Article

Copy Number Variations of Glycoside Hydrolase 45 Genes in *Bursaphelenchus xylophilus* and Their Impact on the Pathogenesis of Pine Wilt Disease

Xiaolei Ding ^{1,2,3}, Qingtong Wang ^{1,2}, Yunfei Guo ³, Yulong Li ^{1,2}, Sixi Lin ^{1,2}, Qingwei Zeng ⁴, Feijian Sun ^{1,2}, De-Wei Li ⁵ and Jianren Ye ^{1,2,*}

- ¹ Co-Innovation Center for Sustainable Forestry in Southern China, College of Forestry, Nanjing Forestry University, Nanjing 210037, China; dxl@njfu.edu.cn (X.D.); qtwang@njfu.edu.cn (Q.W.); yulongli@njfu.edu.cn (Y.L.); lsx@njfu.edu.cn (S.L.); feijian_sun@sina.com (F.S.)
- ² Jiangsu Key Laboratory for Prevention and Management of Invasive Species, Nanjing 210037, China
- ³ Zilkha Neurogenetic Institute, University of Southern California, Los Angeles, CA 90089, USA; guoyunfei1989@gmail.com
- ⁴ School of Life Science and Food Engineering, Huaiyin Institute of Technology, Huaian 223003, China; qwzeng@hyit.edu.cn
- ⁵ The Connecticut Agricultural Experiment Station Valley Laboratory, Windsor, CT 06095, USA; Dewei.Li@ct.gov
- * Correspondence: njfu_jrye@163.com

Citation: Ding, X.; Wang, Q.; Guo, Y.; Li, Y.; Lin, S.; Zeng, Q.; Sun, F.; Li, D.-W.; Ye, J. Copy Number Variations of Glycoside Hydrolase 45 Genes in *Bursaphelenchus xylophilus* and Their Impact on the Pathogenesis of Pine Wilt Disease. *Forests* **2021**, *12*, 275. <http://doi.org/10.3390/f12030275>

Academic Editor: Paulo Zaini

Received: 20 January 2021

Accepted: 8 February 2021

Published: 28 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The pine wood nematode *Bursaphelenchus xylophilus* parasitizes millions of pine trees worldwide each year, causing severe wilt and the death of host trees. Glycoside hydrolase 45 genes of *B. xylophilus* are reported to have been acquired by horizontal gene transfer from fungi and are responsible for cell wall degradation during nematode infection. Previous studies ignored the possibility of copy number variations of such genes. In this study, we determined that two of the glycoside hydrolase 45 genes evolved to maintain multiple copies with distinct expression levels, enabling the nematode to infect a variety of pine hosts. Additionally, tandem repeat variations within coding regions were also detected between different copies of glycoside hydrolase 45 genes that could result in changes in protein sequences and serve as an effective biological marker to detect copy number variations among different *B. xylophilus* populations. Consequently, we were able to further identify the copy number variations of glycoside hydrolase 45 genes among *B. xylophilus* strains with different virulence. Our results provide new insights into the pathogenicity of *B. xylophilus*, provide a practical marker to genotype copy number variations and may aid in population classification.

Keywords: *Bursaphelenchus xylophilus*; glycoside hydrolase 45; multi-copy gene; copy number variation

1. Introduction

The pine wood nematode (PWN) *Bursaphelenchus xylophilus* (Steiner & Buhrer) Nickle is an invasive species and the causative agent of pine wilt disease, which originated in North America [1]. *B. xylophilus* has not caused disease in the local pine species in North America but it poses a major threat to all pine forests in Asia and Europe [2]. In China, the infected area has reached over 660 counties in 18 provinces according to the State Forestry and Grassland Administration of China. Researchers have been struggling to elucidate the pathogenic mechanism of *B. xylophilus* in order to efficiently manage this parasite [3–6]. To better understand this nematode, the whole genome of *B. xylophilus* was published in 2011 [7]. Since the announcement of the first short-read sequencing-based genome assembly, high-throughput sequencing has frequently been used to identify pathogenic traits of *B. xylophilus*. This effort focused on genes, noncoding RNAs and genomic variations, which expanded our knowledge of its molecular pathogenesis [6,8–12]. However, most

of these studies usually focused on analyzing genes with potential regulatory roles. In contrast, only a few reports have shed light on *B. xylophilus* multi-copy genes [9,13,14], which have been suggested to be associated with diseases in humans and other model organisms [15]. Glycoside hydrolase 45 (GH45) gene families are known to be *B. xylophilus*-specific and have been shown to have been acquired by horizontal gene transfer (HGT) and to be associated with cell-wall-degradation processes during infection [7,13]. However, there is little information available about the genome architectures and copy number variations of certain GH45 genes. Consequently, research is needed to investigate the copy numbers of these important genes and to gain novel insights into the pathogenesis of *B. xylophilus*.

