

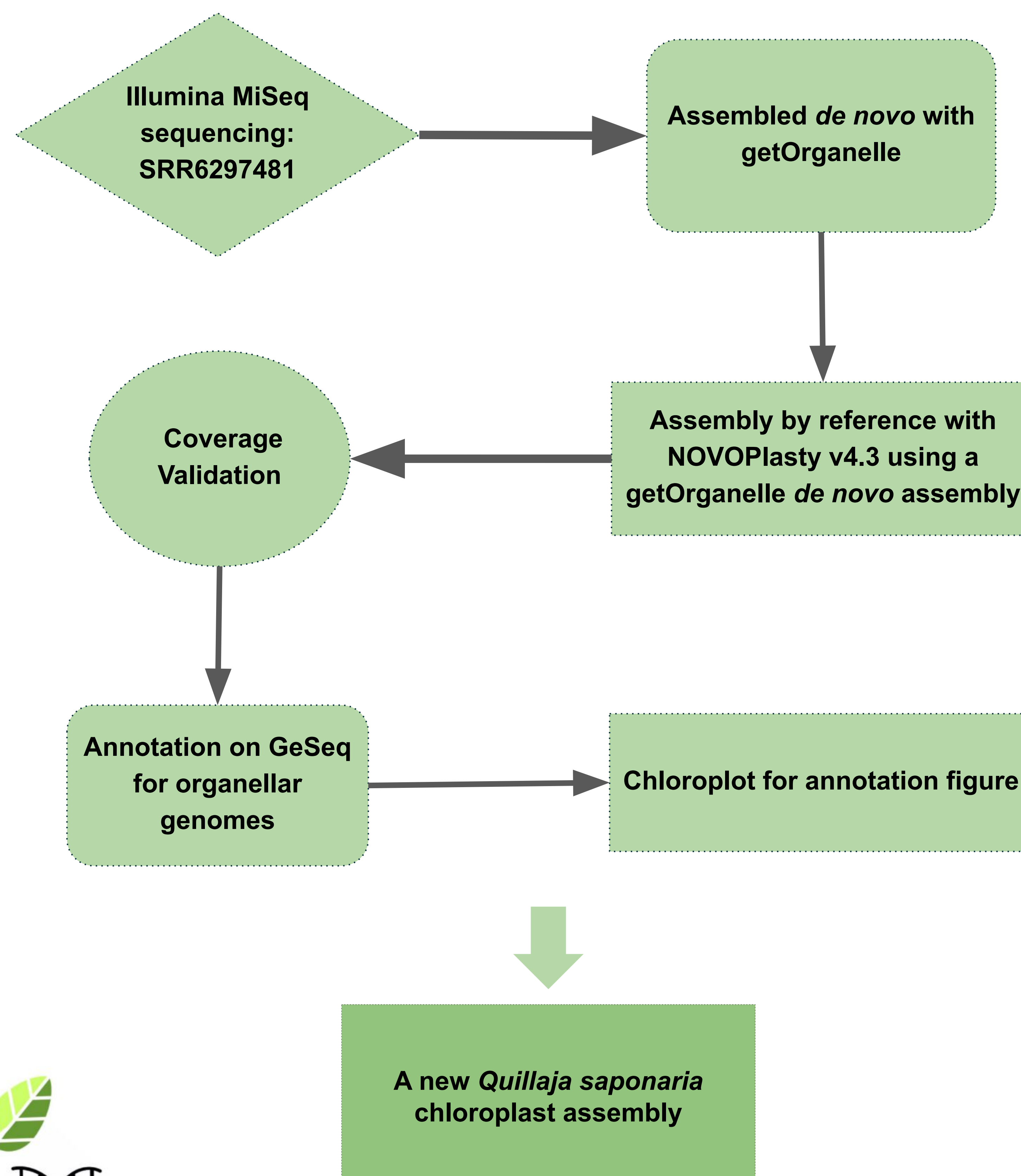
Quillaja saponaria plastome revisited: sequence update and re-annotation

