

中国松材线虫种群遗传分析和疫源追溯研究

简雯婧

(南京林业大学 江苏南京 210037)

摘要: 由松材线虫 (*Bursaphelenchus xylophilus*) 引起的松材线虫病(pine wilt disease)是一种国际性重大森林病害, 中国是松材线虫病流行危害最严重的地区。至 2024 年初我国已有 18 个省(自治区、直辖市)的 663 个县级行政区沦为松材线虫病疫区。由于松材线虫的入侵时间和入侵地环境的不同, 导致我国松材线虫种群在遗传结构上形成了许多不同的地理株系。本研究主要是利用全基因组数据分析我国不同疫区松材线虫的种群遗传多样性, 根据其遗传差异来探究我国各个疫区松材线虫的种群遗传结构, 并建立松材线虫的遗传追溯体系, 来追溯松材线虫的传播来源和扩散路径, 为我国松材线虫的检疫、防控工作提供科学依据。主要研究结果如下:

1、本实验室自 2014 年到 2024 年实验室的虫株资源库中共收集保存 701 株松材线虫虫株资源。对 701 株虫株的重测序数据进行分析, 共鉴定到 9,508,393 个 SNPs 位点。其中国外虫株的 SNP 总数、纯合子数、特有 SNPs 数、转换和颠换数量均高于国内大部分虫株。各省份间不同虫株的 SNPs 数量均存在明显的差异, SNPs 的数量最多的是广东, 其次是吉林, 四川省 SC、湖北 HB、浙江 ZJ、江西 JX、湖南 HN、云南 YN、广西 GX 的虫株平均数均大于 20 万, 而福建 FJ、贵州 GZ、安徽 AH、甘肃 GS、陕西 SX、江苏 JS、河南 HEN、重庆 CQ、辽宁 LN、山东 SD 的虫株的 SNPs 数相对较小, 均小于 10 万, 其中最小的福建省; 纯合子数与 SNPs 数情况相同, 依旧是广东省虫株的纯合子数最多; 而特有 SNPs 数中, 最多的是云南省的虫株, 有 128,356, 广东、湖北、四川、江西、广西的虫株的特有 SNPs 数也相对较多。SNP 基因型种类共有 12 种: A→C, A→G, A→T, C→A, C→G, C→T, G→A, G→C, G→T, T→A, T→C 和 T→G。其中, A→G, C→T, G→A 和 T→C 这四种类型的 SNP 要显著高于其他 8 个类型的 SNP 基因型。

2、对 701 个松材线虫虫株的 SNP 位点进行主成分分析 (PCA) 和系统发育树分析, 可将其归为 8 个类群: 类群 1 共计 9 个虫株, 为日本虫株 (C14-5), 美国虫株 (USA01-06) 和 2 株湖北虫株 (HB36、HB37); 类群 2 共计 41 个虫株, 包括广东虫株 37 株、湖南虫株 3 株和湖北虫株 1 株; 类群 3 共有 117 个虫株, 包括广东虫株 22 株, 湖南省虫株 1 株, 广西壮族自治区 33 株, 湖北省 16 株, 河南省虫株 1 株, 江西省虫株 11 株, 辽宁省虫株 2 株; 云南省虫株 4 株, 浙江省虫株 27 株; 类群 4 共计 47 株, 分别来自广东省的 3 株, 广西壮族自治区 9 株, 湖北省的 5 株, 浙江省的 8 株, 吉林省的 7 株, 四川省的 14 株以及山东省的 1 株; 类群 5 共计有虫株 100 株, 包括广东省的 45 株, 广西壮族自治区的 7 株, 湖北省的 1 株, 江西省的 25 株, 辽宁省的 1 株, 浙江省 16 株, 四川省的 3 株, 江苏省的 2 株; 类群 6 共计有 167 株虫株, 广东省的 13 株, 广西壮族自治区的 11 株, 湖北省虫株 42 株, 河南省虫株 2 株, 江西省虫株 3 株, 辽宁省虫株 9 株, 浙江省虫株 30 株, 山东省 16 株, 四川省虫株 15 株, 陕西省虫株 2 株, 江苏省虫株 12 株, 安徽省虫株 5 株, 福建省虫株 3 株, 重庆市 3 株以及甘肃省 1 株; 类群 7 共计有虫株 102 株, 有广东省的 1 株, 河南省的 3 株, 广西壮族自治区的 2 株, 湖北省虫株 3 株, 河南省虫株 11 株, 江西省虫株 2 株, 辽宁省虫株 15 株, 浙江省虫株 1 株, 四川省虫株 3 株, 陕西省虫株 12 株, 江苏省虫株 32 株, 安徽省虫株