Multi-copy genes are usually found within repeated regions, such as segment duplications (SDs) and tandem repeats (TRs) [16]. Many genomes of simple organisms contain similar sequence repeats ranging in size from TRs to SDs [17]. SDs are usually considered to be highly similar DNA fragments (>90% identity) between 1 and 400 kb, which occur more than once in the genome. They are known to be gene regulators and are responsible for gene functional diversification, gene family expansion and the adaption of the species to certain environments [18]. TRs are unstable genomic elements that are repeated head to tail and are usually prone to variation [19]. They have been used as genetic markers due to their unique polymorphic nature. Extensive studies have also revealed their potential role in the regulation of gene products associated with pathogenicity variation in microbes and in polygenic disorders in humans [15,20,21]. Despite their crucial biological importance, the accurate identification and profiling of multi-copy genes is difficult due to their highly similar sequence content [22].

To fill this knowledge gap, we set out to document the existence and copy number variations of GH45 in *B. xylophilus* based on our recently generated long-read genome assembly (AH1, PRJNA524063, manuscript under review). More importantly, the expression patterns of different GH45 copies were quantified during the pathogenesis process of pine wilt disease. Our results provide a new perspective for these pathogenesis-related genes that will inform current research on this devastating disease.

2. Materials and Methods

2.1. Sampling and Inoculation Test

The *B. xylophilus* strains AMA3 (strong virulent), USA (weak virulent) and GZ1 (weak virulent) were maintained in the laboratory of Prof. Jianren Ye for all downstream analyses. OKD-1 (weak virulent) was kindly provided by the University of Miyazaki [9]. All nematodes were cultured on a fungus (*Botrytis cinerea* Pers.) at 25 °C for 7–10 days to obtain sufficient sample sizes. The nematodes were isolated by the Baermann funnel method and were stored in a water suspension using a 1.5 mL centrifuge tube for downstream sequencing. Approximately 5000 AMA3 were inoculated into a small wound of *Pinus thunbergii* stems for 6, 12 and 24 h and cultivated at temperatures between 28 and 32 °C. Subsequently, the nematodes were extracted from the chips of the entire seedlings for 3 h using the Baermann funnel method. After they were washed three times in double-distilled (dd) H₂O, the nematodes were collected by centrifugation at 3500 rpm for 3 min and frozen in liquid nitrogen for subsequent experimental validation.

2.2. Identification and Quantification of Multi-Copy GH45 Genes

The genomic coordinates and annotation of the four GH45 genes were obtained by the long-read sequencing assembly (PRJNA524063). The identification of SDs and TRs was performed by BLAST based on the annotation result. Multiple gene comparison was conducted by DNAMAN (Lynnon Corp, Quebec, Canada). On the transcriptome level, four specific primer pairs were designed on the first and last exons of each GH45 gene based on previous alignment results. Each of the primer pairs contained > two specific nucleotide sites. PCR was conducted by the following procedure: 94 °C for 2 min, followed by 37 cycles (94 °C for 30 s, 58 °C for 30 s and 72 °C for 30 s) and finally 72 °C

for 10 min. The PCR products were examined using 1% agarose gel electrophoresis. At the genome level, five primers (Supplementary Table S2) were designed in the intergenic regions (P1, P2, P4 and P5) of the four GH45 genes and different combinations were utilized to validate their existence (P2F+P2R, P2F+P4R, P1F+P2R, P1F+P4R and P1F+P5R). Highly sensitive Takara PrimeStar Max enzyme was employed to perform long-range PCR with the three-step procedure described in the manufacture manual. The extension time was adjusted to 25s/kb as needed. The signal peptide cleavage site was predicted by (<http://www.cbs.dtu.dk/services/SignalP/> (accessed on 20 January 2021)).

The RT-PCR test was conducted using an ABI StepOne plus real-time system (Applied Biosystems, Carlsbad, CA, USA) with SYBR Premix Ex Taq (Takara, Dalian, China). The results were normalized to the expression level of the *B. xylophilus* Actin gene (GenBank: EU100952). All RT-PCR experiments were conducted with four replicates (Supplementary Table S3).

