

RESEARCH ARTICLE

Analysis of the lineage of *Phytophthora infestans* isolates using mating type assay, traditional markers, and next generation sequencing technologies

Ramadan A. Arafa^{1,2}, Said M. Kamel¹, Mohamed T. Rakha³, Nour Elden K. Soliman⁴, Olfat M. Moussa⁴, Kenta Shirasawa^{2*}

1 Plant Pathology Research Institute, Agricultural Research Center, Giza, Egypt, **2** Department of Frontier Research and Development, Kazusa DNA Research Institute, Chiba, Japan, **3** Department of Horticulture, Faculty of Agriculture, University of Kafrelsheikh, Kafr El-Sheikh, Egypt, **4** Department of Plant Pathology, Faculty of Agriculture, Cairo University, Giza, Egypt

* shirasaw@kazusa.or.jp



OPEN ACCESS

Citation: Arafa RA, Kamel SM, Rakha MT, Soliman NEK, Moussa OM, Shirasawa K (2020) Analysis of the lineage of *Phytophthora infestans* isolates using mating type assay, traditional markers, and next generation sequencing technologies. PLoS ONE 15(1): e0221604. <https://doi.org/10.1371/journal.pone.0221604>

Editor: Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

Received: August 6, 2019

Accepted: December 17, 2019

Published: January 21, 2020

Copyright: © 2020 Arafa et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Double digest restriction-site associated DNA sequencing (ddRAD-Seq) and whole genome resequencing (WGRS) data are available at the DDBJ Sequence Read Archive database under the accession number DRA007610. These data may also be accessed via the NCBI Sequence Read Archive under the same accession number.

Funding: This work was supported by The Ministry of Higher Education and Scientific Research

Abstract

Phytophthora infestans (Mont.) de Bary, a hemibiotrophic oomycete, has caused severe epidemics of late blight in tomato and potato crops around the world since the Irish Potato Famine in the 1840s. Breeding of late blight resistant cultivars is one of the most effective strategies to overcome this disruptive disease. However, *P. infestans* is able to break down host resistance and acquire resistance to various fungicides, possibly because of the existence of high genetic variability among *P. infestans* isolates via sexual and asexual reproduction. Therefore, to manage this disease, it is important to understand the genetic divergence of *P. infestans* isolates. In this study, we analyzed the genomes of *P. infestans* isolates collected from Egypt and Japan using various molecular approaches including the mating type assay and genotyping simple sequence repeats, mitochondria DNA, and effector genes. We also analyzed genome-wide single nucleotide polymorphisms using double-digest restriction-site associated DNA sequencing and whole genome resequencing (WGRS). The isolates were classified adequately using high-resolution genome-wide approaches. Moreover, these analyses revealed new clusters of *P. infestans* isolates in the Egyptian population. Monitoring the genetic divergence of *P. infestans* isolates as well as breeding of resistant cultivars would facilitate the elimination of the late blight disease.

Introduction

Phytophthora infestans (Mont.) de Bary is a hemibiotrophic oomycete that causes late blight disease in tomato (*Solanum lycopersicum*) and potato (*Solanum tuberosum*). Globally, many epidemics of tomato and potato late blight have been reported since the end of the 19th century, among which the Irish Potato Famine in the 1840s was probably the worst. *P. infestans* is reported to have originated either in South America [1, 2] or in central Mexico [3, 4]. Although

(MHESR), Egypt (Grant Number: 2013/2014-547), and the Kazusa DNA Research Institute Foundation, Japan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

numerous studies have been conducted on late blight [5], *P. infestans* remains a major threat to global food security [6]. *P. infestans* is a heterothallic pathogen with two mating types (A1 and A2), which reproduce sexually through oospores [7]. *P. infestans* also exhibits asexual reproduction, with high genetic variability. Since 1980, the A1 mating type was dominant globally, outside the region of its origin [8, 9]. Then, the A2 mating type appeared in many countries [10]. This was followed by the reappearance of the A1 mating type, which replaced A2 and spread in other regions [11]. This repeated appearance and disappearance of mating types might explain why *P. infestans* is able to break down major resistance genes in host plant species and develop resistance against various fungicides [12].

Breeding of late blight resistant cultivars is one of the most effective strategies to overcome this disruptive disease. Resistance genes and/or DNA markers linked to resistance loci are powerful tools for the selection of resistant lines via marker-assisted selection. To the best of our knowledge, genetic loci have been identified in late blight resistant germplasms of wild tomato, *Solanum pimpinellifolium*, in breeding programs [13–15]. More recently, we identified another resistance locus from a susceptible line of cultivated tomato, ‘Castlerock’, which has a minor effect on the resistance phenotype (Arafa et al. under submission). These results suggest that plants harbor many resistance genes and employ several defense mechanisms against the late blight pathogen. Additionally, the pathogenicity of *P. infestans* is complicated and has evolved over time, probably because of sequence variations in virulence genes, also known as effector genes, which encode proteins that suppress the plant immune system and alter the host response to accelerate the infection process. Therefore, knowledge of the genetic divergence of *P. infestans* isolates is important to control the late blight disease.

Over the last several decades, PCR assays for assessing nucleotide sequence variations in mitochondrial DNA (mtDNA), simple sequence repeats (SSRs; also known as microsatellites), and effector genes have been employed to investigate the genetic diversity of *P. infestans* populations [16–20]. Advances in sequencing techniques, based on next generation sequencing (NGS) technology, have enabled the analysis of nucleotide sequence variations on a genome-wide scale. Since the recent release of the *P. infestans* reference genome sequence [21], reduced-representation sequencing approaches, such as genotyping by sequencing [22] and restriction-site associated DNA sequencing [23], have been used to scan genome-wide sequence variations. Moreover, because of the contiguous decrease in the cost of NGS and increase in its throughput, WGRS analysis has enabled the elucidation of the diversity of *P. infestans* populations on a genome-wide level. The genome-wide genotyping methods have been applied to *P. infestans* populations to reveal their phylogenetic relationship [17, 19]; however, the number of *P. infestans* isolates tested has been limited and insufficient to cover the *P. infestans* populations spread across the world.

