BMC Genomics

Whole genome analysis of the koa wilt pathogen (*Fusarium oxysporum* f. sp. *koae*) and the development of molecular tools for early detection and monitoring

John T. Dobbs[1], Mee-Sook Kim[2], Nicklos S. Dudley[3], Ned B. Klopfenstein[4], Aileen Yeh[3], Robert D. Hauff[5], Tyler C. Jones[3], R. Kasten Dumroese[4], Philip G. Cannon[6] and Jane E. Stewart[1*]

Abstract

Background: Development and application of DNA-based methods to distinguish highly virulent isolates of *Fusarium oxysporum* f. sp. *koae* [*Fo koae*; cause of koa wilt disease on *Acacia koa* (koa)] will help disease management through early detection, enhanced monitoring, and improved disease resistance-breeding programs.

Results: This study presents whole genome analyses of one highly virulent *Fo koae* isolate and one non-pathogenic *F. oxysporum* (*Fo*) isolate. These analyses allowed for the identification of putative lineage-specific DNA and predicted genes necessary for disease development on koa. Using putative chromosomes and predicted gene comparisons, *Fo koae*-exclusive, virulence genes were identified. The putative lineage-specific DNA included identified genes encoding products secreted in xylem (e. g., *SIX1* and *SIX6*) that may be necessary for disease development on koa. Unique genes from *Fo koae* were used to develop pathogen-specific PCR primers. These diagnostic primers allowed target amplification in the characterized highly virulent *Fo koae* isolates but did not allow product amplification in low-virulence or non-pathogenic isolates of *Fo*. Thus, primers developed in this study will be useful for early detection and monitoring of highly virulent strains of *Fo koae*. Isolate verification is also important for disease resistance-breeding programs that require a diverse set of highly virulent *Fo koae* isolates for their disease-screening assays to develop disease-resistant koa.

Conclusions: These results provide the framework for understanding the pathogen genes necessary for koa wilt disease and the genetic variation of *Fo koae* populations across the Hawaiian Islands.

Keywords: *Fusarium oxysporum*, Lineage-specific DNA, Virulence genes, PCR primer, Haplotypes, *Acacia koa*

* Correspondence: jane.stewart@colostate.edu
[1]Colorado State University, Department of Agricultural Biology, 1177 Campus Delivery, Fort Collins, CO 80523, USA
Full list of author information is available at the end of the article

Background

Intraspecific genomic comparisons between pathogenic and non-pathogenic fungal isolates can elucidate key factors required for pathogenicity in plant hosts [1, 2]. These factors include unique genes that are not only vital for the pathogen to cause disease, but can also provide tools to differentiate pathogens and non-pathogens for detection and monitoring. Variation in genome size, predicted effectors, and unique genome regions have been used previously to differentiate plant pathogens from saprophytic non-pathogens [3–5]. Just as the hosts exhibit variation in resistance, fungal pathogens (e.g., *Fusarium oxysporum*) also exhibit variation in virulence [6]. Previous analyses of *F. oxysporum* f. sp. *lycopersici* and *F. oxysporum* f. sp. *radicis-cucumerinum* chromosomes identified conditionally dispensable chromosomes containing virulence factors that could be transferred between compatible strains through heterokaryon formation [4, 7]. Other previous studies found that these virulence factors were encoded by diverse genes. *Fusarium* transcription factors and their corresponding virulence genes, such as those encoding proteins secreted in xylem (*SIX*), were found to be essential for pathogenesis and high virulence in most *F. oxysporum* formae species [8–11]. Based on carbohydrate-active enzyme analysis, the copy number and types of plant cell wall-degrading enzymes have been found to be important for pathogenesis of *F. oxysporum* formae species in legumes [12]. In addition, phytotoxic secondary metabolites and associated transporters (i.e., fusaric acid, fumonisin B1, and beauvericin) have been identified as important for pathogenesis in the *F. oxysporum* f. sp. *cubense* [12, 13], and transposable elements (i.e., miniature impala elements) have been documented to enhance effector diversity [14–16]. These factors help distinguish pathogenicity and/or virulence in *Fusaria* and could also provide a basis for early detection of pathogens which would facilitate disease management programs through efforts to reduce pathogen spread to uninfested sites.

Fusarium oxysporum is a cosmopolitan, ascomycete fungus with observed functions that range from an innocuous soil saprophyte to a catastrophic plant pathogen in important agricultural crops and ecologically important tree species [17]. In addition to these diverse functions, *F. oxysporum* is also known to display high levels of host-specificity. These highly host-specific pathogenic strains are classified into formae speciales (forma specialis singular) that exhibit high genomic plasticity, which has been attributed to effector profiles housed on one to many conditionally dispensable, lineage-specific accessory chromosomes, where the vast majority of genomic variation exists, in comparison to a set of shared "core" chromosomes containing essential genes for normal cellular function [16, 18–20]. These lineage-specific chromosome(s) house genes involved

in virulence and/or evasion of plant host detection and have been described as repeat- and transposon-rich [1, 3]. Comparisons of these virulence genes through whole-genome analyses are useful for distinguishing pathogens from non-pathogens on the susceptible host of a specific *F. oxysporum* forma specialis [1]. In this study, we analyzed and compared the genomes of pathogenic and non-pathogenic *F. oxysporum* strains, which were both isolated from the roots of *Acacia koa* A. Gray (koa).

Gardner (1980) described Fusarium wilt disease in low- to mid-elevation, below 610 m (2000 ft), forests in the Hawaiian archipelago affecting koa, an endemic tree species [21, 22]. Koa trees are ecologically, economically, and culturally valuable to Hawaiʻi. Koa is the second most abundant tree in Hawaiʻi's forests and many of Hawaiʻi's indigenous flora and fauna rely on koa for food and shelter [21]. Due to koa's importance to Hawaiʻi, efforts have been increased to reincorporate koa into lowland forests, where its populations have declined [23]. Currently, koa wilt disease [caused by *F. oxysporum* f. sp. *koae* (*Fo koae*)] represents the predominant threat to koa across its native range [24].

Management methods for koa wilt disease include fumigation and pesticides that are environmentally harmful and economically costly when used in forests and/or timber stands. As a result, current efforts are focused primarily on producing disease-resistant seedlings derived from seed collected from natural stands [23]. To find naturally occurring resistance to the pathogenic strains of the fungus (*Fo koae*), wild-collected seed is used for extensive disease resistance-screening assays in the greenhouse [25]. These screening trials are costly and require diverse pathogenic *Fo koae* strains to ensure that the disease resistance is robust in the selected germplasm. Consequently, these disease resistance-screening programs would greatly benefit from faster, more cost-effective strategies for screening pathogenic *Fo koae* isolates. This screening process is further complicated by considerations that pathogenic fungi can evolve quickly; much faster than their plant hosts [1, 26, 27]. For this reason, new, genetically diverse strains of the pathogen must be continually incorporated in screening programs to help ensure that disease resistance remains stable against newly evolved or introduced strains of the pathogen in Hawaiʻi's forest and agricultural lands. Collecting, isolating, and characterizing pathogenic *Fo koae* isolates is an arduous and time-consuming process that requires 3-month greenhouse virulence assays to confirm pathogenicity. Pathogen-specific, PCR primers would allow for rapid detection of pathogenic *Fo koae* strains, which would expedite collections of new pathogenic strains for disease resistance-breeding programs to meet the needs of growers and land managers for restoration and/or timber purposes.

Dobbs *et al. BMC Genomics*   (2020) 21:764

Page 3 of 15

Objectives for this study were to 1) conduct and analyze whole genome sequences of one highly virulent *Fo koae* isolate and one non-pathogenic *F. oxysporum* (*Fo*) isolate both collected from koa; 2) identify virulence-associated genes with putative roles in koa wilt disease; 3) use the identified differences to develop *Fo koae*-specific PCR primers; and 4) test the robustness of the developed primers to identify highly virulent *Fo koae* isolates among uncharacterized *F. oxysporum* populations.

Results

Whole genome sequencing and assembly
To conduct whole genome comparisons, one pathogenic *Fo koae* isolate (44) and one non-pathogenic *Fo* isolate (170) were whole genome sequenced using Illumina TruSeq short-read sequencing (150 bp). Libraries resulted in 397,390,528 reads with 1245x coverage for pathogenic isolate *Fo koae* 44 and 385,805,256 reads with 1150x coverage for non-pathogenic isolate *Fo* 170. De novo assembly of both *Fo koae* 44 and *Fo* 170 identified genome sizes of 48 and 50 Mb, respectfully. The QUAST report metrics are shown in Table 1.

Whole genome phylogenetic analysis
A whole-genome maximum likelihood phylogeny was constructed to elucidate the evolutionary relationships of *Fo koae* 44 and *Fo* 170 compared with other *Fusarium* spp. and *F. oxysporum* formae speciales (Additional File 9). The phylogeny showed that both *Fo koae* 44 and *Fo* 170 grouped into a well-supported [Bootstrap (BS) = 100] clade with other *F. oxysporum* formae speciales (Additional File 1). Interestingly, both isolates grouped into separate sub-clades with distinct *F. oxysporum* formae speciales, rather than grouping together (Additional File 1). *Fo koae* 44 clustered in a sub-clade (BS =100) that included *F. oxysporum* f. sp. *cubense* race 1 and *Fo* 170 clustered in a sub-clade (BS =100) with *F. oxysporum* f. sp. *gladioli*. This result suggests that these isolates were indeed *F. oxysporum* and that *Fo koae* 44 and *Fo* 170 were sufficiently, genetically different to allow comparison for identifying lineage-specific DNA necessary for pathogenicity in koa.