2.3. Detection of Copy Number Variations

Suitable primers were designed from the flanking sequences of TRs. Exact copy numbers can be assessed appertaining to amplified product lengths (sense primer: CGTGGTG-GAAGGTAAGAC, antisense primer: GTTGGAGTGGTGCTTGAT). PCR was conducted by the following procedure: 94 °C for 2 min, followed by 35 cycles (94 °C for 30 s, 58 °C for 30 s and 72 °C for 30 s) and finally 72 °C for 10 min. The PCR products were visualized using 3% agarose gel electrophoresis.

3. Results

3.1. Identification of Novel HGT Gene Families, Glycoside Hydrolase 45

In this study, four GH45 genes were determined (GH45_1: 1177 bp, GH45_2: 1091 bp, GH45_3: 912 bp and GH45_4: 1101 bp), located on the same contig adjacent to each other and separated by ~1.5 kb intergenic intervals (Figure 1a, Table 1). Based on the long-read assembly (AH1), GH45_1 and GH45_2 were considered to be different copies of the same gene with only 42 bp insertions (3' region) and 19 bp mutations (5' region), which were possibly caused by segmental duplications during evolution. GH45_3 and GH45_4 shared more abundant mutations at both 5' and 3' ends with no insertions and deletions (INDELs) and were considered two other copies of the GH45 gene. The existence of the four genes and their sequence variations was confirmed by PCR amplification followed by clone sequencing (Figure 1b, Supplementary Figure S1, Supplementary Table S1). In addition, specific primers were designed in the up and downstream intergenic regions of the four genes (P1, P2, P4 and P5) to indicate their genomic locations and architectures. The two gel bands from the same primer pairs (P2F+P2R, P2F+P4R and P1F+P2R) proved that the segmental duplication event occurred in P2 and P3. Additionally, the ~8000 bp (P1F+P4R) and ~9000 bp (P1F+P5R) fragments further indicated that the aforementioned four genes were sequentially located (Figure 1c). The genome comparison results suggested that GH45_2 and GH45_4 were not found in both the short-read assembly using Japanese strain (JP) and the newly reported long-read assembly [23] (PRJEB40022). Furthermore, only GH45_1 and GH45_3 were reported (Figure 1d). GH45_1 was determined to have six copies of 21mer TRs while GH45_2 had four copies. The extra two copies of TRs were in accordance with the aforementioned 42 bp insertions (Figures 1a and 2a–c).

Table 1. An overview of GH45 genes with different copies.

	GH45_1	GH45_2	GH45_3	GH45_4
Length (bp)	1177	1091	912	1101
Exon number	2	2	2	2
SP site *	Yes	Yes	NA	NA
Copies of TR	6	4	NA	NA
ORF Length	1056	1014	162	288

* SP site indicates the signal peptide cleavage site.

Figure 1. Identification of multi-copy Glycoside hydrolase 45 (GH45) genes. (a). Visualization of genome locations of GH45 genes, (b). PCR amplification of 4 GH45 genes using AMA3 cDNA, (c). PCR amplification of intergenic regions among GH45 genes and (d). Genome comparison of GH45 genes found in AH1 against JP; the missing copies GH45_2 and GH45_4 are highlighted in the red rectangle.

Figure 2. Identification of tandem repeat (TR) variations within multi-copy GH45 genes. (a) Clone sequencing results of TR regions in GH45_1, (b) Clustal analysis of GH45_1 and GH45_2 clone sequencing results; highlighted blue rectangles indicate the TR variations between GH45_1 and GH45_2 and (c) Clone sequencing results of TR regions in GH45_2; dashed double headed arrows indicate the position where 2 copies of TRs (21 bp) were deleted.

3.2. Sequence Characterizations of Different GH45 Copies

The ORFs among different copies were altered by nucleotide variations and TRs. The longest ORF of GH45_1 was 1056 bp while that of GH45_2 was 1014 bp. The TR variations found in GH45_1 and GH45_2 altered the ORF regions and included two extra copies of repeated peptides (Figure 3). In addition, the ORFs between GH45_3 (162 bp) and GH45_4 (288 bp) diverged significantly. They shared the same start codons but the stop codons were modified by nucleotide variations. Motif locations and types were almost the same between GH45_1 and GH45_2, whilst GH45_2 missed one repeat motif, which was mostly due to the lack of two copies of TRs. Motifs between GH45_3 and GH45_4 varied on the 3' region. Among the 5' regions, different copies were more conserved (Supplementary Figure S2). GH45_1 and GH45_2 were predicted to possess the same signal peptide cleavage site with 21–22 amino acids, whilst GH45_3 and GH45_4 had no such site (Figure 3).