A decision support system in agricultural domains would be essential for successful plant disease management. Knowledge of the evolution of *P. infestans* clonal lineages, based on DNA analysis, epidemiology, and population dynamics of this destructive plant pathogen [19], would be quite useful for the system. Therefore, in this study, we aimed to characterize *P. infestans* populations collected from Egypt and Japan; neither of these regions has been represented in the genome-wide genotyping analysis of *P. infestans* isolates to date. We investigated the mating types, mtDNA haplotypes, and SSR marker genotypes of the isolates as well as genome-wide genotypes, based on double-digest restriction-site associated DNA sequencing (ddRAD-Seq) and WGRS approaches, to gain insights into the diversity of *P. infestans* populations at the genome-wide level.

Results

Mating types in the tested isolates

A total of 80 *P. infestans* isolates were used in this study; these included 62 isolates collected from seven counties in Egypt, and 18 isolates obtained from the National Institute of Agrobiological Sciences (NIAS) in Japan, which were collected from three prefectures (S1 Table). Among these 80 isolates, 67 were of the A1 mating type, 12 were of the A2 mating type, and one was self-fertile (SF). All three mating types were identified among the 18 isolates obtained from the NIAS, whereas only the A1 mating type was observed among the 62 isolates collected from Egypt (Fig 1 and S1 Table). In the NIAS collection, the A1 mating type isolates were collected from potato leaves in Hokkaido prefecture in 2013, while the A2 mating type isolates, except four, were collected from tomato leaves and fruits in Ibaraki prefecture in 1991–1992. The remaining three A2 mating type were collected from potato in Hokkaido and from tomato in Sizuoka prefecture. The SF was isolated from tomato in Ibaraki.

Analysis of mtDNA polymorphisms

The 80 isolates were analyzed by PCR using four restriction fragment length polymorphism (RFLP) markers to detect sequence variations in P1, P2, P3, and P4 regions of the mitochondrial genome (Fig 1 and S1 and S2 Tables). The mtDNA polymorphisms classified the isolates into two groups, Ia and IIa. Isolates in each group showed a strong association with their area of collection; isolates in group Ia were collected from Egypt, while those in group IIa were obtained from Japan. Additionally, polymorphisms in two hypervariable regions (HVRs) in the mitochondrial genome were analyzed, which divided the isolates into five groups. The IR1 type was the most dominant among Egyptian isolates (61/62), while three IIR2 and fourteen IIR3 types were observed among the 18 Japanese isolates. The exceptions were one IR2 type in Egyptian isolates (EG_11) and one IR1 type in Japanese isolates (K92-62).

Analysis of effector gene polymorphism

We tested the PCR assay to detect five effector genes, *AVR1*, *AVR2*, *AVR2-like*, *AVR3a*, and *AVR4*, across all 80 isolates. DNA amplicons were obtained for four genes (*AVR2*, *AVR2-like*, *AVR3a*, and *AVR4*) (Fig 1 and S1 and S2 Tables) and a positive control (*NitRed*), but not for *AVR1*. The *AVR4* and *NitRed* genes were amplified from all 80 isolates. However, the remaining three genes showed polymorphism; *AVR3a*, *AVR2*, and *AVR2-like* were not detected in 4, 12, and 19 isolates, respectively. All of these three genes (*AVR3a*, *AVR2*, and *AVR2-like*) were not detected in the Egyptian isolate EG_36, while *AVR2-like* was not amplified in all of the Japanese isolates.

SSR genotyping analysis

We tested 16 SSR (or microsatellite) markers in the 80 isolates. A total of 54 peaks were detected, ranging from two peaks in five markers (Pi33, Pi70, Pi89, SSR2, and SSR8) to seven peaks in SSR4 (S1 and S2 Tables). The 62 Egyptian and 18 Japanese isolates showed 45 and 38 peaks, respectively, for the 16 SSR markers. Clustering analysis divided the 80 isolates into four clusters (Fig 1 and S1 Table). Two clusters, EGY1 and EGY2, contained 50 and 12 Egyptian isolates, respectively. All of the isolates in the EGY2 cluster were collected in the 2014–2015 growing season, while those in the EGY1 were sampled in 2014–2015 and 2015–2016. The EGY1 and EGY2 clusters showed no relationship with the geographic location, host species, and infected organs, and all isolates in the EGY2 cluster lacked the *AVR2* amplicon. The other two clusters contained only Japanese isolates, with 13 isolates in JPN1 and five in JPN2. Isolates in the JPN1 cluster were collected from 1987–1992, while those in the JPN2 cluster were collected in 2013.