Table 1 QUAST assembly statistics of the SPAdes de novo genome assembly of the pathogenic *Fusarium oxysporum* f. sp. *koae* 44 (*Fo koae* 44) and non-pathogenic *F. oxysporum* 170 (*Fo* 170) isolates
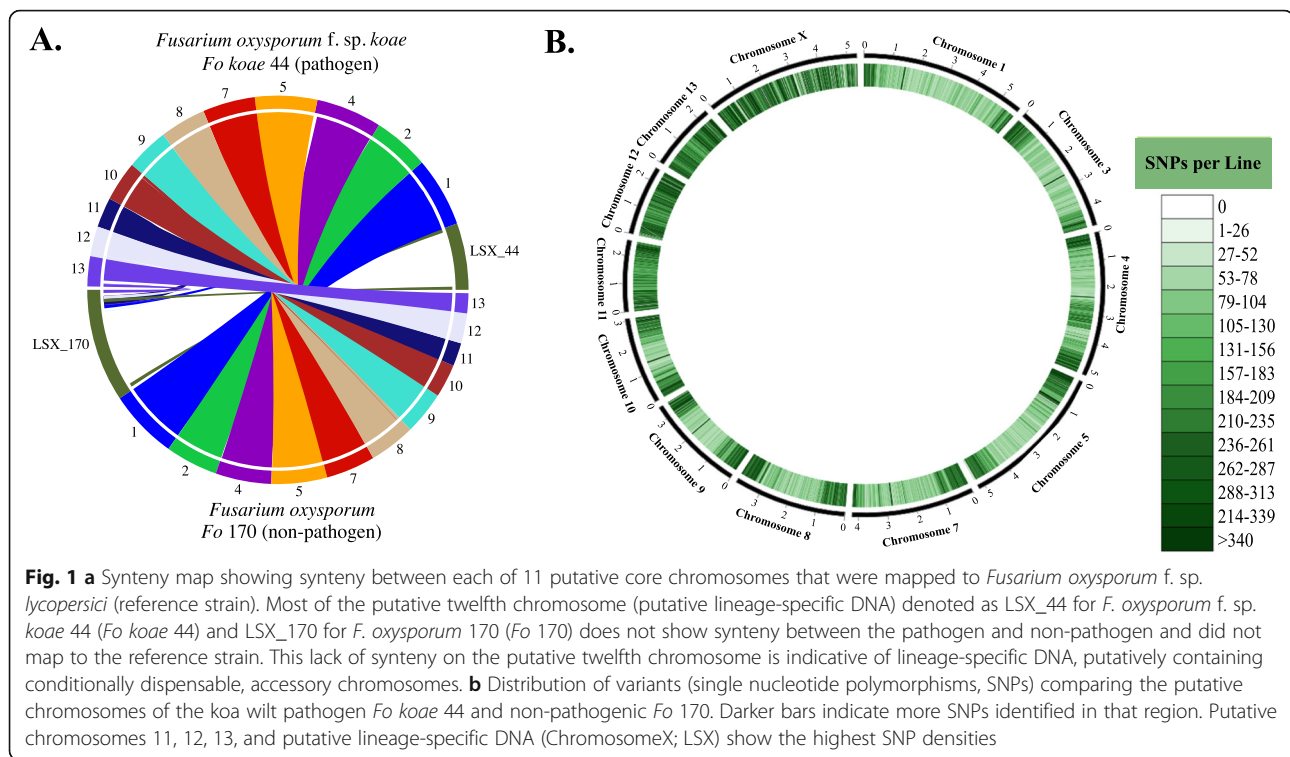
	Fo koae 44	*Fo* 170
Genome size (Mb)	48,183,513	50,628,816
Contig Number	492	964
Scaffold Number	330	910
N50	721,547	562,269
GC%	47.5	47.2

Genomic analyses of *Fo koae* 44 and *Fo* 170
Genomic sequences of *Fo koae* 44 and *Fo* 170 were aligned with reference genome sequences of a well-characterized *F. oxysporum* f. sp. *lycopersici* isolate (GenBank accession AAXH00000000.1) to identify regions of highest dissimilarity and identify lineage-specific DNA, with potential use for pathogen-specific primer development. Based on the reference genome, 11 putative "core" chromosomes shared synteny with *Fo koae* 44 and *Fo* 170. Putative lineage-specific DNA (LSX) was identified in each genome that did not map to the genome of *F. oxysporum* f. sp. *lycopersici* or show synteny with each other (*Fo koae* 44 and *Fo* 170) or genomes of other *F. oxysporum* formae speciales (Fig. 1a). Low-coverage sequences and non-*Fusarium* DNA were discarded to remove possible DNA contamination. The average nucleotide identity (ANIm) between *Fo koae* 44 and *Fo* 170 was 96.5%. The remaining 3.5% genetic dissimilarity is largely due to the identified LSX. Further analyses comparing the putative chromosomes of *Fo koae* 44 and *Fo* 170 indicated that variants (SNPs and indels) were localized on specific chromosomes (LSX_44 and LSX_170) and at chromosome ends (Fig. 1b). Putative core chromosomes 1, 2, 4, 5, 7, 8, 9, and 10 were the most similar with identified variants localized on chromosome ends, likely due to higher recombination observed at telomeres or assembly errors [20, 28]. Putative chromosomes 11, 12, 13, and LSX had the most variation (Fig. 1b). Analyses of the distribution of repeats and transposons determined that *F. oxysporum* transposable elements (e.g., *Foxy*, *Skippy*, *Impala*, *Hop*, *Mariner*, *Hat*, and *Helitron*) occurred on the LSX of both *Fo koae* 44 and *Fo* 170 [29]. The average core chromosome was comprised of 3 and 4% transposable elements and repeats in *Fo koae* 44 and *Fo* 170, respectively. In contrast, the transposable elements and repeats comprised 19 and 20% of LSX for *Fo koae* 44 and *Fo* 170, respectively (Fig. 2a & b). This localization of transposable elements provides evidence that differences within the LSX are likely attributable to one or more conditionally dispensable chromosome(s) in both *Fo koae* 44 and *Fo* 170.

Genome annotation
Using the Maker pipeline, 15,380 and 15,763 transcripts and corresponding proteins were predicted for *Fo koae* 44 and *Fo* 170, respectively, for use in identifying genes associated with pathogenicity in koa wilt disease. Using OrthoVenn2, 1760 non-orthologous proteins from 78 gene clusters and 1576 single genes, were identified as exclusive to *Fo koae* 44 when compared to *Fo* 170. In comparison, *Fo* 170 had 2123 exclusive non-orthologous proteins from 137 gene clusters and 1780 single genes. *Fo koae* 44 had enrichment of gene ontology (GO) terms for nucleic acid binding (Additional File 2), whereas *Fo*

Fig. 1 a Synteny map showing synteny between each of 11 putative core chromosomes that were mapped to *Fusarium oxysporum* f. sp. *lycopersici* (reference strain). Most of the putative twelfth chromosome (putative lineage-specific DNA) denoted as LSX_44 for *F. oxysporum* f. sp. *koae* 44 (*Fo koae* 44) and LSX_170 for *F. oxysporum* 170 (*Fo* 170) does not show synteny between the pathogen and non-pathogen and did not map to the reference strain. This lack of synteny on the putative twelfth chromosome is indicative of lineage-specific DNA, putatively containing conditionally dispensable, accessory chromosomes. **b** Distribution of variants (single nucleotide polymorphisms, SNPs) comparing the putative chromosomes of the koa wilt pathogen *Fo koae* 44 and non-pathogenic *Fo* 170. Darker bars indicate more SNPs identified in that region. Putative chromosomes 11, 12, 13, and putative lineage-specific DNA (ChromosomeX; LSX) show the highest SNP densities

170 had enrichment of GO terms for GTP binding, regulation of transcription from RNA polymerase II promoter during meiosis, sporulation resulting in formation of a cellular spore, and regulation of DNA-template transcription (Additional File 3). The 13,500 protein clusters that were shared between *Fo koae* 44 and *Fo* 170 had no identified GO enrichment.

Using InterProScan, 16,429 and 16,487 putative protein family domains were identified in *Fo koae* 44 and *Fo* 170, respectively. Of these putative proteins, 30 were unique to *Fo koae* 44, and 32 were unique to *Fo* 170. Interestingly, a sterigmatocystin biosynthesis gene was identified as unique in *Fo koae* 44. Of further note is that the genome of *Fo* 170 also encoded some proteins involved in pathogenesis, including oxidoreductases (i.e., monooxygenase) and putrescine biosynthesis.

Using the antiSMASH database, the genes associated with secondary metabolites were similar in both genomes of *Fo koae* 44 and *Fo* 170; however, variation did exist in gene copy number (Additional File 4). Non-ribosomal peptide synthetases (NRPS), NRPS-like, and indoles had higher gene copy numbers for *Fo koae* 44. *Fo* 170 had two more copies of genes associated with the production of the *Fusarium* mycotoxin enniatin.

Similar results were observed from genes identified using the dbCAN2 database, except that two copies of a glycoside hydrolase, GH13_22 (characterized glucan synthase), were found in *Fo koae* 44, but not in *Fo* 170 [30]. In *Fo* 170, two gene copies encoding proteins

GH141 (α-L-fucosidase or xylanase), one copy encoding GH30 (endo-β-1,4-xylanase or β-glucosidase), and one copy encoding GH43_12 (a glycoside hydrolase involved in hydrolysis and/or rearrangement of glycosidic bonds; arabinosidase) were identified.

Using the PHI-base, shared virulence-associated genes that were identified in both *Fo koae* 44 and *Fo* 170 include the following: transport associated genes, such as the ATP-binding cassette (*Gpabc1*; PHI:258); effectors, such as the *Ace1* (PHI:325) and *Mocdip1* (PHI: 3213); *Fusarium*-specific transcription factors, such as the *Ftf2* (PHI:5484) and the *SIX* gene (*SGE1*; PHI:3168); toxin detoxifying genes, such as the *GzmetE* (PHI:355) and *Tom1* (PHI:438); a chitin synthase gene, *CHS2* (PHI: 336); and a beta-1,3-glucanosyltransferase, *GAS1* (PHI: 522). Nine virulence-associated genes were identified as unique to *Fo koae* 44. These genes encoded multiple *Fusarium* transcription factors, mitogen-activated protein kinases, and ATP-binding cassette (ABC) transporters. The *SIX1* and *SIX6* genes, were identified on the LSX of *Fo koae* 44, but were absent in the *Fo* 170 genome.

These *Fo koae* 44 virulence-associated genes showed sequence similarity to other *F. oxysporum* formae speciales (GenBank Accession numbers in Additional Files 10 and 5B). The *SIX1* gene of *Fo koae* 44 showed highest sequence similarity with *F. oxysporum* f. sp. *fragariae* (85%) and the *SIX6* gene showed highest sequence similarity to *F. oxysporum* f. sp. *pisi* (97%) and *F. oxysporum* f. sp. *cucumerinum* (97%). There were seven virulence-
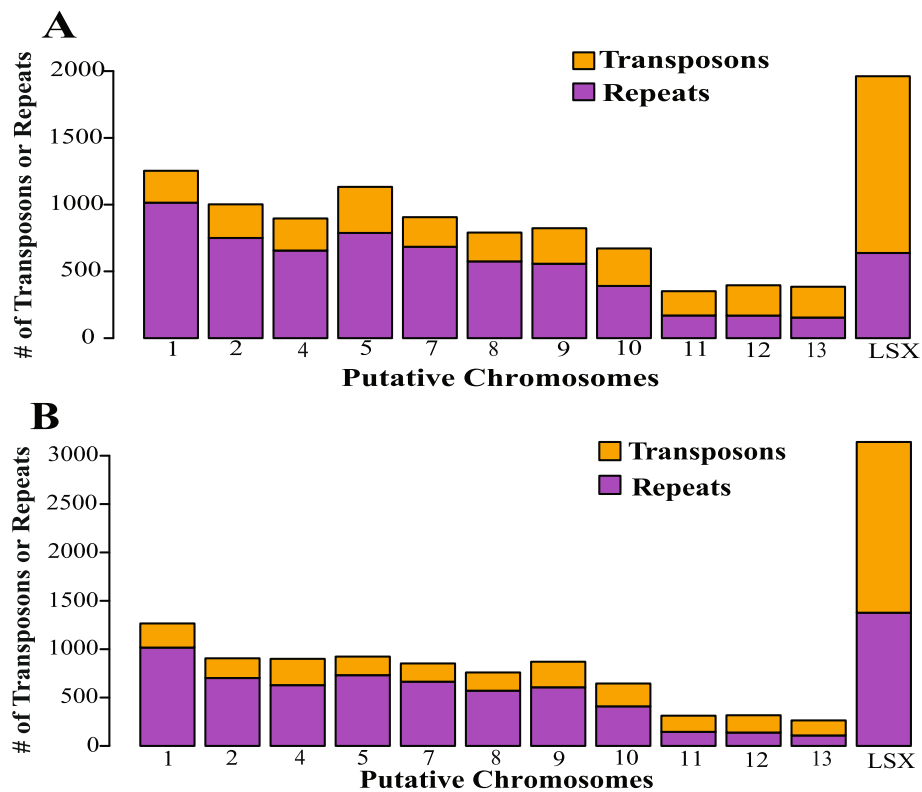
Fig. 2 a Distribution of transposons (orange bars) and repeats (purple bars) across the putative chromosomes and putative lineage-specific DNA (LSX) of *Fusarium oxysporum* f. sp. *koae* (*Fo koae* 44; koa wilt pathogen). **b** Distribution of transposons (orange bars) and repeats (purple bars) across the putative chromosomes LSX in *Fusarium oxysporum* (*Fo* 170). Majority of the transposons were identified in the LSX

associated genes identified as unique to *Fo* 170 that consisted of signaling proteins, ATP binding cassette transporter, and one of unknown function but related to developing root rot symptoms (*PEP2*).