Figure 3. Sequence analysis of GH45 genes with different copies. Linear view of ORFs and TR locations among different GH45 copies—the pink triangle indicates the signal peptide cleavage site.

3.3. Expression Profiling of Novel Multi-Copy GH45 Genes

Due to their highly similar sequences, the GH45 gene copies cannot be well quantified by RNAseq data. This is because sequence aligners are incapable of handling sequencing reads within highly repetitive regions. To confirm the real expressions among different GH45 copies, four multi-copy genes were quantified by a conventional qPCR method using specific primer pairs. The sensitivity of the four qPCR primer pairs was validated by clone sequencing in advance (Figure 4a,b, Supplementary Table S3, Supplementary Figure S3). Subsequent expression profiling of the aforementioned GH45 genes indicated that they shared similar expression patterns with distinct expression levels in early infection stages. All GH45 genes reached their climax at 12 h after inoculation and decreased significantly at 24 h. The expression level of GH45_1 was significantly higher than the other copy, GH45_2. GH45_3 and GH45_4 were expressed at a very low scale when compared with GH45_1 and GH45_2, whilst the expression of GH45_3 was still higher than that of GH45_4 (Figure 4c). These results suggest that expressions of different GH45 copies vary significantly.

Figure 4. Expression profiling of four GH45 copies with the qPCR method. (a) Validation of GH45_1 qPCR primers by clone sequencing, (b) Validation of GH45_2 qPCR primers by clone sequencing and (c) Relative expressions of 4 GH45 genes.

3.4. Copy Number Variations of GH45 Genes Among Different *B. xylophilus* Strains

Accurately genotyping the copy numbers of certain genes was difficult due to their highly similar sequence contents. However, the copy numbers of GH45_1 and GH45_2 can be assessed based on the aforementioned TR variations (Figure 1). The conventional PCR method would be adequate in identifying different copies with different TR combinations by designing suitable primers located in the flanking sequences of TRs. The PCR results combined with Sanger sequencing data indicate that there were two copies in AMA3 (six TRs; four TRs), three copies in USA (six TRs; five TRs; four TRs), one copy with five TRs in OKD-1 and one copy with seven TRs in GZ1 (Figure 5a–c). In addition, AMA3 was verified as a strong virulent strain while USA, OKD-1 and GZ1 were verified as weak virulent strains [9,12]. Obviously, the copy number variations and the TR variations of the GH45 gene existed among different nematode strains. A potential method would be to distinguish *B. xylophilus* strains that have a different virulence, as more strains with distinguished virulence variations should be included to provide more concrete conclusions in future works.

Figure 5. Copy number variations among different *B. xylophilus* strains. (a) PCR validation of copy number variations, (b) Illustrations of various TR combinations within different *B. xylophilus* strains and (c) Validation of various TR combinations within different *B. xylophilus* strains.

4. Discussion

The short-read-based assembly generally did not show an extraordinary capacity in characterizing long repetitive sequences. Consequently, few genes annotated as multiple copies with highly similar sequence contents and copy number variations have been previously reported in *B. xylophilus* [7,14]. In this study, we found that some *B. xylophilus*-specific GH45 genes had multiple copies with the help of long-read-based (PRJNA524063) assembly (the misidentification of GH45 copies in another long-read assembly may be caused by copy number variations—the sequencing strains may not contain any different copies). Hitherto, GH45_1 and GH45_3 were reported as Bx-eng-3 and Bx-eng-2, respectively, whilst the other copies (GH45_2 and GH45_4) were ignored in other studies [14] (Figure 1d). This indicates that some GH45 genes were not only acquired from horizontal gene transfer but also evolved to maintain multiple copies within SD regions.

Additionally, GH45_1 and GH45_2 were proved to locate within the SD region and experimentally verified to possess TR variations within their ORFs (Figures 1a and 2a–c). Relevant research has suggested that TRs could be formed when SDs occurred [24–26] and TRs found within ORFs were usually associated with gene expression, protein structure and function [27–29]. At the same time, TRs were also reported as mediators for phenotypic change, which facilitated rapid adaptation of the organisms to the environment [30]. It was clear that GH45 cellulase was one of the major plant cell-wall-degradation enzymes of *B. xylophilus* [14]. Thus, the TR variations (unit size: 21 mer) of GH45_1 and GH45_2, along with their distinct expressions (Figures 3 and 4), would provide potential evidence for the diverse virulence and wide distribution range of *B. xylophilus*. Additionally, the expression of GH45 genes was boosted 12 h after inoculation and decreased at 24 h. This expression pattern was negatively correlated with the host response activities that occurred at the corresponding time points, which possibly inhibited the pathogenesis-related gene expressions [31]. GH45_3 and GH45_4 did not have TRs but their expression levels were also significantly variable. It is possible that the *B. xylophilus*-specific GH45 cell-wall-degradation genes evolved to maintain their parasitism by introducing various copies and TRs to achieve various gene functions and protein structures.