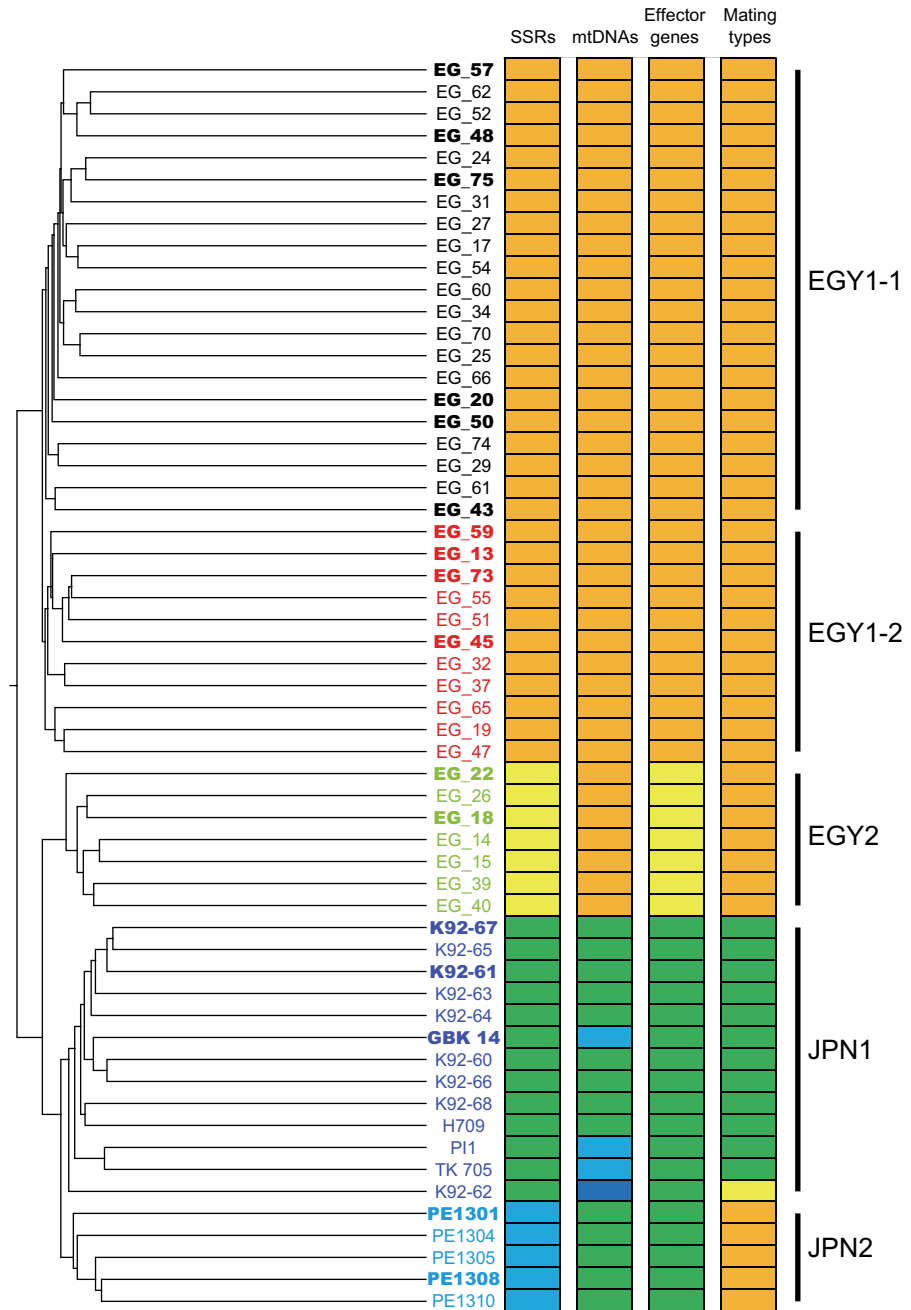


Fig 1. Clustering analysis of *Phytophthora infestans* isolates collected from Egypt and Japan using double-digest restriction-site associated DNA sequencing (ddRAD-Seq) data. Clusters (EGY1-1, EGY1-2, EGY2, JPN1, and JPN2) are indicated with solid bars. Clusters determined using simple sequence repeats (SSRs), effector genes, mitochondrial DNA (mtDNA) polymorphisms, and mating types are shown as colored charts; same colors indicate the members of the same clusters. Isolates indicated in bold were used to perform whole genome resequencing (WGRS) analysis.

<https://doi.org/10.1371/journal.pone.0221604.g001>

Discovery of genome-wide single nucleotide polymorphisms (SNPs) using ddRAD-Seq

The ddRAD-Seq approach was used to analyze genome-wide SNPs. A total of 23 Egyptian isolates were removed from subsequent analysis because of low-quality data, probably because

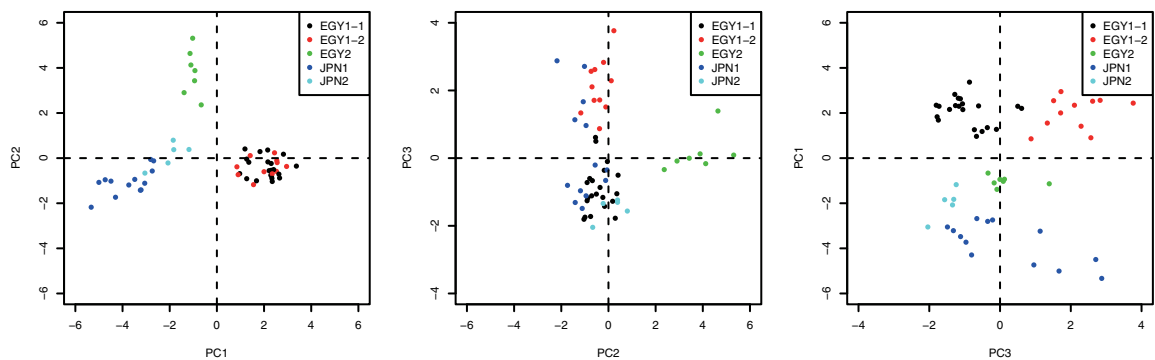


Fig 2. Principal component analysis (PCA) of *P. infestans* isolates. Different colors indicate the five clusters, EGY1-1, EGY1-2, EGY2, JPN1, and JPN2, identified by clustering analysis.

<https://doi.org/10.1371/journal.pone.0221604.g002>

the DNA available for analysis was of poor quality and insufficient quantity. An average of 430,489 high-quality ddRAD-Seq reads were obtained per sample from 57 isolates (S3 Table). Reads were mapped onto the reference genome sequence of the *P. infestans* strain T30-4 (version ASM14294v1) at an average mapping rate of 92.4%, and 996 loci were detected as high-confidence SNPs based on the read alignments.

Clustering analysis, based on SNPs, grouped the tested isolates into four clusters, EGY1, EGY2, JPN1, and JPN2, containing 32, 7, 13, and 5 isolates, respectively, in accordance with the results of the SSR marker analysis (Fig 1). Furthermore, the EGY1 cluster was divided into two sub-clusters, EGY1-1 and EGY1-2, containing 21 and 11 isolates, respectively. Principal component analysis (PCA) supported this result, and PC1, PC2, and PC3 axes explained 11.2%, 4.2%, and 3.8% of the variance, respectively (Fig 2).

Whole genome SNP and copy number variation (CNV) analysis

Seventeen isolates representative of each cluster, based on ddRAD-Seq data and sampling location, host, and year, were selected for WGRS analysis. An average of 7.9 Gb sequence data per sample (34.6× coverage) was obtained (S4 Table). The reads were mapped onto the reference genome sequence of *P. infestans* at an average mapping rate of 94.5%, and 760,928 high-confidence SNPs, including 479,784 transitions and 281,144 transversions (Ts/Tv ratio = 1.7), were detected. The Egyptian isolates EG_18 and EG_73 contained the lowest and highest number of SNPs, respectively. The average density of SNPs was 3.3 per kb. Analysis of SNP annotations using SnpEff indicated that intergenic variants were the most common (606,900; 79.8%), followed by synonymous variants (57,024; 7.4%) and missense variants (56,758; 7.4%) (S5 Table). A total of 1,267 (0.6%) high- and 56,758 (7.4%) moderate-impact SNPs were identified in 1,085 and 11,289 genes, respectively.

Among the Egyptian isolates, CNVs were detected between EG_18, as a reference, and each of the remaining 16 isolates. Between any two isolates, an average of 5,402 CNVs were detected in each isolate, ranging from 533 CNVs in EG_22 to 6,735 CNVs in K92-61. The average size of a CNV was 4,068 bp, with the longest of 79 kb in EG_75.