Fo koae-specific PCR primer development

To increase the likelihood of developing a pathogen-specific primer, two approaches were used to classify genomic regions and genes unique to *Fo koae* 44 for PCR primer development. First, unique genomic regions, that may be non-coding DNA sequences, were identified in *Fo koae* 44. Second, predicted transcripts unique to *Fo koae* 44 were identified.

To identify unique genomic regions, *Fo koae* 44 chromosomes were aligned to chromosomes of *Fo* 170, other *Fusarium* spp., and *F. oxysporum* formae speciales obtained from GenBank (Additional File 9). Analysis of these alignments identified a total of 445 DNA sequences unique to *Fo koae* 44. Of these 445 DNA sequences, 33 had no BLAST hits to the NCBI database, and these 33 DNA sequences were selected for design of PCR primers.

To identify unique genes, *Fo koae* 44 and *Fo* 170 were analyzed using OrthoVenn2. One thousand seven hundred sixty-two predicted proteins were identified as exclusive to *Fo koae* 44. The corresponding transcripts

were BLASTed to the NCBI database and 137 of these transcripts had no BLAST hits. These 137 transcripts were also selected for design of PCR primers.

Taken together, each of the 33 unique genomic DNA sequences and the 137 predicted transcripts (ca. 170 genome regions) were BLASTed against the chromosomes to determine which chromosome contained the selected sequence target for potential amplification. From these sequences, 3750 primers were developed in silico. Of those, 35 primer pairs had no non-target hits in silico through NCBI primer BLAST. These primers were tested on characterized pathogenic *Fo koae* isolates and non-pathogenic *F. oxysporum* (Table 2), *F. commune* (FO21 and 85), and *F. proliferatum* (1 and 80) isolates; all but one *F. commune* isolate (FO21) were collected from koa. Of the 35 primer pairs, 7 (one designed to amplify the "core" chromosome 2 and 6 designed to amplify the LSX) did not amplify DNA from non-target species. The 6 primers designed to amplify the LSX did not result in amplification using DNA from *Fo* 170 or other characterized non-pathogenic *F. oxysporum*, *F. commune*, and *F. proliferatum* isolates. The primer developed on the "core" chromosome 2 (P4) resulted in amplification of DNA from two of the characterized non-pathogenic isolates (72 and 81). Of the primer pairs

Dobbs et al. BMC Genomics   (2020) 21:764

Page 6 of 15

Table 2 *Fusarium oxysporum* f. sp. *koae* and *F. oxyporum* isolates that have been screened for pathogenicity through greenhouse virulence assays at the Hawai'i Agriculture Research Center

Isolate ID	Virulence	% Mortality*	Collection Site (Island)
44	High	95	Hawai'i
78	High	92	Kauai
77	High	80	Kauai
79	High	72	Kauai
0540 K	Moderate	70	Hawai'i
90	Moderate	68	Oahu
20	Moderate	66	Hawai'i
17	Moderate	62	Hawai'i
166	Moderate	60	Hawai'i
76	Moderate	60	Kauai
8	Moderate	58	Hawai'i
34	Low	12	Hawai'i
72	Nonpathogenic	8	Kauai
27	Nonpathogenic	8	Maui
53	Nonpathogenic	4	Maui
81	Nonpathogenic	4	Kauai
170	Nonpathogenic	0	Maui

*% mortality of 24 inoculated seedlings per isolate displaying foliar wilt, chlorosis, and/or necrotic symptoms (Dudley et al. 2007; Dudley et al. 2017)

that did not amplify non-target *Fusarium* species, 2 of primer pairs P4 ("core") and P6 (LSX) were individually subjected to BLAST against the NCBI database to further test primers in silico for specificity (Additional File 5A). The "core" chromosome primer pair P4 had BLAST hits to other *Fusarium* spp. P4 forward and reverse primers had hits of 20 out 20 nucleotides to three genomes of *F. xyrophilum*, and 15 out of 20 nucleotides to *F. fujikuroi* and *F. graminearum* genomes, indicating this genomic region is likely found in other *Fusarium* spp. The LSX primer pair P6 also had partial BLAST hits to other *Fusarium* spp. in GenBank, including *F. venenatum*, *F. fujikuroi*, and *F. graminearum*, and the forward and reverse primers had 19 and 11 out of 20 nucleotides hits, respectively, to *F. culmorum* (Additional File 5B). Therefore, further testing of these primers, P4 and P6, on a broader suite of *Fusarium* spp. is warranted.

Field isolate collection
To test the specificity of primers P4 and P6 on a novel population of *Fusarium* isolates, root samples were collected from symptomatic koa trees. Root samples from 13 of the 14 sites surveyed in Hawai i yielded *Fusarium* isolates. A total of 359 *Fusarium* isolates, including *F. oxysporum*, *F. solani*, *F. proliferatum*, *F. concolor*, *F.*

lateritium, *F. commune*, and *F. fujikuroi* (based on *tef1* sequence data; GenBank Accession numbers and locations in Additional File 11), were collected from 100 symptomatic trees (70.4% of surveyed symptomatic trees) across the three islands (Additional File 6). Most of the *Fusarium* isolates were obtained from Maunawili, Kalopa, Wood Valley, and Pa'auilo, respectively. The least represented sites were Kona, Wung Ranch, and Kaiwiki, respectively.
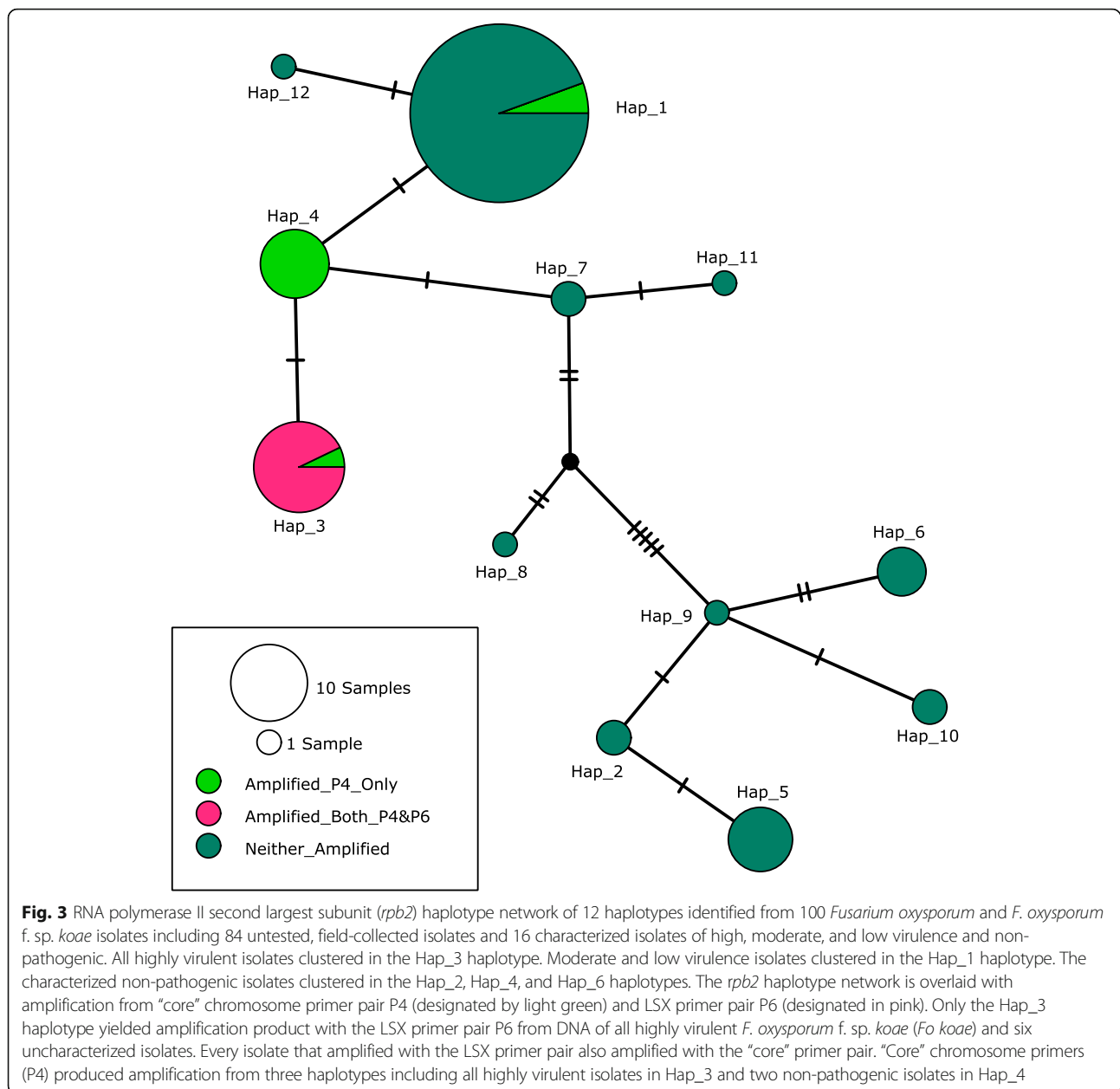
Genetic characterization of field-collected isolates
To genetically characterize field-collected isolates with known virulence isolates, the *rpb2* was sequenced for isolates from individual trees. Sequencing of *rpb2* was conducted on 100 isolates, including 16 characterized for pathogenicity (4 highly virulent, 6 moderately virulent, 1 low-virulence, and 5 non-pathogenic isolates) and 84 field-collected isolates (Additional File 7). Variation at the *rpb2* locus was used to construct a TCS haplotype network that resulted in 12 haplotypes (Fig. 3).

