Accurate and Rapid DNA Barcoding Technique for Herbal and Botanical Products

• High Accuracy   • Fast Turn Around Time   • Streamlined Processes

Jane Luo and Handy Yowanto
SCIEX, Brea, California, USA

Introduction
The global herbal supplement market is estimated to reach $107 billion dollars by 2017. Surveys indicated that commonly used medicinal plants belong to 11,146 species from 2,309 genera of 383 families. Accurate and rapid authentication of these plants and their adulterants is very important for international trade and safe use of medicinal plants and adherence to regulatory compliance. While the conventional chemical analysis can not identify the genus and species of plants; compounding to the challenge, contaminants, mislabeling and adulteration of processed raw materials could post safety risks to consumers and economic loss to manufacturers. Certain conserved regions such as internal (ITS) and external (ETS) ribosomal DNA transcribed spacer regions have been used as markers within the nuclear genome for plant population and evolution studies. In this technical note, we describe a process for sequencing ITS and ETS regions of herb samples using the GenomeLab GeXP™ Genetic Analysis System that offers fast turnaround time (8 hrs) and high identification accuracy. Upon isolation of nucleic acid from each herb sample, the gene target is amplified using polymerase chain reaction (PCR), followed by Dye Terminator Cycle Sequencing (DTCS) on the GeXP. Results are compared against the National Center for Biotechnology Information (NCBI) plant sequence database to identify the genus and species of the plant.

Materials and Methods
Genomic DNA Isolation: Dried herb samples were soaked in nuclease free water overnight and grounded with a small mortar and pestle. About 100 mg of the grounded sample was used for isolation of genomic DNA using Qiagen DNeasy Plant Mini Kit. DNA concentration was determined by a UV spectrophotometer.

Target Gene Amplification and Post PCR Cleanup: Amplification of ITS and ETS regions were performed as described by Sonnante et al. Amplicons were purified using Agencourt AMPure XP or Qiaquick Gel Extraction Kit.

Sequencing and Data Analysis: Sanger Sequencing (Figure 2) was carried out per instructions for the DTCS Quick Start Kit from SCIEX. Sequencing fragments were purified using Agencourt CleanSeq. Separation and data analysis are fully automated on the GenomeLab GeXP Genetic Analysis System (Figure 1). Sequences were submitted to NCBI Database for BLAST search to identify the herb.

Work Flow for Herb Identification: The entire work flow for herb identification using GeXP is shown in Figure 3.
Figure 2. Sanger Dye Terminator Cycle Sequencing.

Figure 3. Work Flow for Herb Identification using GenomeLab GeXP Genetic Analysis System.
Results

Identification of an herbal plant sample as Momordica charantia (bitter melon): The ITS region was amplified from the genomic DNA isolated from the sample; and sequenced using DTCS Quick Start Kit. Sequencing fragments were purified by CleanSeq and separated on the GeXP instrument (Figure 4).

After quality based trimming, sequence results were exported in .scf format; imported to GeneStudio for contig assembly to generate consensus sequence (Figure 5).

Figure 4. Sequencing results showing analyzed data, Quality Values and DNA sequence. The Quality Value (QV) is a measure of confidence for base calling. It relates to estimated error rate. The lower the error rate, the higher the Quality Value. QVs are useful in identifying the highest quality portions of reads for use in sequence assembly. Before the assembly, trimming using QV is preferred than other trimming methods since it provides longer readlength. During assembly, QV can be used to guide determining the consensus sequence.

Figure 5. Contig Assembly in GeneStudio to create consensus sequence through alignment of DNA sequences obtained in two independent sequencing reactions with the same sample. Consensus sequence was exported as a .txt file.
Consensus sequence was entered in BLASTN as "Query Sequence" for BLAST search (Figure 6A). The search result (Figure 6B) indicated that the sample is identified as "Momordica charantia" (bitter melon).

Figure 6A. Blast search in NCBI database. Consensus sequence was copy- and-pasted as "Query Sequence". "Others (nr etc.)" was selected as the database for BLAST search. The search was initiated by clicking on "BLAST" button.

Figure 6B. Blast search results showed the ITS sequence from sample is 99% identical to ITS of Momordica charantia.

Identification of an herbal plant sample as Silybum marianum (milk thistle): Both the ITS region and the ETS region of the sample were sequenced (Figure 7). BLAST search results with ITS sequence (Figure 8A) and ETS sequence (Figure 8B) both indicated that the sample is Silybum marianum (milk thistle).

Figure 7. Sequencing results of the ITS region and the ETS region.
Conclusion

DNA Sequencing targeting the ITS and ETS regions using SCIEX GenomeLab GeXP Genetic Analysis System is an accurate, simple, fast and reliable method to identify genus and species of many different herbal plants.

References