To determine the lineage of *P. infestans* isolates, we used the SNP data and data publicly available on 11 isolates [19]. On the basis of 90,7671 SNPs detected in 28 isolates including the 17 isolates from this study and 11 from the publica data, four clusters were generated (Fig 3). Two Egyptian isolates belonging to the EGY2 cluster, EG_18 and EG_22, were grouped with 06_3928A, an isolate collected from England. Two clusters included Japanese and Egyptian

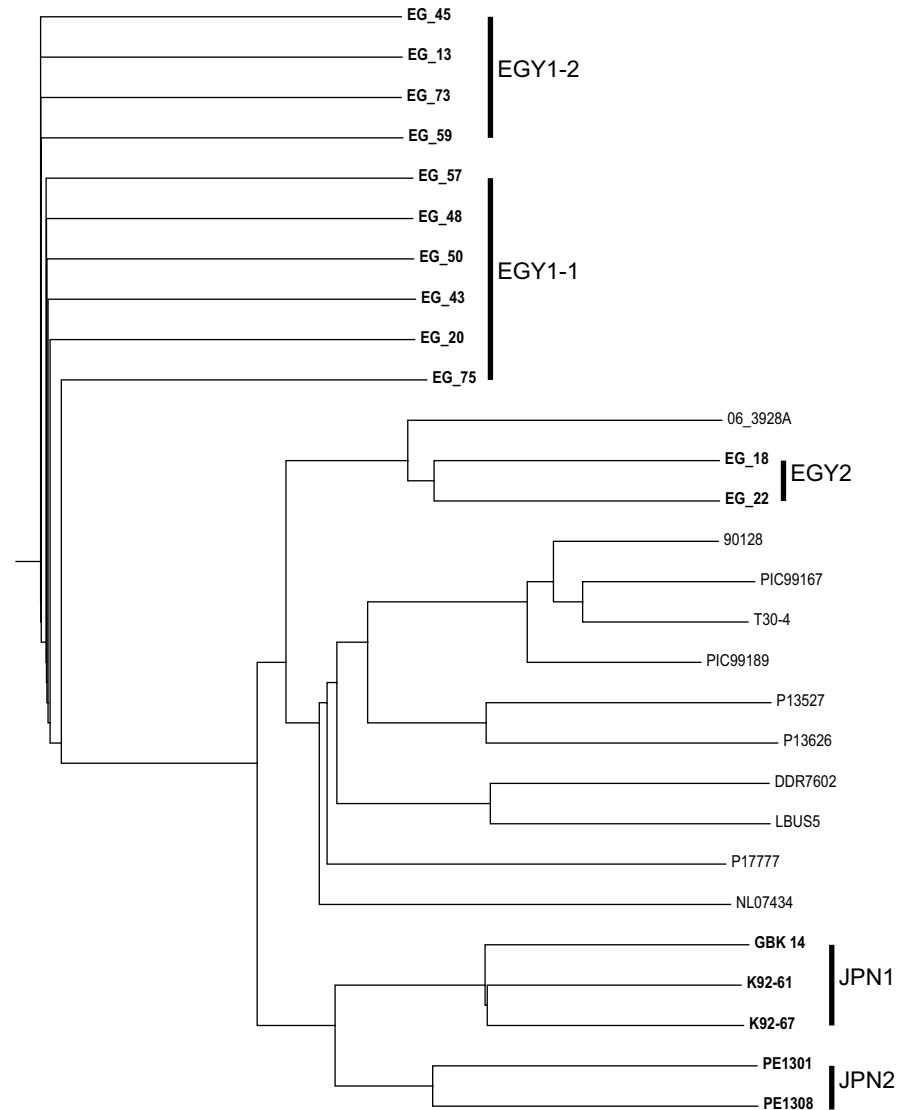


Fig 3. WGRS-based clusters of *P. infestans* isolates. *P. infestans* isolates sequenced in this study are indicated in bold. The remaining isolates represent those whose genome sequences were publicly available.

<https://doi.org/10.1371/journal.pone.0221604.g003>

isolates belonging to JPN1, JPN2, and EGY1 clusters. The fourth cluster consisted of ten isolates, all of which have been reported previously as members of US-1 and HERB-1.

Polymorphisms in effector genes

In the *P. infestans* T30-4 reference genome sequence, 594 and 194 genes have been annotated as encoding RxLR- and Crinkler-type effector proteins, respectively. Among these, four genes were identified in the *P. infestans* T30-4 genome sequence: *Avr1* (PITG_16663) in the contig sequence of supercont1.51, *Avr2* (PITG_08943 and PITG_22870) on supercont1.16, *Avr3a* (PITG_14371) on supercont1.34, and *Avr4* (PITG_07387) on supercont1.11. A total of six non-synonymous SNPs causing missense mutations were identified. Four of the six missense mutations caused amino acid substitutions at positions 19, 80, 103, and 124 of PITG_14371 (*Avr3a*). While the first three mutations were observed in the ten isolates belonging to clusters

EGY1-1 and EGY1-2, the fourth mutation was identified in all 12 isolates belonging to the EGY2 cluster, in addition to the isolates in EGY1-1 and EGY1-2 clusters. Among the two remaining missense mutations, one was detected at amino acid position 31 of PITG_08943 (Avr2) in the two isolates belonging to the JPN2 cluster and one isolate (GBK 14) belonging to JPN1, and the other was detected at amino acid position 139 of PITG_07387 (Avr4) in seven isolates belonging to clusters EGY2, JPN1, and JPN2. No SNPs were detected in PITG_16663 (Avr1).

CNVs were detected in *Avr1*, *Avr2*, *Avr3a*, and *Avr4*. CNVs of *Avr1* and *Avr4* were observed in 12 isolates of EGY1-1, EGY1-2, and JPN2, while CNVs of *Avr2* were detected in three isolates belonging to the cluster JPN1 and 12 isolates belonging to EGY1-1, EGY1-2, and JPN2. The CNV of *Avr3a* was identified in only one isolate, EG_75, belonging to the sub-cluster EGY1-1.