The haplotype network of the *rpb2* revealed that the characterized highly virulent isolates (**44**, 77, 78, 79, 90, and 166) belonged to the same haplotype, Hap_3 (Fig. 3). Moderately virulent isolates clustered within haplotypes Hap_1 and Hap_3, while the low-virulence isolate clustered in Hap_1. Non-pathogenic isolates clustered within four haplotypes, Hap_1, Hap_2, Hap_4, and Hap_6. The developed primer pair P4, developed to amplify DNA from "core" chromosome 2, amplified DNA from 3 of the 12 *rpb2* haplotypes (Hap_1, Hap_3, and Hap_4), whereas the developed primer pair P6, developed to amplify LSX DNA, only amplified DNA from one haplotype (Hap_3) (Fig. 3). P4 produced amplification of DNA from highly virulent and moderately virulent isolates; however, P4 also amplified DNA from two non-pathogenic isolates (72 and 81) in Hap_4. DNA from uncharacterized isolates in Hap_1, Hap_3, and Hap_4 was amplified using the P4 primer pairs. The P6 primer pair did not produce amplification using DNA from non-pathogenic or lowly virulent isolates. The P6 primer pair amplified DNA from all but one uncharacterized isolate in Hap_3.

Discussion
Plant pathogenic fusaria have been assessed for interspecific genetic variation to identify genomic features important for plant host infection [1, 2]. In addition, intraspecific genetic variation is a powerful tool to identify potential mechanisms for disease development by pathogenic fungi [31, 32]. This study presents a whole genome comparison between two *F. oxysporum* isolates collected from koa roots, one pathogenic (*Fo koae* 44) and one non-pathogenic (*Fo* 170) isolates. Through whole genome sequencing, genomic features were assessed bioinformatically for features that

Dobbs *et al. BMC Genomics*   (2020) 21:764

Page 7 of 15

Fig. 3 RNA polymerase II second largest subunit (*rpb2*) haplotype network of 12 haplotypes identified from 100 *Fusarium oxysporum* and *F. oxysporum* f. sp. *koae* isolates including 84 untested, field-collected isolates and 16 characterized isolates of high, moderate, and low virulence and non-pathogenic. All highly virulent isolates clustered in the Hap_3 haplotype. Moderate and low virulence isolates clustered in the Hap_1 haplotype. The characterized non-pathogenic isolates clustered in the Hap_2, Hap_4, and Hap_6 haplotypes. The *rpb2* haplotype network is overlaid with amplification from "core" chromosome primer pair P4 (designated by light green) and LSX primer pair P6 (designated in pink). Only the Hap_3 haplotype yielded amplification product with the LSX primer pair P6 from DNA of all highly virulent *F. oxysporum* f. sp. *koae* (Fo koae) and six uncharacterized isolates. Every isolate that amplified with the LSX primer pair also amplified with the "core" primer pair. "Core" chromosome primers (P4) produced amplification from three haplotypes including all highly virulent isolates in Hap_3 and two non-pathogenic isolates in Hap_4

distinguish the pathogen from the non-pathogen. Similar to other intraspecific comparison studies, the two genomes in this study were very similar [3]. They shared 96% average nucleotide identity (ANIm), indicating that they belong to the same species [33]. The constructed whole-genome phylogeny also confirmed that these isolates were among the *F. oxysporum* species complex (Additional File 1). Synteny analysis identified that the putative chromosomes in both isolates were syntenic amongst the 11 shared chromosomes of the well-characterized *F. oxysporum* f. sp. *lycopersici* isolate [18]. Analysis of the syntenic chromosomes coincided with a recent study that observed higher recombination at the telomeres and amongst smaller

chromosomes, while larger chromosomes were more conserved [28]. A non-syntenic "chromosome" was identified in both genomes, which is likely where most of the genes necessary to differentiate the koa pathogen are localized (Fig. 1b). Genome size has been shown to be a putative indicator of pathogenicity in filamentous fungi. Compared to their non-pathogenic counterparts, pathogens generally utilize larger genomes, which is associated with the presence of conditionally dispensable, lineage-specific accessory chromosomes [19]. These accessory chromosomes are known to be ca. 2 Mb, and they frequently house effectors necessary for disease development or evasion of host detection [4]. Despite these reports, this study found that the

non-pathogen (*Fo* 170) had a slightly larger genome compared to the pathogen (*Fo koae* 44) at 50 and 48 Mb, respectively. Though *Fo* 170 had a larger genome, this study identified unique genes and DNA sequences, most notably the well-characterized secreted in xylem effectors (*SIX* genes), that were exclusive to *Fo koae* 44. Most of these unique genomic features of *Fo koae* 44 were identified amongst the LSX.

As *F. oxysporum* does not have an observed teleomorphic life stage, it relies on asexual means of recombination. As such, conditionally dispensable, accessory chromosomes, and horizontal gene transfer have been described as being responsible for the host-specificity of the various formae speciales (ff. spp.) [19, 34, 35]. The predominant variation between *F. oxysporum* ff. spp. has been found in effector profiles on these conditionally dispensable, accessory chromosomes [16]. Previous studies have suggested that repeats and transposons on the accessory chromosomes increase the occurrence of point mutations, which expands diversity in virulence genes through higher mutation and duplication rates via chromosomal rearrangements and altered gene expression through chromatin restructuring [19, 35]. The identified LSX of *Fo koae* 44 in this study was transposon-rich and contained important wilt-inducing effector genes, *SIX1* and *SIX6*, providing evidence of one or more pathogen-specific, accessory chromosome(s). Interestingly, the identified LSX of *Fo* 170 was also transposon-rich, but it contained genes associated with root disease including the *PEP2* gene. The LSX identified here likely represent a combination of multiple lineage-specific chromosomes, but due to the methodologies used in this study, the LSX likely also contains assembly contigs that did not map to the reference, but could still belong on core chromosomes. However, a portion of DNA found within the LSX likely represents one or more lineage-specific chromosome(s) for both *Fo koae* 44 and *Fo* 170. Future karyotyping studies could determine presence and quantity of accessory chromosomes. Virulence-associated genes within the LSX are likely the determining factors that distinguishes pathogenicity of *Fo koae* 44 on koa.

Even though 1353 and 2369 predicted proteins were identified on the LSX of *Fo koae* 44 and *Fo* 170, respectively, the function of 58 and 50% of genes for *Fo koae* 44 and *Fo* 170, respectively, could not be determined from the available databases used for this study (Additional File 8). Perhaps these genes are highly specific to *Fo koae* 44 and *Fo* 170 and are possibly important for pathogenicity but have yet to be described. Since we did not conduct transcriptomic analyses in this study, we could not validate expression of these genes. Amongst the identified genes, most virulence-associated genes, secondary metabolites, and carbohydrate active enzymes were found to be similar for *Fo koae* 44 and *Fo* 170. Several well-characterized, virulence-associated genes

(e.g., *Gpabc1, GzmetE, Tom1, Ace1, Mocdip1, Ftf2,* and *SGE1*) were identified in both *Fo koae* 44 and *Fo* 170. Interestingly, these genes suggest that *Fo koae* 44 and *Fo* 170 both have the capacity for host recognition and disease development, even though *Fo* 170 does not cause wilt disease on koa [36].

Of the 57 and 59 unique predicted genes from the LSX of *Fo koae* 44 and *Fo* 170, respectively, the major distinctions found in *Fo koae* 44 were a mitogen-activated protein kinase pathway (associated with genes *Fmk1, Ste11,* and *Ste7*), which were previously identified as important for host penetration and pathogenicity, and the *SIX1* and *SIX6* [37]. *SIX* genes have been well-characterized as important host-specific pathogenicity factors in various *F. oxysporum* formae speciales [15, 18]. *SIX1* and *SIX6* genes together have been reported in *F. oxysporum* f. sp. *lycopersici, F. oxysporum* f. sp. *melonis, F. oxysporum* f. sp. *niveum, F. oxysporum* f. sp. *canariensis, F. oxysporum* f. sp. *phaseoli* and *F. oxysporum* f. sp. *cubense* [3, 16, 20, 38–40]. With the banana pathogen, *F. oxysporum* f. sp. *cubense* (*Fo cubense*), tropical race 1 and tropical race 4 can be distinguished by *SIX* genes [41, 42]. The same *SIX1* and *SIX6* genes reported here in *Fo koae* 44 have been previously reported in *Fo cubense* race 1 [5]. *SIX* genes are surrounded by repeat- and transposon-rich regions that increase the chances of mutation and duplication of these important effectors [19]. Our results concur with these previous observations because we observed the localization of the two *SIX* genes amongst the transposon-rich LSX of *Fo koae* 44 [29]. These *SIX* genes, in conjunction with their associated transcription factor (*SGE1*) and the mitogen-activated protein kinase pathway, may be key characteristics that contribute to pathogenicity on koa.

Although *Fo* 170 lacked *SIX* genes, a stronger putative association was observed for secondary metabolite production and transport. The identification of the trehalose, putrescine, cyclodipeptide, and monooxygenase genes in *Fo* 170 signified that toxin production may be more important for disease development on a potential non-koa host of *Fo* 170. Interestingly, another secreted peptide, PEP2 (PHI:224), was identified as exclusive to *Fo* 170, and it has been found to be important for development of root rot on pea. These results also suggest that while *Fo* 170 does not cause koa wilt disease, it may be a pathogen on another host. To differentiate and monitor the koa wilt pathogen from its non-pathogenic counterparts, pathogen-specific primers were developed.

Pathogen-specific primers have previously been shown to be effective at detecting and monitoring pathogenic strains of fungi [43–47], and they are powerful tools in disease management, including disease resistance-breeding programs. Similar to recent studies, developed primers were based on unique non-coding sequences and exclusive

predicted genes of *Fo koae* 44 [43–45]. To reduce the risk of false positive amplification of other *F. oxysporum* formae speciales isolated from koa, we chose to develop primers that would amplify genes unique to *Fo koae* 44 that did not have identified gene ontology terms.

Due to the importance of lineage-specific DNA and the virulence-associated genes found within, the LSX was targeted for primer development to detect the presence of the lineage-specific DNA that contains putative genes necessary for development of koa wilt disease [3, 16]. We observed that primer pair P6 developed for the LSX was more reliable to distinguish pathogenic isolates than unique genes identified on the "core" chromosomes, because no non-target amplification of DNA from characterized non-pathogenic isolates was observed with the LSX primer pair (P6). Of the 12 identified *rpb2* haplotypes, primer pair P4 resulted in target amplification of three haplotypes (Hap_1, Hap_3, and Hap_4), but primer pair P6 produced target amplification of only one haplotype (Hap_3). These results indicate that isolates in haplotype Hap_3 may possess some putative lineage-specific DNA necessary for development of koa wilt disease. DNA of haplotypes that only amplified with primer pairs P4, may contain a homologous gene in their "core" chromosomes that may not be required for virulence that is also shared with other *Fusarium* spp.