14 株，福建省虫株 3 株；类群 8 共计有 118 株虫株，包含广东省的 8 各虫株，湖南省的 5 个虫株，广西壮族自治区的 3 个虫株，湖北省的 13 个虫株，河南省的 2 个虫株，江西省的 8 个虫株，辽宁省的 1 个虫株，云南省的 1 个虫株，浙江省的 20 个虫株，山东省 7 株，四川省 8 株，陕西省 5 株，江苏省 16 株，安徽省 6 株，福建省 6 株，重庆市 5 株，贵州省 4 株。

3、华中地区 30 个虫株中共获得 8,333,375 个 SNPs 位点，所有 SNP 位点统计分析显示，HB（湖北）15、HEN（河南）20、HN（湖南）07、HN08 和 HN10 的 SNP 位点数量与纯合子数量均显著高于其他虫株。此外，在 HB15、HEN20、HN07、HN08 和 HN10 虫株中出现频率最高的突变基因型为 A>G、C>T、G>A、T>C，而其他虫株则为 A>G、C>G、C>T、G>A、G>C、T>C。聚类结果表明，华中地区 30 株松材线虫可以分为 4 个类群：HB15、HEN20、HN07、HN08 和 HN10 为同一类群；HB02、HEN02、HEN04 和 HN06 为类群 2；类群 3 虫株最多，共有 16 个虫株；类群 4 全部为河南虫株（HEN06、HEN09、HEN10、HEN14 和 HEN15）。与广东省和江苏省虫株进行种群遗传差异分析发现，HB15、HEN20、HN07、HN08 和 HN10 与广东省虫株具有亲缘关系，其他虫株与江苏省虫株的遗传距离相对较近，其中河南省虫株间遗传差异显著。华中地区虫株之间存在明显的遗传分化现象，推测华中地区大部分疫区和江苏省的松材线虫具有相同的疫源，少部分地区和广东省虫株具相同的疫源。

4、本研究通过对中国两广地区的松材线虫进行遗传多样性分析，发现两广地区的大部分虫株与中国另一个松材线虫病传播中心——江苏省的虫株的遗传距离很远，两广地区的虫株仅有一部分与江苏省的虫株聚为一类。两广地区的虫株的遗传多样性普遍比江苏省的虫株高，且广东省的松材线虫的遗传多样性最丰富。广东省具有自己的传播源，推测来自国外；而部分的广东和广西的虫株可能来自江苏。选择性清除结果表明，两广地区的类群 B 的虫株的候选基因富集在果胶裂解酶活性等相关基因上，猜测该类群具有更强的毒力，可能危害更多的松树。需要防治松材线虫通过人为途径传播到非疫区，以危害更多松树。本研究为探究两广地区的松材线虫的生存空间、遗传结构、潜在传播空间和制定相关防疫政策提供理论依据。

5、在前人构建的松材线虫疫源追溯数据库（PWNSD），具有数据存储、检索、浏览、系统发育树构建、序列比对以及引物设计等功能。提供的 Web 界面，将松材线虫基因组序列以及相关变异位点的基因和蛋白序列等信息进行可视化，提供更直观的数据展示。选取新发疫区甘肃省的虫株对系统功能的实现进行测试，结果与第三章遗传多样性分析结果一致，证实实现情况良好。目前库内包含了实验室组装的基因组数据以及 701 个松材线虫虫株的相关数据。

关键词：松材线虫；SNP 分子标记；遗传差异；种群遗传学；选择性清除分析