

The Genome Assembly and Annotation of *Chamaesiphon* sp. PCC 6505

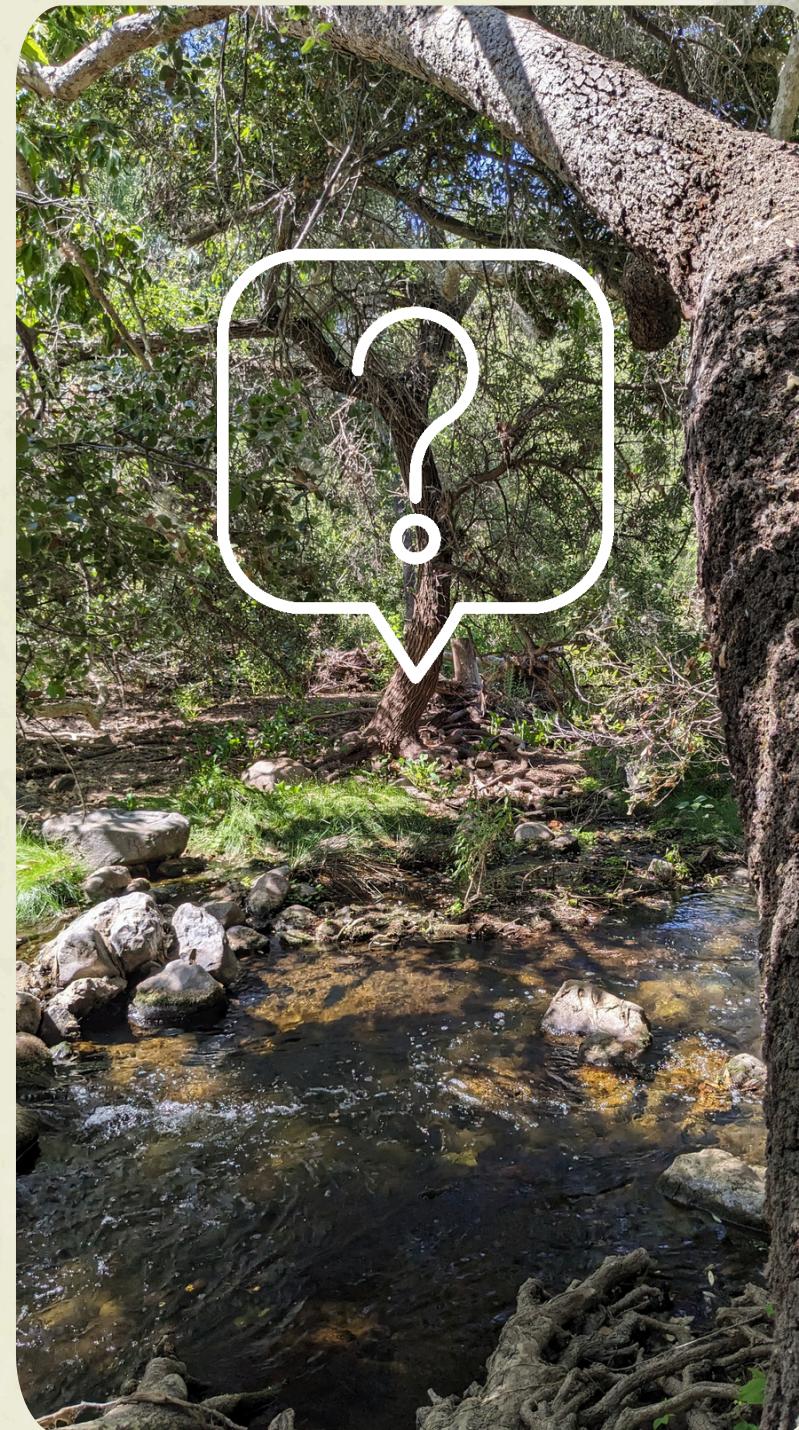
Presentation by Janelle A. Mande

TABLE OF CONTENTS

Introduction	----- 01
Methods and Results	----- 02 - 06
Conclusion	----- 07

Introduction

- During the Summer of 2023, I participated in an REU at California State University, San Marcos
- A toxin was detected at Escondido Lake
 - Hypothesized that the specimen was potentially *Chamaesiphon*
- Simple curiosity motivated me for this final project.

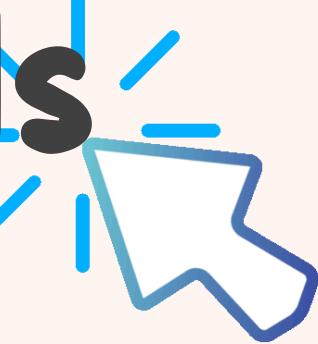


Janelle Mande



Matthews
Robin Matthews

Methods



Acquired
genomic data

Assembly
Quality Check

Conclusion

Genome
Assembly

Genome
Annotation

Acquiring Genomic Data

[SRX202155: Chamaesiphon sp. PCC 6505 library](#)