Traditional population studies on the genetic variations of *B. xylophilus* are usually based on the different enzyme sites within DNA fragments, such as ISSR, AFLP and RFLP [32–34] and other biological markers using mitochondrial DNA, tandem repeats and microsatellites [35–38]. In the current study, a new approach was proposed to classify *B. xylophilus* with different virulence based on copy number variations. Copy number variations of *B. xylophilus* can be easily detected by simple PCR amplifications. This

approach could be utilized as a novel biological marker to characterize the virulence of *B. xylophilus* as well as assist in population classification.

Other studies have also revealed possible relationships between gene copies and complex traits [39]. Global detection of copy numbers in human genomes proved that copy number variations occur within disease loci and functional elements [40]. Evolutionary study of immune-associated genes in *Drosophila* showed that copy number variations tend to vary on signaling proteins, suggesting that molecular interactions between hosts and pathogens may drive adaptive evolution in the immune system [41]. Apart from model organisms, researchers have also found that juvenile hormone esterase-related genes in pests occur concomitantly with the loss of functional motifs during gene expansion [42]. Therefore, it is possible that the copy number variations found in GH45 genes in *B. xylophilus* may correlate with certain pathological traits and mediate the infection process.

5. Conclusions

The accurate identification and quantification of GH45 genes with different copies elucidates a new perspective upon the pathogenicity of *B. xylophilus*. In future studies, the similarities among different copies of GH45 genes should be recognized when their expressions are quantified or relevant studies are conducted. Traditional short-read RNAseq data with a 100–200 bp read-length were incapable of quantifying different copies of GH45 genes with accuracy. Therefore, long-read-based transcriptome sequencing would be a better method to obtain more solid results. However, the interaction between copy number variations of GH45 genes and the pathogenesis of PWN still requires additional experiments with sufficient strains.

Supplementary Materials: The following are available online at <https://www.mdpi.com/1999-4907/12/3/275/s1>, Figure S1: Clustal alignment of 4 GH45 genes. Figure S2: Gene structure variations among different GH45 copies. Figure S3: Clone sequencing results of GH45 qPCR primers; (a) Clone sequencing of GH45_1 qPCR amplification results (~20 clones); (b) Clone sequencing of GH45_2 qPCR amplification results (~20 clones); (c) Clone sequencing of GH45_3 qPCR amplification results, only 1 clone was sequenced since the primers were designed on highly divergent regions; (d) Clone sequencing of GH45_4 qPCR amplification results, only 1 clone was sequenced since the primers were designed on highly specific regions. Table S1. Primer design for the GH45 RT-PCR assay. Table S2. Primer design for the long-range PCR assay. Table S3. Primer design for the GH45 qPCR assay.

Author Contributions: Conceptualization, X.D. and J.Y.; methodology, X.D. and Q.W.; software, X.D. and Y.G.; validation, X.D., Y.L. and S.L.; formal analysis, D.-W.L.; investigation, D.-W.L.; resources, S.L.; data curation, X.D.; writing—original draft preparation, X.D.; writing—review and editing, D.-W.L., Q.Z. and J.Y.; visualization, X.D. and F.S.; supervision, J.Y.; project administration, X.D.; funding acquisition, X.D. All authors have read and agreed to the published version of the manuscript. Authorship must be limited to those who have contributed substantially to the work reported.

Funding: This project is supported by the National Natural Science Foundation of China 31800543 (X.D.) and Z411B17105 (Q.Z.), National Key Research and Development Project 2018YFD0600203 (J.Y.) and Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Strains and primers are available upon request. The authors affirm that all data necessary for confirming the conclusions of the article are present within the article, figures and tables.