The LSX primer pair (P6) resulted in amplification of DNA from only one *rpb2* haplotype which suggests that it may be too specific to amplify all highly virulent isolates within the *Fo koae* population; however, primer pair P6 produced amplified product for all pathogenic *Fo koae* isolates used in HARC's disease resistance breeding program. The seven *Fo koae* isolates used in HARC's disease resistance breeding program were collected on Hawai'i (**44** and 166), Kauai (76, 77, 78, and 79), and Oahu (90); DNA from all of these isolates was amplified with both developed primer pairs [23, 48]. For the ongoing greenhouse pathogenicity trials, these highly virulent isolates (**44**, 77, 78, and 79) and three moderately virulent isolates (76, 90, and 166) are combined evenly by weight and used in combination as a soil inoculum [23]. This methodology (mixing of inoculum from multiple isolates), may allow anastamoses and/or asexual genetic recombination to occur between isolates. Hypothetically, the putative lineage-specific DNA used to develop our primers could have been transferred horizontally among isolates during the inoculation trials conducted by HARC [7], which could have contributed to the observed amplification of a single haplotype with the developed P6 primer pair. Targeted studies of the *Fo koae* population are needed to better understand the role of horizontal gene transfer in pathogenicity.

Future studies should include greenhouse virulence assays to characterize field-collected isolates that amplified

with both the core primer pair (P4) and the LSX primer pair (P6). Population genomic analyses are needed to determine genetic relationships among pathogenic *Fo koae* and non-pathogenic *F. oxysporum* isolates across the Hawaiian archipelago. Such studies can determine if a direct comparison of one pathogen and one non-pathogen is sufficient for developing a primer to detect pathogenic *Fo koae*, or if whole genome sequences of multiple haplotypes are needed to develop more robust, pathogen-specific primers. In addition, karyotyping could validate the presence and quantity of the putative lineage-specific chromosome(s) of *Fo koae* 44.

Conclusion

This study used whole genome analyses to predict virulence-associated genes that may be necessary for the development of koa wilt disease by *Fo koae*. This information was used to develop PCR primers for distinguishing highly virulent *Fo koae* isolates from non-pathogenic *F. oxysporum* isolates. Previously, the *SIX1* and *SIX6* genes have been characterized in other pathogens for their role in causing wilt diseases in other hosts, and it appears that these genes are also necessary for *Fo koae* to cause koa wilt disease. Genomic comparisons between *Fo koae* 44 and *Fo* 170 identified *Fo* 170-exclusive genes, including pathogenicity-related genes, which suggests that *Fo* 170 may be pathogenic to a different host without necessarily causing a wilt disease. Putative lineage-specific chromosomes were identified, but further research is needed to characterize and quantify these chromosomes. Pathogen-specific primers were developed that only amplified DNA of characterized, pathogenic isolates of *Fo koae*.

Because virulence assays have not yet been conducted on *Fo koae* haplotype isolates from the field collection, further research is needed to assess the robustness of the developed primers and determine if the predicted virulence-associated genes are required for pathogenicity. These results provide the framework for understanding the pathogen genes necessary for development of koa wilt disease and determining genetic variation of *Fo koae* populations across the Hawaiian Islands.

Methods

Greenhouse-characterized pathogenic and non-pathogenic isolates of *Fusarium oxysporum*

For this study, *Fusarium oxysporum* isolates that were previously screened for pathogenicity through greenhouse trials were provided by the Hawai'i Agriculture Research Center (HARC) [48] (Table 2). These isolates were screened at HARC using methods described in Dudley et al. [48]. At 90 days post-inoculation of koa seedlings under greenhouse conditions, isolates were classified

as non-pathogenic (< 10% mortality), low virulence (10–30% mortality), moderate virulence (30–70% mortality), or high virulence (70–100% mortality) (Table 2).

Twelve characterized pathogenic *F. oxysporum* isolates of high (**44**, 77, 78, and 79), moderate (8, 17, 20, 76, 90, 166, and 0540 K), and low (34) virulence, and five non-pathogenic (27, 53, 72, 81, and **170**) *F. oxysporum* isolates were used for reference in later analyses. Isolates 44, 76, 77, 78, 79, 90, and 166 are currently used in HARC's disease resistance screening trials. Pathogenic isolate 44 and non-pathogenic isolate 170 were selected for whole genome sequencing.

Whole genome sequencing and assembly

In preparation for DNA extraction, one isolate pathogenic to koa (*Fo koae* 44) and one non-pathogenic isolate (*Fo* 170) were first grown on ¼-strength PDA at 25 °C for 3 days. Hyphal tips were transferred to 100 ml of potato dextrose broth and shaken at 70 rpm for 7 days at room temperature. Mycelium was collected using a Buchner funnel and filtered through a Whatman® No. 1 filter paper under vacuum. Mycelium was placed in 2-ml tubes and frozen at 75 °C for 24 h prior to extraction. DNA was extracted using a cetyl trimethyl ammonium bromide (CTAB) extraction protocol adapted from Cubero et al. [49]. Tissue preparation was modified from the original protocol in that 0.1 g of frozen tissue was pulverized using a FastPrep-24^{Tm} (M.P. Biomedicals LLC, Santa Ana, CA, USA) at 5x speed for three 20-s runs. Between runs, samples were kept frozen using liquid nitrogen. On the third run, 750 μl of CTAB extraction buffer (1% w/v CTAB; 1 M NaC1; 100 mM Tris; 20 mM EDTA; 1% w/v polyvinyl polypyrolidone) was added to each sample. Samples were incubated for 30 min at 70 °C in a heat block. One volume of chloroform:isoamyl alcohol (24:1) was added to each sample and mixed on a shaker for 20 min then centrifuged at 10,000 x g for 5 min. Upper aqueous phase (600 μl) was transferred to a new 2-ml tube. Two volumes of Precipitation buffer (two volumes; 1200 μl) was added to sample, which was mixed in a shaker for 5 min, then centrifuged at 13,000 x g for 15 min. Supernatant was removed and re-suspended in 350 μl 1.5 M NaCl. To each sample, 2 μl of 10 mg/ml RNase was added, followed by incubation at 37 °C for 30 min. One volume of chloroform:isoamyl was added to each sample and mixed on a shaker for 20 min, then centrifuged at 10,000 x g for 5 min. Upper phase was transferred to a new 1.5-ml tube and 0.6 volume of ice-cold isopropanol was added to each sample. Samples were mixed by inversion and incubated at − 20 °C for 1 h. After incubation, samples were centrifuged at 14,000 x g for 5 min at room temperature. Supernatant was removed and 1 ml 70% ethanol was added to each sample. After 1-min incubation at room temperature, samples were centrifuged

at 14,000 x g for 3 min at room temperature. Pellet was dried and resuspended in 50 μl TE buffer (10 mM Tris-HCl, 1 mM EDTA). Extracted DNA was gel electrophoresed in a 2% agarose gel and quantified using a Qubit™ fluorometer (Invitrogen, Carlsbad, CA, USA). A Zymo-clean™ Gel DNA Recovery Kit (Zymo Research, Irvine CA, USA) and a Zymo clean and concentrate column were used to recover and purify the DNA. Extracted DNA was sent to Macrogen (Seoul, South Korea) for TruSeq Illumina shotgun sequencing with reads of 2x151bp.

De novo genome assemblies of both *Fo koae* 44 and *Fo* 170 were constructed using SPAdes 3.11.1 [50] using default parameters. The SPAdes genome assembly was assessed for quality using QUAST [51]. Contigs with than 100x coverage were removed, and the remaining contigs were checked for similarity with *Fusarium* spp. sequences using the National Center for Biotechnology Information (NCBI) BLAST to screen for presence of bacterial DNA in samples.

Phylogenetic analysis

A whole-genome, maximum likelihood phylogeny was constructed to elucidate evolutionary relationships among *Fo koae* 44, *Fo* 170, other *Fusarium* spp., and other *F. oxysporum* formae speciales. REALPHY 112 [52] was used to generate whole genome alignments including the *F. oxysporum* f. sp. *lycopersici* reference genome and *Magnaporthe oryzae* (Both obtained from GenBank; Additional File 9). PhyML 3.1 [53] using bowtie2 2.3.4.2 [54] and the general time reversal (GTR) substitution model [55], was used to construct the whole genome phylogeny with bootstrap support (pseudoreplicates = 200) and *M. oryzae* as the outgroup. Whole genomes of *Fusarium* spp. and *F. oxysporum* formae speciales were obtained from the NCBI GenBank database (Additional File 9).

Genome annotation and analysis

To determine overall sequence similarity of *Fo koae* 44 and *Fo* 170 isolates, average nucleotide identity was calculated using Pyani [56]. R package 'coRdon' was used to determine codon usage bias [57, 58]. Transposable-ELMT, a wrapper script for transposable element identification, and creation of a comprehensive repeat library using RepeatModeler, RepeatClassifier, LTR_finder, ltr_harvest, and TransposonPSI (http://transposonpsi.sourceforge.net) and subsequently RepeatMasker with this library were used to identify transposable elements and repeat-rich regions for both *Fo koae* 44 and *Fo* 170 [59–61].

To identify unique sequences of *Fo koae* 44 and *Fo* 170, putative chromosomes were constructed using a closely related reference genome, *F. oxysporum* f. sp. *lycopersici*. Paired reads of the genomes of *Fo koae* 44 and *Fo* 170 were trimmed using BBDuk (decontamination using kmers; part of BBTools package) [62]. The trimmed reads

Dobbs *et al. BMC Genomics*   (2020) 21:764

Page 11 of 15

were mapped to the reference genome using Geneious (Geneious Prime 2019.2, http://www.geneious.com/) progressiveMauve genome mapper software at medium-low sensitivity [63, 64]. Low and high coverage sites were removed using Geneious' built-in program. Variants and SNPs were identified using FreeBayes [65] within Geneious. Variant calling format (VCF) was exported for use in comparison of SNPs and indels between *Fo koae* 44 and *Fo* 170. Synteny maps and MUMmer alignments were made using the synteny mapping and analysis program (SyMAP) [66] to create putative chromosome contigs based on the reference genome.

The de novo assembled genomes were annotated for both *Fo koae* 44 and *Fo* 170 using the MAKER 2.31.8 annotation pipeline [67] with RepeatMasker 4.0.8 [60]. A *F. oxysporum*-specific repeat library was constructed using RepeatModeler 1.0.11 [68] to mask interspersed repeats and low complexity DNA sequences. Three gene predictors were used in the pipeline: GeneMark-ES [69], SNAP [70], and AUGUSTUS [71]. *Fusarium graminearum* was used as a species model for AUGUSTUS. To identify tRNA genes, tRNAscan-SE 1.3.1 was used with default settings.