Discussion

In this study, we analyzed a panel of 80 *P. infestans* isolates collected from Egypt and Japan using a series of thorough genotyping assays at different levels, including the target gene level (SSRs and RFLPs), genome-wide level (ddRAD-Seq and WGRS), and reproduction biology level (mating type assay). While A1 and A2 mating types were observed among Japanese isolates, the A1 mating type was predominant among Egyptian isolates tested. This suggests that the A1 mating type might be adapted to the Egyptian environment [24], resulting in the absence of A2 mating types needed for sexual reproduction to produce oospores. Nevertheless, Egyptian isolates are more divergent than the Japanese isolates [24], which was supported by this study. The unstable nature of *P. infestans* genome may be one of the possible drivers of genetic diversity [21]. In this study, we observed SNPs and CNVs in the genomes of isolates. Some genomic variations may alter the host specificity and geographical location of isolates [17, 25]. According to a recent report [26], CNVs are associated with the emergence of super-fit clones of *P. infestans*. However, effector genes do not diverge among the Egyptian isolates, possibly because most of the tomato and potato cultivars in Egypt are susceptible to late blight, leading to minimal selection pressure and low mutation rates. Nonetheless, the population structure of *P. infestans* isolates, regardless of their reproduction mode, was broader than that suggested by previous studies on American and European isolates (Fig 3), which was able to directly compare the isolates from other studies [19, 27–29]. The ddRAD-Seq analysis also provides information on population structure [30–32]; however, the ability to compare data depends on the choice of restriction enzymes used for library preparation. The analysis of SSRs, effector genes, and mtDNA was insufficient to reveal the population structure of *P. infestans* isolates (Fig 1 and S1 Table). Besides, for gel-based methods, it would be difficult to detect small sequence variations inside amplicons, e.g., base substitutions and small deletions.

Sequence variations in the *P. infestans* genome might contribute to its wide host range. In this study, while two isolates in the EGY2 cluster were members of 06_3928A from England [27], isolates in JPN1, JPN2, and EGY1 clusters were not represented in the major cluster comprising ten isolates analyzed in previous studies (Fig 1). Therefore, it is important to identify and classify the isolates correctly for successful plant disease management. Additionally, further investigation of effector genes is needed to explain the virulence profile of *P. infestans* isolates. Our results suggest that gene-based analyses are insufficient for isolate identification, and genotyping the isolates using genome-wide analyses, such as ddRAD-Seq and WGRS, is more powerful for the classification of new *P. infestans* isolates together with the previously reported isolates. Moreover, since sequence data could be compared with others, the ddRAD-Seq and WGRS reads from this study would be useful for future genetic and genomic studies involving broad *P. infestans* collections.

The number of isolates with a sequenced genome is limited; however, genome-wide sequence data of different isolates from this study would provide more detailed phylogenetic information, which would help control isolates that could soon become endemic. As proposed by Arafa [13], genomics-based characterization of *P. infestans* will enhance the monitoring of aggressive isolates propagated via infected plant materials from different geographic regions. Breeding late blight resistant tomato and potato cultivars is also a promising strategy to overcome this disease [33, 34]. Surveillance systems such as Asiablight, Euroblight, and USAblight have played important roles in monitoring the late blight disease [35]; however, similar programs are unavailable for other regions such as Africa. The plants were in fact completely damaged once infected by new isolates [36]. Overall, a short-term combat protocol based on prediction, diagnostics, and therapeutics is needed for sustainable farming and food production, and a long-term approach based on genetics, genomics, and plant breeding is needed for the effective control of late blight.

Materials and methods

P. infestans isolates and DNA extraction

A total of 62 isolates were sampled from the leaves, stems, and fruits of tomato and potato plants growing in 16 geographical locations in Egypt during two growing seasons (2014–2015 and 2015–2016) (S1 Table). Additionally, 18 isolates were obtained from the NIAS Genebank, Japan (<http://www.gene.affrc.go.jp>); these isolates were collected from the leaves and fruits of tomato plants and leaves of potato plants growing in three prefectures, Hokkaido, Ibaraki, and Shizuoka, from 1987–2013 (S1 Table). *P. infestans* isolates were extracted from the plants as described previously [33]. The DNA of isolates was extracted using the GeneJET Plant Genomic DNA Purification Mini Kit (Thermo Fisher Scientific) or DNeasy Plant Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions.

Mating type assay

Mycelium plugs of *P. infestans* isolates and those of EG_136 [24] and K92-65 (NIAS Genebank) as standard isolates of A1 or A2 mating type, respectively, were placed on rye agar medium in a Petri dish at a spacing of 3 cm. All samples were incubated at 18°C in the dark for 3 weeks. The interaction domain of oospores between two isolates was investigated under a microscope at 20× magnification. Isolates that produced oospores (sexual spores) of the A2 mating type were identified as A1, while isolates that produced oospores of the A1 mating type were identified as A2 [37].

Analysis of mtDNA and effector genes

The mtDNA of isolates was amplified from the genomic DNA using primers listed in S2 Table. The thermal cycling conditions used for PCR were based on previous studies [38, 39]. Additionally, the presence/absence of five effector genes, including *AVR1*, *AVR2*, *AVR2-like*, *AVR3a*, and *AVR4*, in *P. infestans* isolates was tested in the current populations (S2 Table). The PCR products were separated by electrophoresis on 2% agarose gels and visualized using a UV gel documentation system (Biocraft, Co. Japan) after staining with ethidium bromide.

SSR marker analysis

Sixteen SSR markers [40–42] were used to genotype *P. infestans* isolates using sequence-specific primers (S2 Table). DNA amplification by PCR and capillary electrophoresis using ABI3730 DNA analyzer (Applied Biosystems) were performed as described previously [24].

GeneMarker (SoftGenetics) was used to score the genotype. A phylogenetic tree was constructed based on the SSR data set using graphical genotypes (GGT) 2.0 software [43].

Analysis of ddRAD-Seq data

Genomic DNA of isolates was double-digested with *Pst*I and *Msp*I restriction enzymes and used to construct ddRAD-Seq libraries. The libraries were sequenced using MiSeq (Illumina) in paired-end 251 bp mode, as described previously [33, 44].

Data processing was performed as described previously [33, 44]. Low-quality sequences and adapters (S2 and S3 Tables) were removed from the raw sequence reads using PRINSEQ (version 0.20.4) [45] and fastx_clipper (FASTX Toolkit version 0.0.13; http://hannonlab.cshl.edu/fastx_toolkit), respectively. The cleaned reads were mapped onto the reference genome sequence of *P. infestans* strain T30-4 (ASM14294v1; accession number AATU00000000) [21] using Bowtie 2 (version 2.2.3) [46]. To obtain a variant call format (VCF) file containing SNP information, sequence alignment/map format (SAM) files were converted to binary sequence alignment/map format (BAM) files and subjected to SNP calling using the mpileup command of SAMtools (version 0.1.19; [47] and the view command of BCFtools [47]. High-confidence biallelic SNPs were selected using VCFtools (version 0.1.12b) [48], based on the following criteria: 1) $\geq 10\times$ coverage in each isolate, 2) > 999 SNP quality value, 3) ≥ 0.1 minor allele frequency, and 4) < 0.5 missing data rate.