Acknowledgments: We are grateful to James LaMondia for reviewing the manuscript, and Ruiwen Zhao for their technical assistance in preparing the nematode strains used in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mamiya, Y. Pathology of the pine wilt disease caused by *Bursaphelenchus xylophilus*. *Annu. Rev. Phytopathol.* **1983**, *21*, 201–220. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Soliman, T.; Mourits, M.C.M.; van der Werf, W.; Hengeveld, G.M.; Robinet, C.; Lansink, A.G.J.M.O. Framework for modelling economic impacts of invasive species, applied to pine wood nematode in Europe. *PLoS ONE* **2012**, *7*, e45505. [\[CrossRef\]](#)
3. Espada, M.; Eves-van den Akker, S.; Maier, T.; Vijayapalani, P.; Baum, T.; Mota, M.; Jones, J.T. STATAWAARS: A promoter motif associated with spatial expression in the major effector-producing tissues of the plant-parasitic nematode *Bursaphelenchus xylophilus*. *BMC Genom.* **2018**, *19*, 1–13. [\[CrossRef\]](#)
4. Espada, M.; Silva, A.C.; Eves van den Akker, S.; Cock, P.J.; Mota, M.; Jones, J.T. Identification and characterization of parasitism genes from the pinewood nematode *Bursaphelenchus xylophilus* reveals a multilayered detoxification strategy. *Mol. Plant Pathol.* **2016**, *17*, 286–295. [\[CrossRef\]](#)
5. Tanaka, S.E.; Dayi, M.; Maeda, Y.; Tsai, I.J.; Tanaka, R.; Bligh, M.; Takeuchi-Kaneko, Y.; Fukuda, K.; Kanzaki, N.; Kikuchi, T. Stage-specific transcriptome of *Bursaphelenchus xylophilus* reveals temporal regulation of effector genes and roles of the dauer-like stages in the lifecycle. *Sci. Rep.* **2019**, *9*, 1–13. [\[CrossRef\]](#)
6. Tsai, I.J.; Tanaka, R.; Kanzaki, N.; Akiba, M.; Yokoi, T.; Espada, M.; Jones, J.T.; Kikuchi, T. Transcriptional and morphological changes in the transition from mycetophagous to phytophagous phase in the plant-parasitic nematode *Bursaphelenchus xylophilus*. *Mol. Plant Pathol.* **2016**, *17*, 77–83. [\[CrossRef\]](#)
7. Kikuchi, T.; Cotton, J.A.; Dalzell, J.J.; Hasegawa, K.; Kanzaki, N.; McVeigh, P.; Takanashi, T.; Tsai, I.J.; Assefa, S.A.; Cock, P.J.A.; et al. Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. *PLoS Pathog* **2011**, *7*, e1002219. [\[CrossRef\]](#)
8. Figueiredo, J.; Simoes, M.J.; Gomes, P.; Barroso, C.; Pinho, D.; Conceicao, L.; Fonseca, L.; Abrantes, I.; Pinheiro, M.; Egas, C. Assessment of the geographic origins of pinewood nematode isolates via single nucleotide polymorphism in effector genes. *PLoS ONE* **2013**, *8*, e83542. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Palomares-Rius, J.E.; Tsai, I.J.; Karim, N.; Akiba, M.; Kato, T.; Maruyama, H.; Takeuchi, Y.; Kikuchi, T. Genome-wide variation in the pinewood nematode *Bursaphelenchus xylophilus* and its relationship with pathogenic traits. *BMC Genom.* **2015**, *16*, 1–13. [\[CrossRef\]](#)