Predicted transcripts (≥150 bp) and proteins (≥50 amino acids) were analyzed using five databases for analysis of putative genes: OrthoVenn2 for putative non-orthologous proteins [72], InterProScan for protein family domains [36], antiSMASH for putative secondary metabolites [73], dbCAN2 for putative carbohydrate-active enzymes [30], and PHI-base for putative virulence-associated proteins [74]. OrthoVenn2, antiSMASH, and dbCAN2 databases' online servers were used to analyze the predicted genes at default settings. A local BLAST was used for InterProScan and the PHI-base database. Selection of putative proteins from PHI-base was based on a ≥ 85% grade (metric of combining the query coverage, e-value, and identity values for each hit with weights of 0.5, 0.25 and 0.25, respectively). Putative virulence-associated proteins were tested for exclusivity to *Fo koae* 44 by aligning these sequences to sequences obtained from the NCBI database using MUSCLE in Geneious [75] (Additional File 10).

Primer development

Putative chromosomes from SyMAP were aligned using progressiveMauve [63] in Geneious. Using these alignments, unique sequences of *Fo koae* 44 were identified and selected. Also, predicted transcripts specific to *Fo koae* 44 were identified by comparing the predicted transcriptomes of *Fo koae* 44 and *Fo* 170 from the Maker annotation identified from Orthovenn2. *Fo koae* 44 predicted transcripts were further identified as unique by using a local BLAST against *Fo* 170 and the NCBI database. Sequences with no sequence similarity to sequences from *Fo* 170, other *Fusarium* spp., and *F. oxysporum* formae speciales were used to develop primer pairs. Primer3 [76] was used to develop PCR primers. *Fo*

koae 44 unique sequences and predicted transcripts were used to generate ca. ten primer sets (forward and reverse primers) for each sequence and transcript.

Primers were tested for specificity using template DNA from four characterized highly virulent *Fo koae* isolates (isolates used in HARC greenhouse screening trials: **44**, 77, 78, and 79), seven characterized moderately virulent *Fo koae* isolates (HARC isolates 8, 17, 20, 76, 90, 166, and 0540 K), two characterized non-pathogenic *Fo* isolates (HARC isolates 27 and **170**), two *F. commune* isolates (HARC isolate 85 and Stewart collection FO21), and one *F. proliferatum* isolate (HARC isolate 1). Each 25-ul reaction contained 10 ng template DNA or sterile, molecular grade water for the negative control. Primers were tested using a Mastercycler ProS thermocycler (Eppendorf, Hamburg, Germany) at a program of 94 °C for 2 min, 30 cycles of 94 °C for 40 s, 59 °C for 40 s, and 72 °C for 30 s, and 72 °C for 5 min. PCR products for each primer set and template DNA were electrophoresed using a standard 1.5% agarose gel at 60 V for 60 min, and visualized using GelRed® (Biotium, Freemont, CA, USA).

Field collection of *Fusarium* spp.

To assemble an expanded population of *F. oxysporum* isolates for *Fo koae* specific-primer tests, 14 collection sites were selected based on a previous survey conducted by HARC from 2004 to 2007 across four of the main Hawaiian Islands, including Kauai, Oahu, Maui, and Hawai'i [77]. Sites for new isolate collections were selected based sites on where previously characterized pathogen isolates of moderate to high virulence were collected. In addition, selected collection sites were geographically distant from each other to increase the likelihood of collecting a more genetically diverse pathogen population. On Kauai, the three collection sites included Kokee, Kapa'a, and Hanalei (Princeville). On Oahu, three collection sites included Maunawili, Poamoho, and Kahana. On Hawai'i, eight collection sites included Wood Valley, Pu'u wa'a wa'a, Kona, Wung Ranch, Kaiwiki, Kalopa, Paauilo, and Volcano National Park. Root samples were collected from symptomatic trees at each site. Symptomatic segments of 4- to 5-mm thick roots exhibiting discoloration and black streaking, were targeted for isolation of *Fusarium* spp. Five fine roots and/or pencil-thick roots were plated on selective Komada media [78] for isolation of *Fusarium* spp. Plates were incubated at 25 °C for 3 days then checked for mycelial growth emanating from the root samples. Hyphal tips were taken from the radiating mycelium and transferred to ¼-strength PDA and incubated at 25 °C for another 3 days. Colony morphology was used to determine if the isolate was *Fusarium*. To further characterize these isolates to species, isolates were sequenced at the translation elongation factor 1 alpha (*tef1*) region using a Mastercycler ProS thermocycler and a PCR

cycle program of 94 °C for 2 min, 40 cycles of 94 °C for 40 s, 58 °C for 40 s, and 72 °C for 30 s, and 72 °C for 5 min [79]. Products were run on a 1.5% agarose gel to visualize amplified PCR product using GelRed®.

Genetic characterization of field-collected isolates

To determine the genetic relationships among the field-collected isolates and the previously characterized pathogenic and non-pathogenic isolates, *F. oxysporum* isolates, based of *tef1* data. Isolates were selected from separate, individual trees. The selected isolates were then sequenced at the RNA polymerase II second largest subunit (*rpb2*) [80] using previous methods above for *tef1*. *Rpb2* sequences of both characterized and field-collected isolates were aligned using MUSCLE alignment in Geneious [75]. A statistical parsimony (TCS) haplotype network was constructed based on this *rpb2* sequence data using PopART with a 95% genetic cut-off [81]. These isolates were tested with the developed *Fo koae*-specific primers to identify haplotypes that may contain the putative lineage-specific DNA identified from *Fo koae* 44.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12864-020-07156-y.

Additional file 1. Whole genome maximum likelihood phylogeny based on *Fusarium* spp. and *F. oxysporum* (*Fo*) formae speciales. Pathogenic isolate *F. oxysporum* f. sp. *koae* (*Fo koae* 44) indicated in red and non-pathogenic isolate *F. oxysporum* (*Fo* 170) indicated in blue. Bootstrap = 100.

Additional file 2. Putative biological, molecular, and cellular function of the predicted non-orthologous proteins identified as unique to the pathogenic isolate of *Fusarium oxysporum* f. sp. *koae* (*Fo koae* 44) when compared to non-pathogenic isolate of *F. oxysporum* (*Fo* 170). Identified function is based on gene ontology (GO) terms.

Additional file 3. Putative biological process and molecular function of the predicted non-orthologous proteins identified as unique to the non-pathogenic *Fusarium oxysporum* isolate (*Fo* 170) when compared to the pathogenic isolate of *F. oxysporum* f. sp. *koae* (*Fo koae* 44). Function is based on gene ontology (GO) terms.

Additional file 4. Copy number of secondary metabolite genes including nonribosomal peptide synthetases (NRPS), type 1 and type 3 polyketide synthases (T1PKS and T3PKS), terpenes, and indoles for the pathogenic *Fusarium oxysporum* f. sp. *koae* isolate (*Fo koae* 44, red bars) and non-pathogenic *F. oxysporum* isolate (*Fo* 170, blue bars).

Additional file 5. A) Pathogen-specific PCR primers designed from unique genes and sequences in genome of *Fusarium oxysporum* f. sp. *koae* 44. Isolates were confirmed for pathogen specificity through PCR testing on greenhouse-characterized pathogenic and non-pathogenic isolates provided by the Hawai'i Agriculture Research Center. **B)** NCBI BLAST results of primer pairs P4 (designed to amplify "core" chromosome 2) and P6 (designed to amplify the putative lineage-specific DNA, LSX, of *Fo koae* 44).

Additional file 6. Results of 2018 field collection showing number of *Fusarium* isolates collected from symptomatic *Acacia koa*.

Additional file 7. Distribution and virulence ratings for 15 characterized and 85 uncharacterized *Fusarium oxysporum* isolates collected from *Acacia koa* roots used for haplotype and primer amplification analysis.

Additional file 8. Proportion of predicted genes with putative function of the lineage specific DNA (LSX) of the *Acacia koa* wilt pathogen *Fusarium oxysporum* f. sp. *koae* (*Fo koae* 44) and the *F. oxysporum* (*Fo* 170) isolate characterized as non-pathogenic to *A. koa*. Both isolates had similar proportions of transposons, repeats, carbohydrate active enzymes (CAZymes), and genes identitifed through gene ontology terms and protein family as having a biological, molecular, or cellular function. *Fo koae* 44 had a higher proportion of virulence associated genes and *Fo* 170 had a higher proportion of secondary metabolite genes on the LSX.

Additional file 9. Whole genome sequences of *Fusarium* spp. and formae speciales of *F. oxysporum* retrieved from NCBI with GenBank or RefSeq accession numbers. These genomes were used to make the whole genome phylogeny.

Additional file 10. Secreted in xylem (*SIX*) gene sequences of formae speciales of *Fusarium oxysporum* (*Fo*) retrieved from NCBI with GenBank accession numbers. Sequences used to compare identified *SIX* genes in pathogenic isolate *F. oxysporum* f. sp. *koae* 44.

Additional file 11. GenBank accession numbers and collection locations for all isolates used in this study.

Abbreviations

Fo koae: *Fusarium oxysporum* f. sp. *koae*; *Fo*: *Fusarium oxysporum*; koa: *Acacia koa*; *rpb2*: RNA polymerase II second largest subunit; *tef1*: Translation elongation factor 1 alpha; HARC: Hawai'i Agriculture Research Center; LSX: Putative lineage-specific DNA; DNA: Deoxyribonucleic acid; Hap: Haplotype; *SIX*: Secreted in xylem; ANIm: Average nucleotide identity; GO: Gene ontology; NRPS: Non-ribosomal peptide synthetases; ABC: ATP-binding cassette; PCR: Polymerase chain reaction; CTAB: Cetyl trimethyl ammonium bromide; GTR: General time reversal; NCBI: National Center for Biotechnology Information; SNP: Single nucleotide polymorphism; VCF: Variant calling format; SyMAP: Synteny mapping and analysis program

Availability of data and materials

Cultures used in this study are stored at Colorado State University Stewart laboratory culture collection and are available upon request. All sequences have been deposited into National Center for Biotechnology Information (NCBI) BioPoject, BioSample, Nucleotide and GenBank and accession numbers are listed in Additional File 5B and Additional File 11. Site information including GPS locations, where available, are listed in Additional File 7 and Additional File 11 for each isolate. GenBank accession numbers of isolates used for comparison of secreted in xylem genes are listed in Additional File 10. GenBank accession numbers for reference genomes and genomes used for whole genome comparisons are listed in Additional File 9.

Dobbs *et al. BMC Genomics*   (2020) 21:764

Page 13 of 15

Author details
[1]Colorado State University, Department of Agricultural Biology, 1177 Campus Delivery, Fort Collins, CO 80523, USA. [2]USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA. [3]Hawai'i Agriculture Research Center, Maunawili Research Station, Oahu, HI, USA. [4]USDA Forest Service, Rocky Mountain Research Station, 1221 South Main Street, Moscow, ID 83843, USA. [5]Division of Forestry and Wildlife, Department of Land and Natural Resources, 1151 Punchbowl Street, Room 325, Honolulu, HI 96813, USA. [6]USDA Forest Service, Forest Health Protection, 1323 Club Drive, Vallejo, CA 94592, USA.

