



TERMS OF REFERENCE TO CONTRACT A SERVICE PROVIDER FOR ASSEMBLY THE GENOME OF 35 WILD POTATO ACCESSIONS, DO A TRANSCRIPTOME ANALYSIS AND GENE DISCOVERY

1. Main Objective/Background:

The wild potato germplasm conserved at the International Potato Center (CIP) Genebank is important for plant breeding. These wild *Solanum* species originated in varying environments in South America from the deserts to wetlands, and thus, carry genes that are of interest to potato breeders for disease and pest resistance. Additionally, these species also confer resilience to climate change, abiotic stress tolerance, and adaptation to production in extreme environments such as the high altitudes with cold/frost. Because of this, CIP-Genebank beside doing conservation of genetic resources aims to characterize the wide diversity of wild potato species both at phenotype and genotype level. Specifically, genome sequencing of different wild potato species would promote the use and provide understanding to develop strategies for prioritizing conservation of biodiversity. For breeders and scientist genomic sequences would support the understanding of complex trait, discover new sources of genetic diversity, and develop marker and genomic-assisted selection tools to advance development of improved potato varieties.

Currently in a CIP-Genebank research project, 35 accessions of wild potato were selected to generate genome sequences through the Illumina platform at 50X using HiSeq sequencer in paired-end mode for 10 accessions and 100X paired-end for 25 accessions. Additionally, Pac-Bio's Single Molecule RS II system technology and Nanopore sequencing is also going to be generated for 5 of the accessions with 100X Illumina sequences. The genome sequences are being generated by a service request with sequencing provider and would be delivered around September. On the other hand, *Solanum acaule* is a species well documented as cold tolerant, one cold tolerant accession of this species is part of the set of sequences material. A cold treatment under growth chamber conditions is proposed to do a transcriptome analysis of leaf tissue to identify genes of interest to abiotic stress. The objective of this announcement is to contract a bioinformatics service provider with experience in complex genomes to do the genome assembly of 35 accessions of wild potato and some cultivated potato and the cold treatment and transcriptome analysis of a *S. acaule* accession.

2. Main Objective:

The bioinformatics service provider should present a genome assembly strategy with a platform where one can download the sequences, do quality control and genome assembly and technical validation of the 35 genomes as well as the methodology to do cold stress treatment and transcriptome analysis of one cold tolerant accession of *S. acaule* sequenced in this set. The results of this assembled genomes and transcriptome analysis would be used to do comparative genome sequencing to identify regions of the genome of interest to abiotic stress.

3. Scope of the project

The bioinformatics service provider would propose a methodology based on the experience to assemble 35 genomes and do the cold treatment and transcriptome analysis. The proposal should also consider the cost, time required for the analysis, and date of final report. The service provider should document the experience and technical support for the proposed methodology for the bioinformatic analysis supported in published studies.

4. Technical Specification:

At least 26 different species are being sequenced in this project and for many of them this is the first genome sequence generated for that specific species. Additionally, some hybrids with different species genetic background are included in the germplasm sequenced. Most of the species are diploids but also some are tetraploid accessions. The assembly of this type of complex genomes



WWW.
CIPOTATO.ORG
Street Address:
Av. La Molina 1895, La Molina, Lima, Perú

requires several levels of expertise, software and storage platform. We are looking for a provider with previous experience in assembly of at least 20 potato genomes, and that do *de novo* assembly following a hybrid-read method.

- a. *De novo* assembly of organellar genomes using Illumina and PacBio sequencing data
- b. *De novo* assembly of nuclear genomes using Illumina, PacBio, and Nanopore sequencing data
- c. Annotation of the *de novo* genome assemblies
- d. Structural variation and SNP analysis
- e. Transcriptome analyses of cold tolerant *S. acaule* (CIP 761464)

For the data analysis it is recommended the use of these tools: Trimmomatic v0.40, NOVOPlasty v4.3.1, BUSCO v5.1.2, and AUGUSTUS for the annotation.

The starting material for the cold tolerant accession is *in vitro* grown plants that should be transferred and grown in soil (Sogemix) in a growth chamber at 20 C day and 18 C night and 150 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light conditions. When plants reach a vegetative stage three plants should be transferred to a growth chamber at 10 C daytime and 4C night-time for 3 days.

The genomes must be loaded in The Potato Genome Diversity Portal where 13 other CIP potato genomes and their organellar genomes are published.

5. Deliverables and dates for delivery:

- a. Genome and organellar genome assembly of 35 potato accessions (CIP 761464, PI 545864, PI 275236, PI 473243, PI 545751, PI 472837, PI 245939, PI 631212, CIP762126.217, CIP762616.261, CIP590008.29, CIP762573.219, CIP762297.201, CIP762446.201, CIP760917.1, CIP761868.202, CIP760147.7, CIP761403.208, CIP761080.201, CIP761353.009, CIP760368.015, CIP761498.010, CIP761359.032, CIP762830.057, CIP761786.005, CIP761645.014, CIP762866.038, CIP762866.026, CIP762833.025, CIP512010.1, CIP392650.12, variety Wiñay, and three breeding material)
- b. Cold treatment and transcriptome analysis of *S. acaule* (CIP 761464)
- c. Publication of sequences in The Potato Genome Diversity Portal and preliminary report Dec 15, 2021
- d. Final report