10. Li, Y.X.; Meng, F.L.; Deng, X.; Wang, X.; Feng, Y.Q.; Zhang, W.; Pan, L.; Zhang, X.Y. Comparative transcriptome analysis of the pinewood nematode *Bursaphelenchus xylophilus* reveals the molecular mechanism underlying its defense response to host-derived -pinene. *Int. J. Mol. Sci.* **2019**, *20*, 911. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Ding, X.L.; Ye, J.R.; Wu, X.Q.; Huang, L.; Zhu, L.H.; Lin, S.X. Deep sequencing analyses of pine wood nematode *Bursaphelenchus xylophilus* microRNAs reveal distinct miRNA expression patterns during the pathological process of pine wilt disease. *Gene* **2015**, *555*, 346–356. [\[CrossRef\]](#)
12. Ding, X.; Ye, J.; Lin, S.; Wu, X.; Li, D.; Nian, B. Deciphering the molecular variations of pine wood nematode *Bursaphelenchus xylophilus* with different virulence. *PLoS ONE* **2016**, *11*, e0156040. [\[CrossRef\]](#)
13. Kikuchi, T.; Jones, J.T.; Aikawa, T.; Kosaka, H.; Ogura, N. A family of glycosyl hydrolase family 45 cellulases from the pine wood nematode *Bursaphelenchus xylophilus*. *FEBS Lett.* **2004**, *572*, 201–205. [\[CrossRef\]](#)
14. Palomares-Rius, J.E.; Hirooka, Y.; Tsai, I.J.; Masuya, H.; Hino, A.; Kanzaki, N.; Jones, J.T.; Kikuchi, T. Distribution and evolution of glycoside hydrolase family 45 cellulases in nematodes and fungi. *BMC Evol. Biol.* **2014**, *14*, 69. [\[CrossRef\]](#)
15. O'Dushlaine, C.T.; Edwards, R.J.; Park, S.D.; Shields, D.C. Tandem repeat copy-number variation in protein-coding regions of human genes. *Genome Biol.* **2005**, *6*, 1–12. [\[CrossRef\]](#)
16. Brahmachary, M.; Guilmatre, A.; Quilez, J.; Hasson, D.; Borel, C.; Warburton, P.; Sharp, A.J. Digital genotyping of macrosatellites and multicopy genes reveals novel biological functions associated with copy number variation of large tandem repeats. *PLOS Genet* **2014**, *10*, e1004418. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Orgel, L.E.; Crick, F.H. Selfish DNA: The ultimate parasite. *Nature* **1980**, *284*, 604–607. [\[CrossRef\]](#)
18. Duda, T.F., Jr.; Palumbi, S.R. Molecular genetics of ecological diversification: Duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6820–6823. [\[CrossRef\]](#)
19. Verstrepen, K.J.; Jansen, A.; Lewitter, F.; Fink, G.R. Intragenic tandem repeats generate functional variability. *Nat. Genet.* **2005**, *37*, 986–990. [\[CrossRef\]](#)
20. Gymrek, M.; Willems, T.; Guilmatre, A.; Zeng, H.Y.; Markus, B.; Georgiev, S.; Daly, M.J.; Price, A.L.; Pritchard, J.K.; Sharp, A.J.; et al. Abundant contribution of short tandem repeats to gene expression variation in humans. *Nat. Genet.* **2016**, *48*, 22–32. [\[CrossRef\]](#)
21. Quilez, J.; Guilmatre, A.; Garg, P.; Highnam, G.; Gymrek, M.; Erlich, Y.; Joshi, R.S.; Mittelman, D.; Sharp, A.J. Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans. *Nucleic Acids Res.* **2016**, *44*, 3750–3762. [\[CrossRef\]](#)
22. Cantsilieris, S.; Baird, P.N.; White, S.J. Molecular methods for genotyping complex copy number polymorphisms. *Genomics* **2013**, *101*, 86–93. [\[CrossRef\]](#)
23. Dayi, M.; Sun, S.; Maeda, Y.; Tanaka, R.; Yoshida, A.; Tsai, I.J.; Kikuchi, T. Nearly Complete Genome Assembly of the Pinewood Nematode *Bursaphelenchus xylophilus* Strain Ka4C1. *Microbiol. Res. Announc.* **2020**, *9*, e01002–e01020. [\[CrossRef\]](#)
24. Fan, H.; Chu, J.Y. A brief review of short tandem repeat mutation. *Genom. Proteom. Bioinform.* **2007**, *5*, 7–14. [\[CrossRef\]](#)

