

Training instructions for Sukhdeep

1. Install Ubuntu on machine
2. Follow setup instructions on Biowiki
3. Set up LDAP account
4. Set up Biowiki account
5. Set up account for emails
6. Get a basic idea of the project that you will be working on by consulting the links below. You do not have to understand all the details at the moment. For now it should be enough to just familiarize yourself with the project and the basic concepts.

<https://drive.google.com/folderview?id=0B4Vx2hKNcWEaXy1saGRzQnJ1UUU&usp=sharing>

<https://drive.google.com/folderview?id=0B4Vx2hKNcWEaNmp0NmNzeWcwaXM&usp=sharing>

Our clients for this project are Irene Martin (Head, Streptococcus and STI Unit, NML) and Walter Demczuk (Biologist, NML). They can be reached at the following email addresses:

irene.martin@phac-aspc.gc.ca
walter.demczuk@phac-aspc.gc.ca

We contact Irene and Walter when we have questions about what features to implement and how the application should behave (since they are the client/customer that we are developing the application for).

My contact information is also located here:

Irish Medina
Student, University of Manitoba, Department of Computer Science (expected graduation date: February 2015)
Email Address: irish.m.medina@gmail.com

Developed NG-STAR from January 2014 to August 2014.

I will also be developing NG-STAR from January 2015 to July/August 2015 and helping

you with training during this time period. I am also available to answer any of your questions. I will be attending graduate school out of town starting in September 2015. You will be taking over the project as the main developer after I am gone.

There are no other developers in this project, besides help from mentors such as Aaron.

To be able to effectively start work on the project you will have to do some initial training or review. Some of the concepts that you will have to learn or review are Linux, Perl, HTML, jQuery, git, GitLab, some biology concepts, some concepts in epidemiology, Catalyst MVC, Bootstrap, DBIx::Class (ORM), MySQL, MySQL Workbench and many others. There are many tools and topics to learn, so this process should take about 3 to 4 weeks (or more). We do not have any particular deadlines at the moment so there is no hurry.

7. Try out introduction to Linux tutorials for review

Introduction to Linux Part 1:

[http://wwwi.biowiki/index.php/File:Linux Workshop - 2014 - Day 1 %28morning%29 - Introduction To Linux.pptx](http://wwwi.biowiki/index.php/File:Linux_Workshop_-_2014_-_Day_1_%28morning%29_-_Introduction_To_Linux.pptx)

Introduction to Linux Part 2:

[http://wwwi.biowiki/index.php/File:Linux intro 2014.pptx.pdf](http://wwwi.biowiki/index.php/File:Linux_intro_2014.pptx.pdf)

Introduction to the Command Line:

[http://wwwi.biowiki/index.php/File:Linux Workshop - 2014 - Day 1 %28afternoon %29 - Introduction to the Command-line.pptx](http://wwwi.biowiki/index.php/File:Linux_Workshop_-_2014_-_Day_1_%28afternoon%29_-_Introduction_to_the_Command-line.pptx)

You need to know how to install software using apt-get and using APT (Advanced Packaging Tool) in general:

[http://wwwi.biowiki/index.php/File:Linux Workshop - 2014 - Day 1 %28afternoon %29 - Introduction to the Command-line.pptx](http://wwwi.biowiki/index.php/File:Linux_Workshop_-_2014_-_Day_1_%28afternoon%29_-_Introduction_to_the_Command-line.pptx)

The software that you will be developing uses the BLAST algorithm so it will be helpful to know how to use it. You should also try to get familiar with FASTA formatted files:

<http://wwwi.biowiki/index.php/File:UnixWorkshop2014-BLAST.pptx>

It might be helpful to learn how to use vim as a text editor:

<http://wwwi/biowiki/index.php/File:UnixWorkshop2014-Vim.pptx>

Review on SSH and soft links:

<http://wwwi/biowiki/index.php/File:UnixWorkshop2014-SSHLinking.pptx>

Advanced Linux Commands:

<http://wwwi/biowiki/images/3/3d/UnixWorkshop2014-AdvanceLinux.pptx>

8. Try out Perl tutorials for review

The software that you will be developing is developed using Perl programming language. A review on Perl is provided here:

<http://wwwi/biowiki/images/0/02/PerlNotes-PerlApr2012.pdf>

It will be helpful to be familiar with BioPerl and BLAST. You can skip the sections on ClustalW and Glimmer since you won't be using it. You should also make sure to cover CPAN near the end of the slides.

<http://wwwi/biowiki/images/0/02/AdvPerlNotes-Apr2012.pdf>

9. Learn biology basics

If you do not have a biology background, it is strongly recommended that you review the concept behind DNA transcription and translation. You should know what genes are and what a genome is. It may be helpful to know what the Human Genome Project is.

https://www.youtube.com/watch?v=-vZ_g7K6P0 (DNA)

<https://www.youtube.com/watch?v=Kzgnl5-8Wak> (DNA Transcription, Translation, Replication)

https://www.youtube.com/watch?v=41_Ne5mS2ls (DNA Transcription and Translation)

<https://www.youtube.com/watch?v=MvuYATh7Y74> (Human Genome Project)

10. Learn biology in more depth

The major topics surrounding the background on NG-STAR that are important to understand are antimicrobial resistance (what it is, why does it happen, why is it such a problem), how *N. Gonorrhoeae* relates to antimicrobial resistance, what MLST (Multi-locus sequence typing) is, why we use MLST, what alleles, sequence types and loci are in MLST. You may want to research these on your own to get a better understanding.