Analysis of WGRS data

Sequencing libraries with an insert size of 600 bp were prepared for 17 isolates of *P. infestans*, as described previously [33, 49], and sequenced using Illumina NextSeq 500 in paired-end 151 bp mode. Data processing was conducted as described above for ddRAD-Seq data, with the exception of SNP filtering criteria, which were as follows: 1) $\geq 5\times$ coverage in each isolate, 2) > 999 SNP quality value, 3) ≥ 0.1 minor allele frequency, and 4) < 0.2 missing data rate. Annotations of SNP effects on gene functions were predicted using SnpEff (version 4.2) [50]. CNVs were detected using CNV-seq (version 0.2.7) [51].

Clustering analysis

Dendrograms were constructed using the unweighted pair-group method with average linkage (UPGMA). Additionally, PCA was performed to determine the relationship among samples using the TASSEL software [52], with the number of components limited to four.

Accession numbers: Double digest restriction-site associated DNA sequencing (ddRAD-Seq) and whole genome resequencing (WGRS) data are available at the DDBJ Sequence Read Archive database under the accession number DRA007610.

Supporting information

S1 Table. *Phytophthora infestans* isolates used in this study.
(XLSX)

S2 Table. List of primer sequences used in this study.
(XLSX)

S3 Table. Number of ddRAD-Seq reads and rate of mapping against the reference genome sequence of *P. infestans*.
(XLSX)

S4 Table. Number of WGRS reads and rate of mapping against the reference genome sequence of *P. infestans*.

(XLSX)

S5 Table. Numbers of annotated SNPs and indels classified using the SnpEff program.

(XLSX)

Acknowledgments

We thank the NIAS Genebank, Japan, for providing the *P. infestans* isolates. We also thank S. Sasamoto, C. Mimani, H. Tsuruoka, and A. Obara at the Kazusa DNA Research Institute for technical assistance.

Author Contributions

Conceptualization: Ramadan A. Arafa, Kenta Shirasawa.

Data curation: Ramadan A. Arafa, Kenta Shirasawa.

Formal analysis: Ramadan A. Arafa, Kenta Shirasawa.

Funding acquisition: Ramadan A. Arafa, Kenta Shirasawa.

Investigation: Ramadan A. Arafa, Said M. Kamel.

Methodology: Ramadan A. Arafa.

Project administration: Kenta Shirasawa.

Resources: Ramadan A. Arafa, Said M. Kamel, Mohamed T. Rakha, Nour Elden K. Soliman, Kenta Shirasawa.

Supervision: Mohamed T. Rakha, Olfat M. Moussa, Kenta Shirasawa.

Validation: Kenta Shirasawa.

Visualization: Ramadan A. Arafa.

Writing – original draft: Ramadan A. Arafa, Kenta Shirasawa.

Writing – review & editing: Mohamed T. Rakha.

References

1. Martin MD, Vieira FG, Ho SY, Wales N, Schubert M, Seguin-Orlando A, et al. Genomic Characterization of a South American *Phytophthora* Hybrid Mandates Reassessment of the Geographic Origins of *Phytophthora infestans*. *Mol Biol Evol*. 2016; 33(2):478–91. Epub 2015/11/19. <https://doi.org/10.1093/molbev/msv241> PMID: 26576850; PubMed Central PMCID: PMC4866541.
2. Gomez-Alpizar L, Carbone I, Ristaino JB. An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. *Proc Natl Acad Sci U S A*. 2007; 104(9):3306–11. Epub 2007/03/16. <https://doi.org/10.1073/pnas.0611479104> PMID: 17360643; PubMed Central PMCID: PMC1805513.
3. Goss EM, Tabima JF, Cooke DE, Restrepo S, Fry WE, Forbes GA, et al. The Irish potato famine pathogen *Phytophthora infestans* originated in central Mexico rather than the Andes. *Proc Natl Acad Sci U S A*. 2014; 111(24):8791–6. Epub 2014/06/04. <https://doi.org/10.1073/pnas.1401884111> PMID: 24889615; PubMed Central PMCID: PMC4066499.
4. Wang S, Boevink PC, Welsh L, Zhang R, Whisson SC, Birch PRJ. Delivery of cytoplasmic and apoplastic effectors from *Phytophthora infestans* haustoria by distinct secretion pathways. *New Phytol*. 2017; 216(1):205–15. Epub 2017/08/02. <https://doi.org/10.1111/nph.14696> PMID: 28758684; PubMed Central PMCID: PMC5601276.