Gian C. C. Barros¹, Alan Lorenzetti, Bruno Cesar Rossini², Ilara Gabriela Frasson Budzinski, Douglas Silva Domingues
¹Group of Genomics and Transcriptomes in Plants, Institute of Biosciences, São Paulo State University, UNESP,- Rio Claro, SP, Brasil
²Institute of Biotechnology, São Paulo State University, UNESP, Botucatu, SP, Brazil

INTRODUCTION

Quillaja saponaria is an important South American woody species in the production of vaccine adjuvants, from triterpene saponins produced by its specialized metabolism. Few genomic resources are available for this species. Among them is the assembly of its chloroplast genome (published in 2019;NCBI accession MH880827), showing an estimated size of 132,854 bp, considerably smaller than that identified in species of the same order, such as *Suriana maritima* (163,747 bp) and *Polygala japonica* (165,439 bp). The reduced size might indicate assembling errors due to the use of inappropriate genome assembling tools. To test our misassembly hypothesis, the objective of this study was to re-assemble and re-annotate the *Q. saponaria* chloroplast using assemblers and strategies specially developed for the recovery of plastid genomes. We intend to review the annotation and sequencing data of the *Quillaja saponaria* species as a way to obtain a new chloroplast assembly.

METHOD



RESULTS

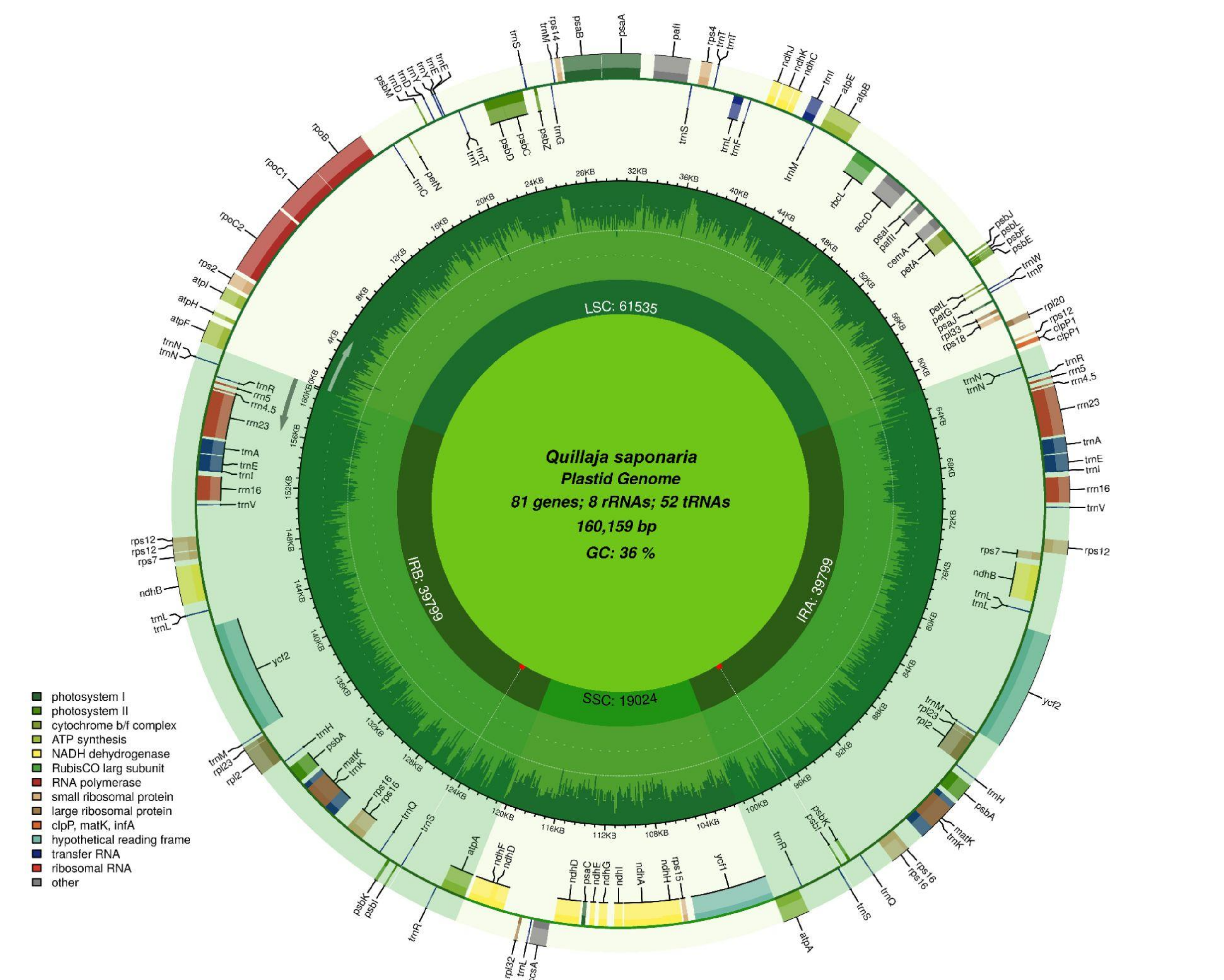


Figure 1: Complete chloroplast of the species *Q. saponaria* using our assembly approach.



Figure 2: Genome-wide sequencing coverage of *Q. saponaria* chloroplast with our assembly strategy. Most regions have between 500-3000 X coverage.