Here are some additional links on background information that may be helpful:

http://en.wikipedia.org/wiki/Multilocus_sequence_typing

[In AMR (antimicrobial resistance) we don't refer to alleles as “housekeeping genes” since these genes are highly recombinant for bacteria such as *N. Gonorrhoeae*. Instead, we refer to these genes as “antimicrobial resistant genes”]

There are a few existing MLST web applications, but these are not concerned with antimicrobial resistance. It may be helpful to try out these exiting web applications to see how MLST works.

<http://pubmlst.org/software/database/bigsdb/>

<http://www.mlst.net/>

<http://www.ng-mast.net/> (this is an MLST scheme for *N. Gonorrhoeae*, however, it only includes 2 genes and these genes are not associated with antimicrobial resistance)

To obtain an even better understanding of MLST, you can try out the tutorials that Aaron prepared on Phyloviz:

<http://wwwi.biowiki/images/1/1e/Day6LabPetkauPhylovizgeoBURSTMay2014.pdf>

<http://wwwi.biowiki/images/2/2b/Day6PetkauMSTLab2014.pdf>

<http://wwwi.biowiki/images/9/9c/Day6PetkauMSTLabAnswers2014.pdf>

11. Learn how to use git and GitLab

We use git for our version control software and GitLab which is a web interface for managing projects using git. You need to be familiar with both git and GitLab. Make sure you work through the examples and be able to perform basic commands like clone, add, commit, pull, and push. You should try creating a simple Hello World project on GitLab, generate SSH public and private keys and try out the Issues functionality. You should try cloning the project that you created, make a change to a file present in the project and be able to add, commit and push changes to this project.

You can read additional tutorials here:

<http://wwwi.biowiki/images/c/c0/PetkauGitJan2013.pdf>

<https://try.github.io/levels/1/challenges/1>

<http://rogerdudler.github.io/git-guide/>

<https://www.atlassian.com/git/tutorials/setting-up-a-repository>

12. Work through the initial project requirements

http://gitlab.corefacility.ca/irish.medina/ng-star/blob/development/Documentation/1_Requirements%20%28Read%20this%20first%29.pdf

You should try to work through this guide thoroughly. Make sure that you try the tutorials on Catalyst MVC, DBIx::Class, MySQL and MySQL Workbench as specified in the guide. This is also a great opportunity to review on SQL, HTML and jQuery. This guide should take a significant amount of time to complete.

13. Try out the Allele Database project

After you have completed the initial project requirements guide in detail, you can proceed with trying out the Allele Database project. It will be helpful to type out the code yourself as opposed to copying and pasting the code.

http://gitlab.corefacility.ca/irish.medina/ng-star/blob/development/Documentation/2_Tutorial.pdf

The architecture of this project and that of NG-STAR differs from a general Catalyst project significantly. The Allele Database project and NG-STAR have a separate BusinessLogic and Data Access Layer (DAL) so that the application has greater separation of concerns. It is very important to be able to make this distinction.

14. Try creating your own mini project that does something interesting. Make sure that your project has a separate BusinessLogic and DAL like NG-STAR. Use Bootstrap, FormHandler and DBIx::Class. Create and modify your database schema using MySQL Workbench.

15. Install NG-STAR on your machine

To install NG-STAR on your machine, just follow this tutorial:

http://gitlab.corefacility.ca/irish.medina/ng-star/blob/development/Documentation/3_Installation%20Guide.pdf

16. Read additional design documentation for NG-STAR

NG-STAR Database Schema:

http://gitlab.corefacility.ca/irish.medina/ng-star/blob/development/Documentation/4_NGSTAR%20Database%20Schema.pdf

Update Schema Guide:

http://gitlab.corefacility.ca/irish.medina/ng-star/blob/development/Documentation/5_Update%20Schema%20Guide.pdf

Integration Testing:

http://gitlab.corefacility.ca/irish.medina/ng-star/blob/development/Documentation/7_Integration%20Testing.pdf

Update NG-STAR Auth Schema:

http://gitlab.corefacility.ca/irish.medina/ng-star/blob/development/Documentation/8_Update%20NGSTAR_Auth%20Schema.pdf

Alleles for Testing:

http://gitlab.corefacility.ca/irish.medina/ng-star/blob/development/Documentation/9_Test%20Alleles.pdf

17. Review the code and examine how everything is structured

Reviewing the code and determining how everything is structured will help you when you are implementing your first feature later on.

<http://gitlab.corefacility.ca/irish.medina/ng-star/tree/development>

18. Start implementing your first feature!