1 ILLUMINA (Illumina Genome Analyzer II) run: 30.2M spots, 4.6G bases, 2.6Gb downloads

Design: Solexa paired ends

Submitted by: DOE JOINT GENOME INSTITUTE (JGI)

Study: Chamaesiphon sp. PCC 6505 Project

- [SRP016908](#)
- [All experiments](#)
- [All runs](#)

[show Abstract](#)

Sample: Chamaesiphon sp. PCC 6505

[SAMN01797692](#) • [SRS373642](#) • [All experiments](#) • [All runs](#)

Organism: [Octopus sp.](#) HBH-6

Library:

Name: HTHC

Instrument: Illumina Genome Analyzer II

Strategy: WGS

Source: GENOMIC

Selection: RANDOM

Layout: PAIRED

Spot descriptor:



Runs: 1 run, 30.2M spots, 4.6G bases, [2.6Gb](#)

Run	# of Spots	# of Bases	Size	Published
SRR610296	30,157,315	4.6G	2.6Gb	2015-07-22

ID: 269024

Genome assembly ASM31714v1

reference

[Download](#)

datasets

curl

FTP

NCBI RefSeq assembly GCF_000317145.1

Submitted GenBank assembly GCA_000317145.1

Taxon Chamaesiphon minutus PCC 6605

Strain PCC 6605

Relation to type material assembly from type material

Submitter JGI

Date Dec 5, 2012

Acquiring Genomic Data



[SRX202155: Chamaesiphon sp. PCC 6505 library](#)
1 ILLUMINA (Illumina Genome Analyzer II) run: 30.2M spots, 4.6G bases, 2.6Gb downloads

Design: Solexa paired ends

Submitted by: DOE JOINT GENOME INSTITUTE (JGI)

Study: Chamaesiphon sp. PCC 6505 Project
• [SRP016908](#) • [All experiments](#) • [All runs](#)
[show Abstract](#)

Sample: Chamaesiphon sp. PCC 6505
[SAMN01797692](#) • [SRS373642](#) • [All experiments](#) • [All runs](#)
Organism: [Octopus sp. HBH-6](#)

Library:
Name: HTHC
Instrument: Illumina Genome Analyzer II
Strategy: WGS
Source: GENOMIC
Selection: RANDOM
Layout: PAIRED

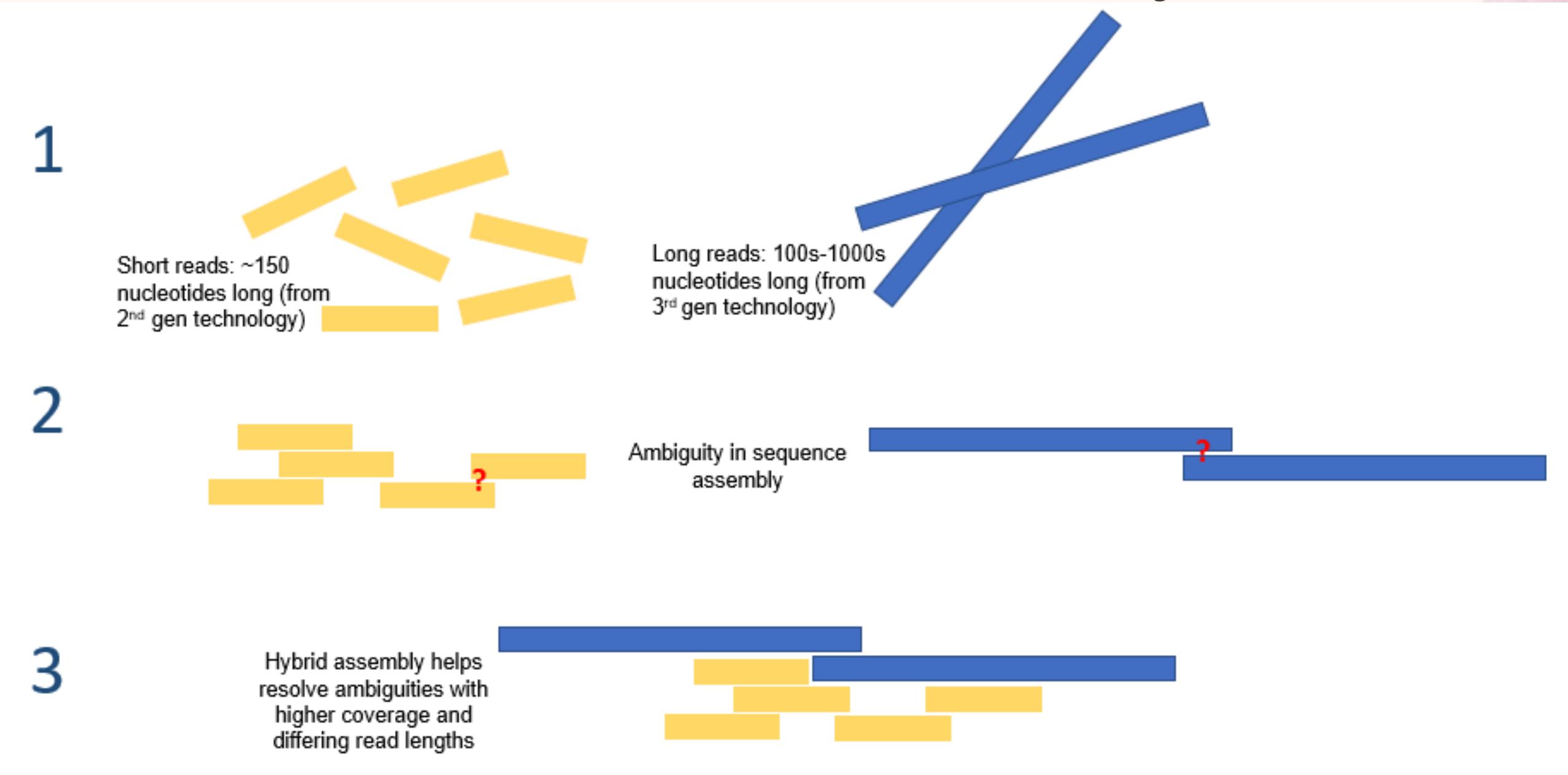
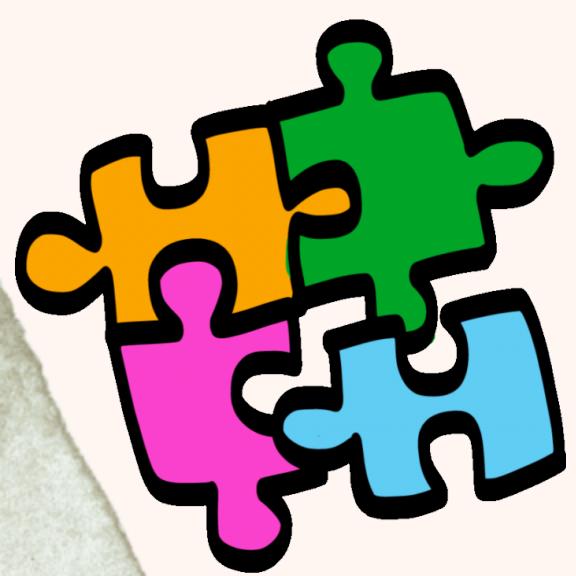
Spot descriptor:
1 forward → ← 77 reverse

