ARPHA Conference Abstracts 5: e87349
doi: 10.3897/aca.5.e87349

Conference Abstract

Full chloroplast sequencing using genome skimming for novel plant DNA barcode discovery in Amaranths

Leonardo Galindo Gonzalez‡, Marie-Claude Gagnon‡, Sarah Kyte‡, Andréanne Charron‡, Adam Colville‡

‡ Canadian Food Inspection Agency, Ottawa, Canada

Corresponding author: Leonardo Galindo Gonzalez (leonardo.galindogonzalez@inspection.gc.ca)

Received: 03 Jun 2022 | Published: 17 Jun 2022

Citation: Galindo Gonzalez L, Gagnon M-C, Kyte S, Charron A, Colville A (2022) Full chloroplast sequencing using genome skimming for novel plant DNA barcode discovery in Amaranths. ARPHA Conference Abstracts 5: e87349. https://doi.org/10.3897/aca.5.e87349

Abstract

DNA barcoding has been established as an efficient, sensitive and reliable methodology for plant identification. However, in spite of efforts to find a universal DNA plant barcode, some taxa are not sufficiently resolved by typical plant barcoding genes like *matK* or *rbcL.* We have used a technique known as genome skimming, which relies on the empiric low coverage sequencing of a full plant genome, resulting in high coverage of the high copy genome fractions such as chloroplast and rDNA. Phylogenetic studies show that these regions are a reservoir of variability which could be further exploited for DNA barcode discovery. We sequenced eight amaranth species using Illumina Next Generation Sequencing technology to test the feasibility of this technique. Amaranths were chosen due to their increasing impact as invasive species bearing multiple herbicide resistance mechanisms. Our results showed that complete chloroplast genomes could be assembled for all of the eight species tested. We obtained an average of 47 million reads for each one of the amaranth nuclear genomes, which range in size between 400-700Mb approximately. These reads provide an average theoretical coverage of 10-15X for each nuclear genome, but resulted in an average chloroplast genome coverage in the range of 500-8000X due to multiple chloroplast genome copies per cell. Alignment of the eight chloroplast genomes shows variability in the single copy regions (Fig. 1), especially on intergenic sections (Fig. 2

). Additional preliminary analyses also show variation among different populations of the same species, demonstrating the importance of studying both inter and intraspecific diversity to design reliable and accurate DNA barcodes that can be used in species identification.

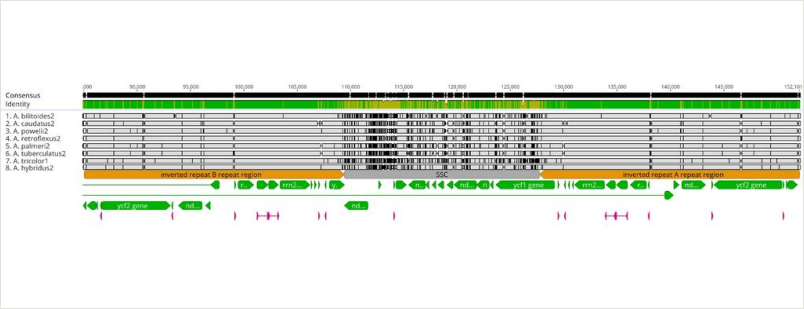

Figure 1. doi

Variability between the Small Single-Copy (SSC) region and the Inverted Repeat (IR) regions. Alignment of eight amaranth chloroplast genomes using MAFFT depicts high levels of polymorphism in the SSC. Each vertical line or gap depicted in each sequence corresponds to a polymorphism between tested species. Genes are depicted in green and tRNAs are shown in pink.

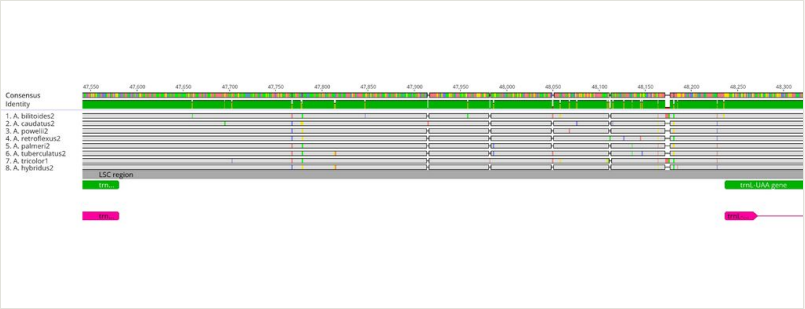

Figure 2. doi

Polymorphism of an intergenic section from the Long Single-Copy (LSC) region. Alignment of eight amaranth chloroplast genomes using MAFFT depicts a close up of a highly polymorphic intergenic region between the tRNAs for Threonine (left) and Leucine (right). Each vertical line or gap depicted in each sequence corresponds to a polymorphism between tested species (the colors correspond to different nucleotides). Genes are depicted in green and tRNAs are shown in pink.

Keywords

Genome skimming, chloroplast sequencing, Next Generation Sequencing (NGS), amaranths, DNA barcoding

Presenting author

Leonardo Galindo-González

Presented at

International Conference on DNA Barcoding and Biodiversity