC12, or MSC16

Parent ^a	Female	Male		Observed		Proportion	
		Genotype	Parent	WT	MSC	WT	WT
C10089E	<i>Psm</i> +/+	MSC16	0	54	0.00		
C10089 K	<i>Psm</i> +/+	MSC16	0	43	0.00		
C10089L	<i>Psm</i> +/+	MSC16	0	41	0.00		
C10089R	<i>Psm</i> +/+	MSC16	0	44	0.00		
C10090E	<i>Psm</i> +/+	MSC16	0	51	0.00		
C10090G	<i>Psm</i> +/+	MSC16	0	48	0.00		
C10090Q	<i>Psm</i> +/+	MSC16	0	45	0.00		
C10090S	<i>Psm</i> +/+	MSC16	0	46	0.00		
C10089E	<i>Psm</i> +/+	MSC12	0	52	0.00		
C10089F	<i>Psm</i> +/+	MSC12	0	51	0.00		
C10089 M	<i>Psm</i> +/+	MSC12	2	46	0.04		
C10089T	<i>Psm</i> +/+	MSC12	1	47	0.02		
C10090F	<i>Psm</i> +/+	MSC12	0	48	0.00		
C10090 N	<i>Psm</i> +/+	MSC12	1	45	0.02		
C10089C	<i>Psm</i> +/+	MSC3	0	50	0.00		
C10089D	<i>Psm</i> +/+	MSC3	0	51	0.00		
C10090C	<i>Psm</i> +/+	MSC3	1	45	0.02		
C10090 M	<i>Psm</i> +/+	MSC3	0	46	0.00		
C10093 N	<i>Psm</i> -/-	MSC16	56	0	1.00		
C10093O	<i>Psm</i> -/-	MSC16	53	0	1.00		
C10093P	<i>Psm</i> -/-	MSC16	55	0	1.00		
C10093Q	<i>Psm</i> -/-	MSC16	55	2	0.96		
C10093U	<i>Psm</i> -/-	MSC16	54	0	1.00		
C10086A	<i>Psm</i> -/-	MSC16	156	0	1.00		
C10086D	<i>Psm</i> -/-	MSC16	125	0	1.00		
C10086I	<i>Psm</i> -/-	MSC16	53	0	1.00		

Female Parent ^a	Genotype	Male		Observed		Proportion	
		Parent		WT	MSC	WT	WT
C10086 J	<i>Psm</i> -/-	MSC16		136	0	1.00	
C10091F	<i>Psm</i> -/-	MSC16		42	5	0.89	
C10093H	<i>Psm</i> -/-	MSC12		53	8	0.87	
C10093P	<i>Psm</i> -/-	MSC12		49	2	0.96	
C10093Q	<i>Psm</i> -/-	MSC12		53	0	1.00	
C10093R	<i>Psm</i> -/-	MSC12		20	3	0.87	
C10091D	<i>Psm</i> -/-	MSC12		41	1	0.98	
C10091G	<i>Psm</i> -/-	MSC12		63	5	0.93	
C10091R	<i>Psm</i> -/-	MSC12		52	1	0.98	
C10091R	<i>Psm</i> -/-	MSC12		55	2	0.96	
C10091R	<i>Psm</i> -/-	MSC12		107	3	0.97	
C10093E	<i>Psm</i> -/-	MSC3		15	10	0.60	
C10091Q	<i>Psm</i> -/-	MSC3		23	18	0.56	
C10091Q	<i>Psm</i> -/-	MSC3		28	26	0.52	
C10091T	<i>Psm</i> -/-	MSC3		22	18	0.55	

^aWhen female parent is listed more than once, multiple fruit were produced and progenies from each fruit were separately evaluated