References
1. Sperschneider J, Gardiner DM, Thatcher LF, Lyons R, Singh KB, Manners JM, Taylor JM. Genome-wide analysis in three Fusarium pathogens identifies rapidly evolving chromosomes and genes associated with pathogenicity. Genome Biol Evol. 2015;7(6):1613–27. https://doi.org/10.1093/gbe/evv092.
2. Walkowiak S, Rowland O, Rodrigue N, Subramaniam R. Whole genome sequencing and comparative genomics of closely related Fusarium head blight fungi: Fusarium graminearum, F. meridionale and F. asiaticum. BMC Genomics. 2016;17(1):1014. https://doi.org/10.1186/s12864-016-3371-1.
3. Armitage AD, Taylor A, Sobczyk MK, Baxter L, Greenfield BPJ, Bates HJ, Wilson F, Jackson AC, Ott S, Harrison RJ, Clarkson JP. Characterization of pathogen-specific regions and novel effector candidates in Fusarium oxysporum f. sp. cepae. Scientific reports. 2018;8(1):13530. https://doi.org/10.1038/s41598-018-30335-7.
4. Bertazzoni S, Williams AH, Jones DA, Syme RA, Tan K-C, Hane JK. Accessories make the outfit: accessory chromosomes and other dispensable DNA regions in plant-pathogenic fungi. Mol Plant-Microbe Interact. 2018;31(8):779–88. https://doi.org/10.1094/MPMI-06-17-0135-FI.
5. Deltour P, França SC, Heyman L, Pereira OL, Höfte M. Comparative analysis of pathogenic and nonpathogenic Fusarium oxysporum populations associated with banana on a farm in Minas Gerais, Brazil. Plant Pathol. 2018;67(3):707–18. https://doi.org/10.1111/ppa.12757.
6. Liu X, Xing M, Kong C, Fang Z, Yang L, Zhang Y, Wang Y, Ling J, Yang Y, Lv H. Genetic diversity, virulence, race profiling, and comparative genomic analysis of the Fusarium oxysporum f. sp. conglutinans strains infecting cabbages in China. Frontiers in. Microbiology. 2019;10:1373. https://doi.org/10.3389/fmicb.2019.01373.
7. Li J, Fokkens L, Conneely LJ, Rep M. A partial pathogenicity chromosome in Fusarium oxysporum is sufficient to cause disease and can be horizontally transferred [preprint]. Microbiology. 2020. https://doi.org/10.1101/2020.01.20.912550.
8. Jiang C, Zhang X, Liu H, Xu J-R. Mitogen-activated protein kinase signaling in plant pathogenic fungi. PLoS Pathog. 2018;14(3):e1006875. https://doi.org/10.1371/journal.ppat.1006875.
9. Ponukumati SV, Elliott ML, Des Jardin EA. Comparison of secreted in xylem (SIX) genes in two Fusarium wilt pathogens of ornamental palms. Plant Pathol. 2019;68(9):1663–81. https://doi.org/10.1111/ppa.13090.
10. Taylor A, Vágány V, Jackson AC, Harrison RJ, Rainoni A, Clarkson JP. Identification of pathogenicity-related genes in Fusarium oxysporum f. sp. cepae: pathogenicity in Fusarium oxysporum f. sp. cepae. Mol Plant Pathol. 2016;17(7):1032–47. https://doi.org/10.1111/mpp.12346.
11. Thatcher LF, Gardiner DM, Kazan K, Manners JM. A highly conserved effector in Fusarium oxysporum is required for full virulence on Arabidopsis. Mol Plant-Microbe Interact. 2012;25(2):180–90. https://doi.org/10.1094/MPMI-08-11-0212.
12. Roy A, Jayaprakash A, Rajeswary TR, Annamalai A, Lakshmi P. Genome-wide annotation, comparison and functional genomics of carbohydrate-active enzymes in legumes infecting Fusarium oxysporum formae speciales. Mycology. 2020;11(1):56–70. https://doi.org/10.1080/21501203.2019.1706656.
13. Guo L, Han L, Yang L, Zeng H, Fan D, Zhu Y, Feng Y, Wang G, Peng C, Jiang X, Zhou D, Ni P, Liang C, Liu L, Wang J, Mao C, Fang X, Peng M, Huang J. Genome and transcriptome analysis of the fungal pathogen Fusarium oxysporum f. sp. cubense causing banana vascular wilt disease. PLoS ONE. 2014;9(4):e95543. https://doi.org/10.1371/journal.pone.0095543.
14. Chang W, Li H, Chen H, Qiao F, Zeng H. Identification of mimp-associated effector genes in Fusarium oxysporum f. sp. cubense race 1 and race 4 and virulence confirmation of a candidate effector gene. Microbiological Research. 2020;232:126375. https://doi.org/10.1016/j.micres.2019.126375.
15. Schmidt SM, Houterman PM, Schreiver I, Ma L, Amyotte S, Chellappan B, Boeren S, Takken FLW, Rep M. MITEs in the promoters of effector genes allow prediction of novel virulence genes in Fusarium oxysporum. BMC Genomics. 2013;14(1):119. https://doi.org/10.1186/1471-2164-14-119.
16. van Dam P, Fokkens L, Schmidt SM, Linmans JHJ, Kistler HC, Ma L-J, Rep M. Effector profiles distinguish formae speciales of Fusarium oxysporum: effector profiles distinguish formae speciales of Fo. Environ Microbiol. 2016;18(11):4087–102. https://doi.org/10.1111/1462-2920.13445.
17. Gordon TR. Fusarium oxysporum and the Fusarium wilt syndrome. Annu Rev Phytopathol. 2017;55(1):23–39. https://doi.org/10.1146/annurev-phyto-080615-095919.
18. Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim W-B, Woloshuk C, Xie X, Xu J-R, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RAE, Chapman S, Coulson R, Coutinho PM, Danchin EGJ, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Hua-Van A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee Y-H, Li L, Manners J, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park S-Y, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC, Turgeon BG, Wapinski I, Yoder O, Young S, Zeng Q, Zhou S, Galagan J, Cuomo CA, Kistler HC, Rep M. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature. 2010;464(7287):367–73. https://doi.org/10.1038/nature08850.
19. Raffaele S, Kamoun S. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol. 2012;10(6):417–30. https://doi.org/10.1038/nrmicro2790.
20. van Dam P, Fokkens L, Ayukawa Y, van der Gragt M, ter Horst A, Brankovics B, Houterman PM, Arie T, Rep M. A mobile pathogenicity chromosome in Fusarium oxysporum for infection of multiple cucurbit species. Sci Rep. 2017;7(1):9042. https://doi.org/10.1038/s41598-017-07995-y.
21. Anderson, R. C., Gardner, D. E., Daehler, C. C., & Meinzer, F. C. (2002). Dieback of Acacia koa in Hawaii: ecological and pathological characteristics of affected stands. For Ecol Manag, 162(2–3), 273–286. https://doi.org/10.1016/S0378-1127(01)00522-9.
22. Gardner DE. Acacia koa seedling wilt caused by Fusarium oxysporum f. sp. koae, f. sp. nov. Phytopathology. 1980;70:594–7.
23. Dudley NS, Jones TC, James RL, Sniezko RA, Cannon P, Borthakur D. Applied disease screening and selection program for resistance to vascular wilt in Hawaiian Acacia koa. Southern Forests: A Journal of Forest Science. 2015;77(1):65–73. https://doi.org/10.2989/20702620.2015.1007263.
24. Dudley, N., Jones, T., James, R., Sniezko, R., Wright, J., Liang, C., Gugger, P. F., & Cannon, P. (2017). Applied genetic conservation of Hawaiian Acacia koa: An eco-regional approach. In: Sniezko, Richard A.; Man, Gary; Hipkins, Valerie; Woeste, Keith; Gwaze, David; Kliejunas, John T.; McTeague, Brianna A., tech. cords. 2017. Gene conservation of tree species—banking on the future. Proceedings of a workshop. Gen. Tech. Rep. PNW-GTR-963. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station, 963, 78–91.
25. Dudley, N., James, R., Sniezko, R., Cannon, P., Yeh, A., Jones, T., & Kaufmann, M. (2012). Operational disease screening program for resistance to wilt in Acacia koa in Hawaii. In: Sniezko, Richard A.; Yanchuk, Alvin D.; Kliejunas, John T.; Palmieri, Katharine M.; Alexander, Janice M.; Frankel, Susan J., tech. coords. Proceedings of the fourth international workshop on the genetics of host-parasite interactions in forestry: Disease and insect resistance in forest trees. Gen. Tech. Rep. PSW-GTR-240. Albany, CA: Pacific Southwest Research Station, Forest Service, US Department of Agriculture, 240, 286–289.

Dobbs *et al. BMC Genomics*   (2020) 21:764

Page 14 of 15

26. Ma L-J, Geiser DM, Proctor RH, Rooney AP, O'Donnell K, Trail F, Gardiner DM, Manners JM, Kazan K. Fusarium pathogenomics. Annu Rev Microbiol. 2013; 67(1):399–416. https://doi.org/10.1146/annurev-micro-092412-155650.

27. Takken F, Rep M. The arms race between tomato and Fusarium oxysporum. Mol Plant Pathol. 2010;11(2):309–14. https://doi.org/10.1111/j.1364-3703. 2009.00605.x.

28. Fokkens L, Shahi S, Connolly LR, Stam R, Schmidt SM, Smith KM, Freitag M, Rep M. The multi-speed genome of Fusarium oxysporum reveals association of histone modifications with sequence divergence and footprints of past horizontal chromosome transfer events [preprint]. Genomics. 2018. https:// doi.org/10.1101/465070.

29. Suga H, Hirayama Y, Morishima M, Suzuki T, Kageyama K, Hyakumachi M. Development of PCR primers to identify Fusarium oxysporum f. sp. fragariae. Plant Dis. 2013;97(5):619–25. https://doi.org/10.1094/PDIS-07-12-0663-RE.

30. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. dbCAN2: A meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2018;46(1):95–101. https://doi.org/10.1093/ nar/gky418.

31. Carmona SL, Burbano-David D, Gómez MR, Lopez W, Ceballos N, Castaño-Zapata J, Simbaqueba J, Soto-Suárez M. Characterization of pathogenic and nonpathogenic Fusarium oxysporum isolates associated with commercial tomato crops in the Andean region of Colombia. Pathogens. 2020;9(1):70. https://doi.org/10.3390/pathogens9010070.

32. Seo S, Pokhrel A, Coleman JJ. The genome sequence of five genotypes of Fusarium oxysporum f. sp. vasinfectum: A resource for studies on Fusarium wilt of cotton. Mol Plant-Microbe Interact. 2020;33(2):138–40. https://doi.org/ 10.1094/MPMI-07-19-0197-A.

33. Figueras MJ, Beaz-Hidalgo R, Hossain MJ, Liles MR. Taxonomic affiliation of new genomes should be verified using average nucleotide identity and multilocus phylogenetic analysis. Genome Announcements. 2014;2(6): e00927–14, 2/6/e00927-14. https://doi.org/10.1128/genomeA.00927-14.