Runs: 1 run, 30.2M spots, 4.6G bases, [2.6Gb](#)

Run	# of Spots	# of Bases	Size	Published
SRR610296	30,157,315	4.6G	2.6Gb	2015-07-22

ID: 269024

Genome Assembly



MEGAHIT

SPAdes

Assembly Quality Check



QUAST - MEGAHIT

QUAST - SPAdes

QUAST Output

MEGAHIT

SPAdes

Complete BUSCO (%)

89.86

89.86

Partial BUSCO (%)

2.70

2.70

N50

44433

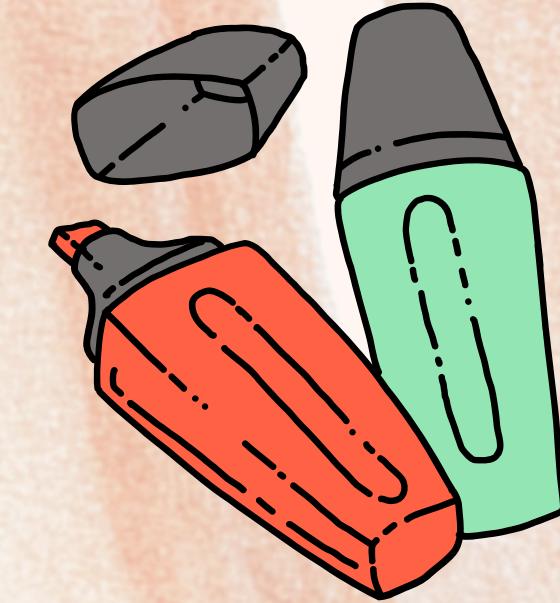
39388

Assembly Quality Check

QUAST - MEGAHIT	MEGAHIT	QUAST - SPAdes
QUAST Output		SPAdes
Complete BUSCO (%)	89.86	89.86
Partial BUSCO (%)	2.70	2.70
N50	44433	39388

Genome Annotation

Prokka



locus_tag	ftype	length_bp	gene	COG	product
BBPJPPNJ_02762	CDS	1080	recA		Protein RecA
BBPJPPNJ_05158	CDS	1538	merA_2	COG3668	Mercuric reductase
BBPJPPNJ_04017	CDS	267	yoeB	COG4115	Toxin YoeB

Genome Annotation

RAST

Gene	Description
RecA	Part of the DNA repair system.
merA	Chamaesiphon is resistant to Mercury, which is toxic to humans, and allows the element to participate in biogeochemical cycles. MerA is found with other metal transporters and reducers.
yoeB	YoeB/YefM: specific endoribonuclease, cleaves free mRNA. The YefM antitoxin of <i>E. coli</i> is unfolded in its native state; the YoeB toxin is folded and forms a complex with the unfolded YefM antitoxin, leading to inhibition of the RNase activity of YoeB.

IN CONCLUSION

My curiosity was satiated :)