5. Avila-Adame C, Gomez-Alpizar L, Zismann V, Jones KM, Buell CR, Ristaino JB. Mitochondrial genome sequences and molecular evolution of the Irish potato famine pathogen, *Phytophthora infestans*. *Curr Genet*. 2006; 49(1):39–46. Epub 2005/12/06. <https://doi.org/10.1007/s00294-005-0016-3> PMID: 16328503.
6. Haverkort AJ, Struik PC, Visser RGF, Jacobsen E. Applied Biotechnology to Combat Late Blight in Potato Caused by *Phytophthora infestans*. *Potato Research*. 2009; 52(3):249–64. <https://doi.org/10.1007/s11540-009-9136-3>
7. Turkensteen LJ, Flier WG, Wannings R, Mulder A. Production, survival and infectivity of oospores of *Phytophthora infestans*. *Plant Pathology*. 2000; 49(6):688–96. <https://doi.org/10.1046/j.1365-3059.2000.00515.x>
8. Goodwin S, Fry W. Global migration of *Phytophthora infestans*. *Phytopathology*. 1991; 82:955–61.
9. Hohl HR, Iselin K. Strains of *Phytophthora infestans* from Switzerland with A2 mating type behaviour. *Transactions of the British Mycological Society*. 1984; 83(3):529–30. [https://doi.org/10.1016/s0007-1536\(84\)80057-1](https://doi.org/10.1016/s0007-1536(84)80057-1)
10. Kato M, Mizubuti ES, Goodwin SB, Fry WE. Sensitivity to Protectant Fungicides and Pathogenic Fitness of Clonal Lineages of *Phytophthora infestans* in the United States. *Phytopathology*. 1997; 87(9):973–8. Epub 2008/10/24. <https://doi.org/10.1094/PHTO.1997.87.9.973> PMID: 18945070.
11. Day JP, Wattier RAM, Shaw DS, Shattock RC. Phenotypic and genotypic diversity in *Phytophthora infestans* on potato in Great Britain, 1995–98. *Plant Pathology*. 2004; 53(3):303–15. <https://doi.org/10.1111/j.0032-0862.2004.01004.x>
12. Fry WE, Birch PR, Judelson HS, Grunwald NJ, Danies G, Everts KL, et al. Five Reasons to Consider *Phytophthora infestans* a Reemerging Pathogen. 2015; 105(7):966–81. <https://doi.org/10.1094/PHTO-01-15-0005-FI> PMID: 25760519.
13. Arafa RA, Shirasawa K. Technical review of molecular markers and next-generation sequencing technology to manage plant pathogenic oomycetes. *African Journal of Biotechnology*. 2018; 17(12):369–79. <https://doi.org/10.5897/ajb2017.16304>
14. Arafa RA, Moussa OM, Soliman NEK, Shirasawa K, Kamel SM, Rakha MT. Resistance to *Phytophthora infestans* in tomato wild relatives. *African Journal of Agricultural Research*. 2017; 12(26):2188–96. <https://doi.org/10.5897/ajar2017.12422>
15. Foolad MR, Merk HL, Ashrafi H. Genetics, Genomics and Breeding of Late Blight and Early Blight Resistance in Tomato. *Critical Reviews in Plant Sciences*. 2008; 27(2):75–107. <https://doi.org/10.1080/07352680802147353>
16. Chowdappa P, Nirmal Kumar BJ, Madhura S, Mohan Kumar SP, Myers KL, Fry WE, et al. Severe outbreaks of late blight on potato and tomato in South India caused by recent changes in the *Phytophthora infestans* population. *Plant Pathology*. 2015; 64(1):191–9. <https://doi.org/10.1111/ppa.12228>
17. Hansen ZR, Knaus BJ, Tabima JF, Press CM, Judelson HS, Grunwald NJ, et al. Loop-mediated isothermal amplification for detection of the tomato and potato late blight pathogen, *Phytophthora infestans*. *J Appl Microbiol*. 2016; 120(4):1010–20. Epub 2016/01/29. <https://doi.org/10.1111/jam.13079> PMID: 26820117.
18. Rekad FZ, Cooke DEL, Puglisi I, Randall E, Guenaoui Y, Bouznad Z, et al. Characterization of *Phytophthora infestans* populations in northwestern Algeria during 2008–2014. *Fungal Biol*. 2017; 121(5):467–77. Epub 2017/04/10. <https://doi.org/10.1016/j.funbio.2017.01.004> PMID: 28390504.
19. Yoshida K, Schuenemann VJ, Cano LM, Pais M, Mishra B, Sharma R, et al. The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. *ELife*. 2013; 2:e00731. Epub 2013/06/07. <https://doi.org/10.7554/eLife.00731> PMID: 23741619; PubMed Central PMCID: PMC3667578.
20. Gilroy EM, Breen S, Whisson SC, Squires J, Hein I, Kaczmarek M, et al. Presence/absence, differential expression and sequence polymorphisms between PiAVR2 and PiAVR2-like in *Phytophthora infestans* determine virulence on R2 plants. *New Phytol*. 2011; 191(3):763–76. Epub 2011/05/05. <https://doi.org/10.1111/j.1469-8137.2011.03736.x> PMID: 21539575.
21. Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, Cano LM, et al. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*. 2009; 461(7262):393–8. Epub 2009/09/11. <https://doi.org/10.1038/nature08358> PMID: 19741609.
22. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*. 2011; 6(5):e19379. Epub 2011/05/17. <https://doi.org/10.1371/journal.pone.0019379> PMID: 21573248; PubMed Central PMCID: PMC3087801.
23. Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet*. 2011; 12(7):499–510. Epub 2011/06/18. <https://doi.org/10.1038/nrg3012> PMID: 21681211.