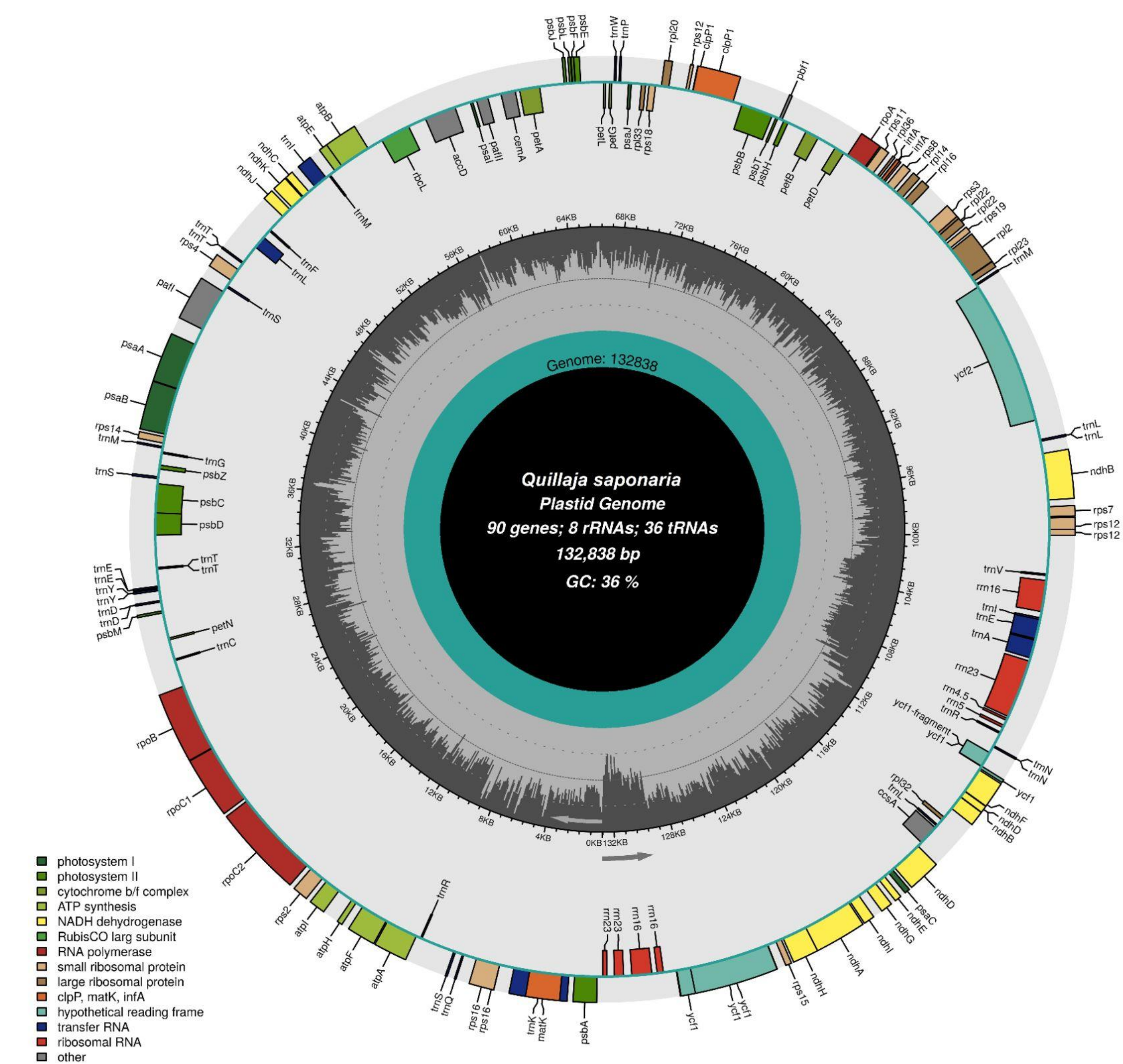


Figure 3: Sequence of *Quillaja saponaria* chloroplast MH880827.1 deposited at NCBI.

Group of genes	Genes
Subunits of ATP synthase	atpA, atpB, atpE, atpF, atpH, atpI
Subunits of photosystem II	psbA, psbC, psbD, psbE, psbF, psbI, psbJ, psbK, psbL, psbM, psbZ,
Subunits of NADH-dehydrogenase	ndhA, ndhB, ndhB, ndhC, ndhD, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK
Subunits of cytochrome b/f complex	petA, petG, petL, petN
Subunits of photosystem I	psaA, psaB, psaC, psaI, psaJ
Subunit of rubisco	rbcL
Large subunit of ribosome	rpl2, rpl2, rpl20, rpl23, rpl23, rpl33, rpl32
DNA dependent RNA polymerase	rpoB, rpoC1, rpoC2
Small subunit of ribosome	rps2, rps14, rps15, rps16, rps16, rps18, rps4, rps7, rps7, rps12, rps12
Subunit of Acetyl-CoA-carboxylase	accD
c-type cytochrom synthesis gene	ccsA
Envelop membrane protein	cemA
Maturase	matK
Conserved open reading frames	ycf1, ycf2, ycf2, ycf4 trnM(2), trnG, trnT(4), trnE(4), trnY(2), trnD(2), trnC, trrL(3), trnR, trnQ, trnK, trnH, trnV, trnI(2), trnA, trnR, trnN(2), trnF, trnS(2)
Transfer RNA	
Ribossomal RNA	rrn16, rrn23, rrn4.5, rrn5
Other proteins	clP1 (2), pfalI, pfalI

Table 1: Genes annotated in *Q. saponaria* chloroplast; in green the genes identified in our assembly approach.

A complete chloroplast sequence of 160,393 bp. From the Chlorobx tool annotation, we found 81 genes of which 52 are tRNAs and eight are rRNAs.

In the unverified chloroplast annotation, only 33 tRNAs and four rRNAs were annotated.

Our results show high similarity to NCBI accession NC_047356.1, an unannotated chloroplast assembly deposit from another specimen of *Q. saponaria*.

Mixed assembly strategies can improve assembly and annotation of the *Q. saponaria* chloroplast genome.