**Abstract**

This is the critical initial contact with the reader. Distill the necessary parts of your proposal to one-half page or less, stating the problem and what you intend to do about it. Make it understandable to the intelligent, but inexpert, reader.

**Specific Aims**

List the major questions that will be answered in your research and the specific approaches that will be used to address those questions. This is typically done in an outline form of no more than one-half page. It should also provide the framework for the Experimental Design section below, so its organization is key to the entire proposal. Try to be realistic and propose an amount of work that you are likely to accomplish in the next 2-3 years; excessively optimistic proposals suggest a lack of critical thought.

It is often advisable to divide the following sections into subsections with titles to orient the reader.

1. Determine the evolutionary dynamics of sequence discrete populations using GFMs

Find pairs/sets of GFM’s that are closely related but different bins

What separates these? Ecotypes? Why are they maintained as separate populations?

Determine the evolutionary dynamics of populations using SAGs, looking at major freshwater lineages and closely related sets

2. Hypothesize ecological explanations using gene contexts for the patterns seen above

maybe propose experiments ?

Looking at the metabolic capabilities in the gene content

Investigate the role of specific phyla in the lakes

3. Oligotrophic vs. Copiotrophic comparison of found GFM’s and SAGs

IMPORTANT LIMITATION TO RECOGNIZE: Completion estimates? – Working with Ben

**Background and Significance**

This section should be several pages long and contain enough information to make the subsequent sections understandable to the reviewer. It should also give the reviewer an understanding of the state of the field before your participation. It should therefore cite any critical information that is either published, or known to you through personal communication. Your accomplishments will be described in the following sections, but it may be necessary to allude to some of your results in this section for clarity or argument. Results from your laboratory, but which you were not involved in, should be described in this section. This section should also serve to convince the reviewer that the general question chosen is an important one.

Driving advances in sequencing and stuff makes this

Sags Ramunas’ paper marine samples vs marine sags

Coverage binning, Banfield, Mads papers

Discovery based, but with new tools we can do more hypothesis driven science

Sequence Discrete populations

Ecotype model of bacterial speciation vs. other model

Sarahi’s paper characterizing L06 and metabolism

Trevor’s acI paper?

MDM paper

Livermore

Oligotrophs paper

Pnec paper

1. The uncultivated majority of microbes(see paper sent by Sarahi
   1. Looking at these microbes using
      1. 16S tags
         1. Newton Guide paper
         2. 16S to ANI – Kostas paper! Find these
      2. metagenomes – Can’t track back specific organisms well
         1. Many binning methods have been previously used – phylogenetic distribution, GC content, tetranucleotide frequency
         2. New binning method incudes using coverage differential
      3. SAGs – Random amplification misses chunks of DNA
         1. Mapping metas to SAGs – Sags Ramus’s paper marine samples vs marine sags
         2. MDM paper
2. Using Genomics to study ecology and evolution
   1. Evolution studies using isolates
      1. Lenski
      2. Whitaker and Vibrio
      3. Kettler Prochlorococcus paper
   2. Evolution studies using SAGs or metagenomes
      1. LD12 recombination paper
      2. Kashtan Prochlorococcus paper
      3. Banfield acid mine drainage – rates of evolution in nature
      4. Rex/Leo Paper
         1. Ecotype model
   3. Tracking populations in nature
      1. Mapping metas to SAGs – Sags Ramus’s paper marine samples vs marine sags
      2. Heinrich…Bertilsson Paper tracking LD12 in lake
      3. Martinez-Garcia..Stepanauskas
      4. Hahn Pnec Paper
      5. acI populations paper
   4. Metabolic potential inferred from genomes
      1. Newton Roseobacter paper?
      2. Sarahi’s L06 paper
      3. Trevor’s paper
      4. Salcher 2013 paper about substrate preferences (MAR-FISH)
      5. Verrucomicrobia paper
3. Genome streamlining and genomic features
   1. Lauro Paper
   2. Livermore Paper
   3. Rappe 2012 SAR11 streamlining
   4. Viklund 2012 SAR11 paper
   5. 2014 Luo- Moran Roseobacter paper
   6. Giovannoni SAR 11 streamlining paper!

**Preliminary Results**

Describe the progress you personally have made while in the lab. The goal of this section is to convince the reader that you have made some progress and/or that you have developed skills that will be necessary to complete the proposed work.

Work at JGI

Classifying set of GFMs from conserved genes

Written programs to help parse

Work on acI paper?

Work to construct core genes

Work to find closest relative

Population dynamics work

Written programs to help

Work on completion estimates??

The old method – should I test this with lower levels of classification?

Ben’s new method -- helping to construct methods paper

Going to write programs to help? Maybe should write synthetic SAG script help Ben with his work.

Learned some R?

Set up server?

Learned to work with IMG

Learned to work with JGI

Experimental Work – Do I need an experimental component needed? ( might come up, maybe talk with each member about this separately) - maybe with Raman FISH MAR

AIM1:

Binned metagenomes…sort of

Classified them using Marker genes with phylosift

Some examples of classification

How many where classified table

How many were classified in each phyla between each lake

Experience tracking populations with acI paper

Mapped all the metagenomes to the acI SAGs w/ BLAST

Parsed results for ANI

Constructed AAI program for comparing all the SAGs to one another

Completion estimates work

Script to test completion estimates by taking out various amounts of the genomes in blocks

**Experimental Plan**

Typically the sections in this part will follow in the order laid out in the Specific Aims. The goal here is to convince the reviewer that the approach you have chosen will yield interpretable results and that you really understand those approaches. If there are intermediate goals that are absolutely critical to the whole project, either defend why your single approach must work, or propose alternative "backup" approaches. Provide enough information to make it clear that you understand the technique; this does not mean an abundance of detail, but a terse description of potential problems and shortfalls in the experiment or its analysis. If there are obvious experiments that will not be done, tersely say why.

Throughout this section, make your priorities clear; not every experiment is as important as the next, and some approaches will be pursued only under certain circumstances. Continually orient the reader by explaining how each intermediate goal fits into the overall plan.

**Aim 1:**

Map meta reads to SAGs and GFMs

Look at abundance patterns

Look at SNP patterns

-gene vs. genome wide seeps

Look gene and loss patterns

What predicted functions genes that are gained or lost are

Recombination

Same method Siv’s group uses in the LD12 recombination paper? w/ metagenomes

**Aim 2:**

Use metabolic mapping programs to assess…

Carbon substrate usage/uptake

Nitrogen substrate usage/uptake

Phosphorous substrate usage/uptake

**Aim 3:**

Characterize

Predicted growth rate

Carbon substrate(done above)

Signal transduction genes

Suggest test other features….

**Timetable**

This short section should be a realistic estimate of when the critical intermediate goals in the proposal will be accomplished. It should also make clear when the primary approaches will be dropped and the alternatives adopted. You wish to convince the reviewer that, no matter what happens, you will return with a "story" suitable for a thesis in a reasonable time period.

**Literature Cited**

Using a standard format (authors' names and journal citation, including titles), list the references cited throughout the proposal. This should not only document your understanding of the state of current information, but also that you know the critical sources of information on the methods you have proposed to use.

**Overall Format**

The proposal should be limited to no more 18 double-spaced text pages including tables and figures. References are not included in the page limit.