24. Arafa RA, Soliman NEK, Moussa OM, Kamel SM, Shirasawa K. Characterization of Egyptian *Phytophthora infestans* population using simple sequence repeat markers. *Journal of General Plant Pathology*. 2018; 84(2):104–7. <https://doi.org/10.1007/s10327-018-0763-x>
25. Li Y, van der Lee TA, Evenhuis A, van den Bosch GB, van Bekkum PJ, Forch MG, et al. Population dynamics of *Phytophthora infestans* in the Netherlands reveals expansion and spread of dominant clonal lineages and virulence in sexual offspring. *G3 (Bethesda)*. 2012; 2(12):1529–40. Epub 2013/01/01. <https://doi.org/10.1534/g3.112.004150> PMID: 23275876; PubMed Central PMCID: PMC3516475.
26. Knaus BJ, Tabima JF, Shakya SK, Judelson HS, Grünwald NJ. 2019. <https://doi.org/10.1101/633701>
27. Cooke DE, Cano LM, Raffaele S, Bain RA, Cooke LR, Etherington GJ, et al. Genome analyses of an aggressive and invasive lineage of the Irish potato famine pathogen. *PLoS Pathog*. 2012; 8(10): e1002940. Epub 2012/10/12. <https://doi.org/10.1371/journal.ppat.1002940> PMID: 23055926; PubMed Central PMCID: PMC3464212.
28. Raffaele S, Farrer RA, Cano LM, Studholme DJ, MacLean D, Thines M, et al. Genome evolution following host jumps in the Irish potato famine pathogen lineage. *Science*. 2010; 330(6010):1540–3. Epub 2010/12/15. <https://doi.org/10.1126/science.1193070> PMID: 21148391.
29. Raffaele S, Win J, Cano LM, Kamoun S. Analyses of genome architecture and gene expression reveal novel candidate virulence factors in the secretome of *Phytophthora infestans*. *BMC Genomics*. 2010; 11:637. Epub 2010/11/18. <https://doi.org/10.1186/1471-2164-11-637> PMID: 21080964; PubMed Central PMCID: PMC3091767.
30. Kim MS, Hohenlohe PA, Kim KH, Seo ST, Klopfenstein NB. Genetic diversity and population structure of *Raffaelea quercus-mongolicae*, a fungus associated with oak mortality in South Korea. *For Pathol*. 2016; 46(2):164–7. Epub 2016/04/19. <https://doi.org/10.1111/efp.12263> PMID: 27087782; PubMed Central PMCID: PMC4827447.
31. Grewe F, Huang JP, Leavitt SD, Lumbsch HT. Reference-based RADseq resolves robust relationships among closely related species of lichen-forming fungi using metagenomic DNA. *Sci Rep*. 2017; 7(1):9884. Epub 2017/08/31. <https://doi.org/10.1038/s41598-017-09906-7> PMID: 28852019; PubMed Central PMCID: PMC5575168.
32. Grewe F, Lagostina E, Wu H, Printzen C, Lumbsch HT. Population genomic analyses of RAD sequences resolves the phylogenetic relationship of the lichen-forming fungal species *Usnea antarctica* and *Usnea aurantiacoatra*. *MycKeys*. 2018;(43):91–113. Epub 2018/12/28. <https://doi.org/10.3897/mycokeys.43.29093> PMID: 30588165; PubMed Central PMCID: PMC6300515.
33. Arafa RA, Rakha MT, Soliman NEK, Moussa OM, Kamel SM, Shirasawa K. Rapid identification of candidate genes for resistance to tomato late blight disease using next-generation sequencing technologies. *PLoS One*. 2017; 12(12):e0189951. Epub 2017/12/19. <https://doi.org/10.1371/journal.pone.0189951> PMID: 29253902; PubMed Central PMCID: PMC5734779.
34. Yang L, Wang D, Xu Y, Zhao H, Wang L, Cao X, et al. A New Resistance Gene against Potato Late Blight Originating from *Solanum pinnatisectum* Located on Potato Chromosome 7. *Front Plant Sci*. 2017; 8:1729. Epub 2017/11/01. <https://doi.org/10.3389/fpls.2017.01729> PMID: 29085380; PubMed Central PMCID: PMC5649132.
35. Cooke L, Zhang R, Wu L, Chen S, Forbes G. Recent developments: late blight in Asia AsiaBlight. EuroBlight workshop. 2017:14–7.
36. Cooke LR, Schepers HTAM, Hermansen A, Bain RA, Bradshaw NJ, Ritchie F, et al. Epidemiology and Integrated Control of Potato Late Blight in Europe. *Potato Research*. 2011; 54(2):183–222. <https://doi.org/10.1007/s11540-011-9187-0>
37. Runno-Paurson E, Fry WE, Myers KL, Koppel M, Mänd M. Characterisation of *Phytophthora infestans* isolates collected from potato in Estonia during 2002–2003. *European Journal of Plant Pathology*. 2009; 124(4):565–75. <https://doi.org/10.1007/s10658-009-9442-2>
38. Griffith GW, Shaw DS. Polymorphisms in *phytophthora infestans*: four mitochondrial haplotypes are detected after PCR amplification of DNA from pure cultures or from host lesions. *Appl Environ Microbiol*. 1998; 64(10):4007–14. Epub 1998/10/06. PMID: 9758833; PubMed Central PMCID: PMC106592.
39. Yang X. *Phytophthora mississippiensis* sp. nov., a New Species Recovered from Irrigation Reservoirs at a Plant Nursery in Mississippi. *Journal of Plant Pathology & Microbiology*. 2013; 04(06). <https://doi.org/10.4172/2157-7471.1000180>
40. Knapova G, Schlenzig A, Gisi U. Crosses between isolates of *Phytophthora infestans* from potato and tomato and characterization of F1 and F2 progeny for phenotypic and molecular markers. *Plant Pathology*. 2002; 51(6):698–709. <https://doi.org/10.1046/j.1365-3059.2002.00762.x>
41. Lees AK, Wattier R, Shaw DS, Sullivan L, Williams NA, Cooke DEL. Novel microsatellite markers for the analysis of *Phytophthora infestans* populations. *Plant Pathology*. 2006; 55(3):311–9. <https://doi.org/10.1111/j.1365-3059.2006.01359.x>

42. Li Y, Cooke DE, Jacobsen E, van der Lee T. Efficient multiplex simple sequence repeat genotyping of the oomycete plant pathogen *Phytophthora infestans*. *J Microbiol Methods*. 2013; 92(3):316–22. Epub 2013/01/15. <https://doi.org/10.1016/j.mimet.2012.11.021> PMID: 23313554.
43. van Berloo R. GGT 2.0: versatile software for visualization and analysis of genetic data. *J Hered*. 2008; 99(2):232–6. Epub 2008/01/29. <https://doi.org/10.1093/jhered/esm109> PMID: 18222930.
44. Shirasawa K, Hirakawa H, Isobe S. Analytical workflow of double-digest restriction site-associated DNA sequencing based on empirical and in silico optimization in tomato. *DNA Res*. 2016; 23(2):145–53. Epub 2016/03/05. <https://doi.org/10.1093/dnares/dsw004> PMID: 26932983; PubMed Central PMCID: PMC4833422.
45. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011; 27(6):863–4. Epub 2011/02/01. <https://doi.org/10.1093/bioinformatics/btr026> PMID: 21278185; PubMed Central PMCID: PMC3051327.
46. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9(4):357–9. Epub 2012/03/06. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286; PubMed Central PMCID: PMC3322381.
47. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25(16):2078–9. Epub 2009/06/10. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943; PubMed Central PMCID: PMC2723002.
48. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011; 27(15):2156–8. <https://doi.org/10.1093/bioinformatics/btr330> PMID: 21653522; PubMed Central PMCID: PMC3137218.
49. Shirasawa K, Hirakawa H, Nunome T, Tabata S, Isobe S. Genome-wide survey of artificial mutations induced by ethyl methanesulfonate and gamma rays in tomato. *Plant Biotechnol J*. 2016; 14(1):51–60. Epub 2015/02/18. <https://doi.org/10.1111/pbi.12348> PMID: 25689669; PubMed Central PMCID: PMC5023996.
50. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012; 6(2):80–92. Epub 2012/06/26. <https://doi.org/10.4161/fly.19695> PMID: 22728672; PubMed Central PMCID: PMC3679285.
51. Xie C, Tammi MT. CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics*. 2009; 10:80. Epub 2009/03/10. <https://doi.org/10.1186/1471-2105-10-80> PMID: 19267900; PubMed Central PMCID: PMC2667514.
52. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. 2007; 23(19):2633–5. Epub 2007/06/26. <https://doi.org/10.1093/bioinformatics/btm308> PMID: 17586829.