34. Mehrabi R, Mirzadi Gohari A, Kema GHJ. Karyotype variability in plant-pathogenic fungi. Annu Rev Phytopathol. 2017;55(1):483–503. https://doi.org/10.1146/annurev-phyto-080615-095928.

35. Rep M, Van Der Does HC, Meijer M, Van Wijk R, Houterman PM, Dekker HL, De Koster CG, Cornelissen BJC. A small, cysteine-rich protein secreted by Fusarium oxysporum during colonization of xylem vessels is required for I-3-mediated resistance in tomato: cysteine-rich protein required for I-3-mediated resistance. Mol Microbiol. 2004;53(5):1373–83. https://doi.org/10. 1111/j.1365-2958.2004.04177.x.

36. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. InterProScan: protein domains identifier. Nucleic Acids Res. 2005;33(web server):W116–20. https://doi.org/10.1093/nar/gki442.

37. Di Pietro A, Madrid MP, Caracuel Z, Delgado-Jarana J, Roncero MIG. Fusarium oxysporum: exploring the molecular arsenal of a vascular wilt fungus. Mol Plant Pathol. 2003;4(5):315–25. https://doi.org/10.1046/j.1364-3703.2003.00180.x.

38. Covey PA, Kuwitzky B, Hanson M, Webb KM. Multilocus analysis using putative fungal effectors to describe a population of Fusarium oxysporum from sugar beet. Phytopathology. 2014;104(8):886–96. https://doi.org/10. 1094/PHYTO-09-13-0248-R.

39. Meldrum RA, Fraser-Smith S, Tran-Nguyen LTT, Daly AM, Aitken EAB. Presence of putative pathogenicity genes in isolates of Fusarium oxysporum f. sp. cubense from Australia. Australas Plant Pathol. 2012;41(5):551–7. https://doi.org/10.1007/s13313-012-0122-x.

40. Niño-Sánchez J, Casado-Del Castillo V, Tello V, De Vega-Bartol JJ, Ramos B, Sukno SA, Díaz Mínguez JM. The FTF gene family regulates virulence and expression of SIX effectors in Fusarium oxysporum: the FTF genomic expansion of Fusarium oxysporum. Mol Plant Pathol. 2016;17(7):1124–39. https://doi.org/10.1111/mpp.12373.

41. Czislowski E, Fraser-Smith S, Zander M, O'Neill WT, Meldrum RA, Tran-Nguyen LTT, Batley J, Aitken EAB. Investigation of the diversity of effector genes in the banana pathogen, Fusarium oxysporum f. sp. cubense, reveals evidence of horizontal gene transfer: effector genes in Fusarium oxysporum f. sp. cubense. Mol Plant Pathol. 2018;19(5):1155–71. https://doi.org/10.1111/mpp.12594.

42. Fraser-Smith S, Czislowski E, Meldrum RA, Zander M, O'Neill W, Balali GR, Aitken EAB. Sequence variation in the putative effector gene SIX8 facilitates molecular differentiation of Fusarium oxysporum f. sp. cubense. Plant Pathol. 2014;63(5):1044–52. https://doi.org/10.1111/ppa.12184.

43. Burkhardt A, Henry PM, Koike ST, Gordon TR, Martin F. Detection of Fusarium oxysporum f. sp. fragariae from infected strawberry plants. Plant Dis. 2019;103(5):1006–13. https://doi.org/10.1094/PDIS-08-18-1315-RE.

44. Kashiwa T, Inami K, Teraoka T, Komatsu K, Arie T. Detection of cabbage yellows fungus Fusarium oxysporum f. sp. conglutinans in soil by PCR and real-time PCR. J Gen Plant Pathol. 2016;82(5):240–7. https://doi.org/10.1007/ s10327-016-0668-5.

45. Kim H, Hwang S-M, Lee JH, Oh M, Han JW, Choi GJ. Specific PCR detection of Fusarium oxysporum f. sp. raphani: A causal agent of Fusarium wilt on radish plants. Lett Appl Microbiol. 2017;65(2):133–40. https://doi.org/10. 1111/lam.12761.

46. Stewart JE, Abdo Z, Dumroese RK, Klopfenstein NB, Kim M-S. Virulence of Fusarium oxysporum and F. commune to Douglas-fir (Pseudotsuga menziesii) seedlings: virulence of Fusarium oxysporum and F. commune. For Pathol. 2012;42(3):220–8. https://doi.org/10.1111/j.1439-0329.2011.00746.x.

47. van Dam P, de Sain M, ter Horst A, van der Gragt M, Rep M. Use of comparative genomics-based markers for discrimination of host specificity in Fusarium oxysporum. Appl Environ Microbiol. 2017;84(1):e01868–17. https://doi.org/10.1128/AEM.01868-17.

48. Dudley NS, James RL, Sniezko RA, Yeh A. Investigating koa wilt and dieback in Hawaii: pathogenicity of Fusarium species on Acacia koa seedlings. Native Plants J. 2007;8(3):259–66.

49. Cubero OF, Crespo A, Bridge P. DNA extraction and PCR amplification method suitable for fresh, herbarium-stored, lichenized, and other fungi. Plant Systematic Evolution. 1999;216:243–9.

50. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77. https://doi.org/10.1089/cmb.2012.0021.

51. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072–5. https://doi.org/10. 1093/bioinformatics/btt086.

52. Bertels F, Silander OK, Pachkov M, Rainey PB, van Nimwegen E. Automated reconstruction of whole-genome phylogenies from short-sequence reads. Mol Biol Evol. 2014;31(5):1077–88. https://doi.org/10.1093/molbev/msu088.

53. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 2003;52(5):696–704. https://doi.org/10.1080/10635150390235520.

54. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9(4):357–9. https://doi.org/10.1038/nmeth.1923.

55. Tavaré S. In: Miura RM, editor. Providence, RI: Amer math Soc Some probabilistic and statistical problems in the analysis of DNA sequences. Some mathematical questions in biology - DNA sequence analysis; 1986. p. 57–86.

56. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal Methods. 2016;8(1):12–24. https://doi.org/10.1039/ C5AY02550H.

57. Fabijanić M, Vlahoviček K. Big data, evolution, and Metagenomes: predicting disease from gut microbiota codon usage profiles. In: Carugo O, Eisenhaber F, editors. Data mining techniques for the life sciences. Methods in molecular biology, 1415. New York, NY: Humana press; 2016. https://doi.org/ 10.1007/978-1-4939-3572-7_26.

58. Core Team R. R: A language and environment for statistical computing. In: R Foundation for statistical computing. Vienna: Austria. URL; 2019. https:// www.R-project.org/.

59. Ellinghaus, D., Kurtz, S., & Willhoef, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. 9(1), 18.

60. Smit, A. F. A., Hubley, R., & Green, P. (2013-2015). RepeatMasker Open-4.0. http://www.repeatmasker.org.

61. Wyka SA. TransposableELMT - wrapper script to create a comprehensive repeat library for newly sequenced organisms. (version v1.0.0). Zenodo; 2019. https://doi.org/10.5281/zenodo.3469661.

62. Bushnell, B. (n.d.). BBMap—Sourceforge.net/projects/bbmap/.

63. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE. 2010;5(6): e11147. https://doi.org/10.1371/journal.pone.0011147.

64. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis

of sequence data. Bioinformatics. 2012;28(12):1647–9. https://doi.org/10.1093/bioinformatics/bts199.

65. Garrsion, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. https://doi.org/arXiv preprint arXiv:1207.3907 [q-bio.GN].

66. Soderlund C, Bomhoff M, Nelson WM. SyMAP v3.4: A turnkey synteny system with application to plant genomes. Nucleic Acids Res. 2011;39(10):e68. https://doi.org/10.1093/nar/gkr123.

67. Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sanchez Alvarado A, Yandell M. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 2007;18(1):188–96. https://doi.org/10.1101/gr.6743907.

68. Smit, A. F. A., & Hubley, R. (2008-2015). RepeatModeler Open 1.0. http://repeatmodeler.org.

69. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. 2008;18(12):1979–90. https://doi.org/10.1101/gr.081612.108.

70. Zaharia, M., Bolosky, W. J., Curtis, K., Fox, A., Patterson, D., Shenker, S., Stoica, I., Karp, R. M., & Sittler, T. (2011). Faster and more accurate sequence alignment with SNAP. http://arxiv.org/abs/1111.5572.

71. Keller O, Kollmar M, Stanke M, Waack S. A novel hybrid gene prediction method employing protein multiple sequence alignments. Bioinformatics. 2011;27(6):757–63. https://doi.org/10.1093/bioinformatics/btr010.

72. Xu L, Dong Z, Fang L, Luo Y, Wei Z, Guo H, Zhang G, Gu YQ, Coleman-Derr D, Xia Q, Wang Y. OrthoVenn2: A web server for whole-genome comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res. 2019;47(1):52–8. https://doi.org/10.1093/nar/gkz333.

73. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 2019;47(W1):W81–7. https://doi.org/10.1093/nar/gkz310.

74. Urban M, Cuzick A, Rutherford K, Irvine A, Pedro H, Pant R, Sadanadan V, Khamari L, Billal S, Mohanty S, Hammond-Kosack KE. PHI-base: A new interface and further additions for the multi-species pathogen–host interactions database. Nucleic Acids Res. 2017;45(D1):D604–10. https://doi.org/10.1093/nar/gkw1089.

75. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. https://doi.org/10.1093/nar/gkh340.

76. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. Primer3—new capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115. https://doi.org/10.1093/nar/gks596.

77. Dudley NS, James RL, Cannon P, Yeh AK, Jones TC, Kaufmann MK. Developing resistant koa-early results, from disease survey to seedling resistance testing in Hawai'i. In: Proceedings of the 7th Meeting of IUFRO Working Party 7.03.04 Diseases and Insects in Forest Nurseries; 2009; Hilo, HI; 2010. p. 39–47.

78. Komada H. Development of a selective medium for quantitative isolation of Fusarium oxysporum from natural soil. Review Plant Protection Research. 1975;8:114–25.

79. O'Donnell K, Sutton DA, Rinaldi MG, Magnon KC, Cox PA, Revankar SG, Sanche S, Geiser DM, Juba JH, van Burik J-AH, Padhye AA, Anaissie EJ, Francesconi A, Walsh TJ, Robinson JS. Genetic diversity of human pathogenic members of the Fusarium oxysporum complex inferred from multilocus DNA sequence data and amplified fragment length polymorphism analyses: evidence for the recent dispersion of a geographically widespread clonal lineage and nosocomial origin. J Clin Microbiol. 2004;42:5109–20.

80. Liu YJ, Whelen S, Hall BD. Phylogenetic relationships among ascomycetes: evidence from an RNA polymerse II subunit. Mol Biol Evol. 1999;16(12):1799–808.

81. Clement M, Posada D, Crandall K. TCS: a computer program to estimate gene genealogies, vol. 10; 2000. p. 1657–9.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.