25. Paques, F.; Leung, W.Y.; Haber, J.E. Expansions and contractions in a tandem repeat induced by double-strand break repair. *Mol. Cell. Biol.* **1998**, *18*, 2045–2054. [[CrossRef](#)]
26. Fiston-Lavier, A.S.; Anxolabehere, D.; Quesneville, H. A model of segmental duplication formation in *Drosophila melanogaster*. *Genome Res.* **2007**, *17*, 1458–1470. [[CrossRef](#)] [[PubMed](#)]
27. Fondon, J.W., 3rd; Garner, H.R. Molecular origins of rapid and continuous morphological evolution. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 18058–18063. [[CrossRef](#)]
28. Nasrallah, M.P.; Cho, G.; Simonet, J.C.; Putt, M.E.; Kitamura, K.; Golden, J.A. Differential effects of a polyalanine tract expansion in *Arx* on neural development and gene expression. *Hum. Mol. Genet.* **2012**, *21*, 1090–1098. [[CrossRef](#)] [[PubMed](#)]
29. Matsushima, N.; Yoshida, H.; Kumaki, Y.; Kamiya, M.; Tanaka, T.; Izumi, Y.; Kretsinger, R.H. Flexible structures and ligand interactions of tandem repeats consisting of proline, glycine, asparagine, serine, and/or threonine rich oligopeptides in proteins. *Curr. Protein Pept. Sci.* **2008**, *9*, 591–610. [[CrossRef](#)] [[PubMed](#)]
30. Young, E.T.; Sloan, J.S.; Van Riper, K. Trinucleotide repeats are clustered in regulatory genes in *Saccharomyces cerevisiae*. *Genetics* **2000**, *154*, 1053–1068.
31. Yu, L.-Z.; Wu, X.-Q.; Ye, J.-R.; Zhang, S.-N.; Wang, C. NOS-like-mediated nitric oxide is involved in *Pinus thunbergii* response to the invasion of *Bursaphelenchus xylophilus*. *Plant Cell Rep.* **2012**, *31*, 1813–1821. [[CrossRef](#)]
32. Mallez, S.; Castagnone, C.; Espada, M.; Vieira, P.; Eisenback, J.D.; Harrell, M.; Mota, M.; Aikawa, T.; Akiba, M.; Kosaka, H.; et al. Worldwide invasion routes of the pinewood nematode: What can we infer from population genetics analyses? *Biol. Invasions* **2015**, *17*, 1199–1213. [[CrossRef](#)]
33. Takemoto, S.; Kanzaki, N.; Futai, K. PCR-RFLP image analysis: A practical method for estimating isolate-specific allele frequency in a population consisting of two different strains of the pinewood nematode, *Bursaphelenchus xylophilus* (Aphelenchida: Aphelenchoididae). *Appl. Entomol. Zool.* **2005**, *40*, 529–535. [[CrossRef](#)]
34. Cheng, X.Y.; Cheng, F.X.; Xu, R.M.; Xie, B.Y. Genetic variation in the invasive process of *Bursaphelenchus xylophilus* (Aphelenchida: Aphelenchoididae) and its possible spread routes in China. *Heredity* **2008**, *100*, 356–365. [[CrossRef](#)]
35. Zhou, L.F.; Chen, F.M.; Xie, L.Y.; Pan, H.Y.; Ye, J.R. Genetic diversity of pine-parasitic nematodes *Bursaphelenchus xylophilus* and *Bursaphelenchus mucronatus* in China. *Forest Pathol.* **2017**, *47*, e12334. [[CrossRef](#)]
36. Pereira, F.; Moreira, C.; Fonseca, L.; van Asch, B.; Mota, M.; Abrantes, I.; Amorim, A. New insights into the phylogeny and worldwide dispersion of two closely related nematode species, *Bursaphelenchus xylophilus* and *Bursaphelenchus mucronatus*. *PLoS ONE* **2013**, *8*, e56288. [[CrossRef](#)]
37. Jung, J.; Han, H.; Ryu, S.H.; Kim, W. Microsatellite variation in the pinewood nematode, *Bursaphelenchus xylophilus* (Steiner and Buhrer) Nickle in South Korea. *Genes Genom.* **2010**, *32*, 151–158. [[CrossRef](#)]
38. Vieira, P.; Castagnone, C.; Mallez, S.; Espada, M.; Navas, A.; Mota, M.; Castagnone-Sereno, P. Sequence variability of the MspI satellite DNA family of the pinewood nematode *Bursaphelenchus xylophilus* at different geographic scales. *J. Nematol.* **2013**, *45*, 324. [[CrossRef](#)] [[PubMed](#)]
39. Hahn, M.W. Distinguishing Among Evolutionary Models for the Maintenance of Gene Duplicates. *J. Hered.* **2009**, *100*, 605–617. [[CrossRef](#)]
40. Redon, R.; Ishikawa, S.; Fitch, K.R.; Feuk, L.; Perry, G.H.; Andrews, T.D.; Fiegler, H.; Shapero, M.H.; Carson, A.R.; Chen, W.; et al. Global variation in copy number in the human genome. *Nature* **2006**, *444*, 444–454. [[CrossRef](#)]
41. Sackton, T.B.; Lazzaro, B.P.; Schlenke, T.A.; Evans, J.D.; Hultmark, D.; Clark, A.G. Dynamic evolution of the innate immune system in *Drosophila*. *Nat. Genet.* **2007**, *39*, 1461–1468. [[CrossRef](#)]
42. Kontogiannatos, D.; Swevers, L.; Kourti, A. Recent gene multiplication and evolution of a juvenile hormone esterase-related gene in a lepidopteran pest. *Gene Rep.* **2016**, *4*, 139–152. [[CrossRef](#)]