Hokkaido University
Program for Leading Graduate Schools
Fostering Global Leader in Veterinary Science for contributing to One Health

([✓ Abroad] Domestic) Internship report form (Student) 2016/08/09

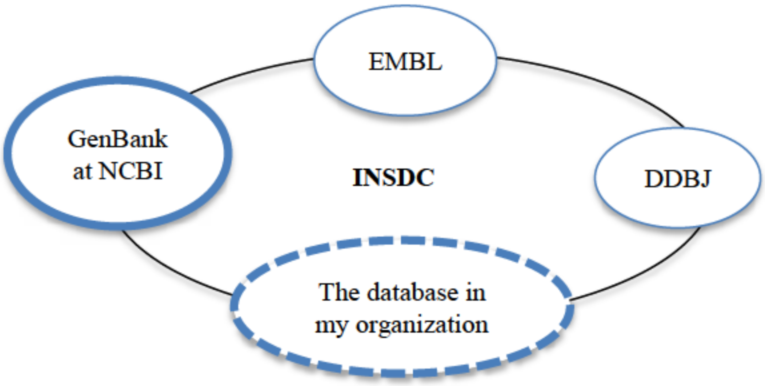
(Year/Month/Day)

Name	Nipawit Karnbunchob
Laboratory	Bioinformatics
Year (Grade)	4
Internship institution	National Center for Biology Information (NCBI), a division of the National Library of Medicine (NLM) at the U.S. National Institute of Health (NIH), USA
Internship period	Internship period: 06/27/2016 - 07/24/2016 (Departure Date from Sapporo: 06/26/2016, Arrival Date in Sapporo: 07/26/2016)
Purpose	1. To obtain knowledge and new experiences on the management of NCBI database, which contribute me to create the genetic sequence database for public service 2. To get experience of job activities at NCBI 3. To increase the opportunities to get a job at university and international organization that I am interested in 4. Networking and to meet new people in my field

This report should be submitted within 2 weeks after you return to Japan.

- The reason why you chose this institute

I am planning to create the genetic sequence database in my country (Thailand) after I graduated. This database will be a part of the International Nucleotide Sequence Database Collaboration (INSDC) and the first genetic sequence database of South-East Asia. INSDC are comprised from the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI, these three organizations exchange data on a daily basis. For NCBI, they retrieve the data from Genbank and provide the public access to biomedical and genomic information via databases of NCBI. Additionally, NCBI also provide the sequence analytic tools such as Basic Local Alignment Search Tool (BLAST). Therefore the internship at NCBI gave me the necessary information to create the genetic sequence database in my organization, which will be one more organization of INSDC.



- Result of the activity（about 800 words, provide photos, tables and figures that clearly show the activities during the period）

For Internship at NCBI, which is an institute of NIH, I prepared the certificates of online training modules, B-1 visa, and the Arrival-Departure Record (Form I-94) of particular aliens. The three specific requirements were the important documents to use to entry into NIH and get my ID badge. Before departure, I must completed the following five online training modules:

1. Information Security Awareness
2. Privacy Training
3. Information Security and Privacy Awareness Refresher
4. Securing Remote Computers
5. Remote Access User Certification Agreement

I also obtained a B-1 visa stamp from a U.S. consulate or embassy to entry into the U.S. After arrival, I have to print out my Form I-94 as soon as I have computer access upon entering the U.S. My arrival to the U.S. will be electronically captured by generation of a form Arrival/Departure record.

Note: B-1 status is for a visitor who plans to engage in business consultations, independent research or other professional activities.

The details of my daily activities, detail for learning about resource of viral genomes in the viral genome webpage of NCBI, and meeting with the virus group. My host supervisor at NCBI is Dr. Yiming Bao. He is a staff of virus group of NCBI. I learnt many things about the resource of viral genomes during my internship at NCBI. I focused on the virus part and use of tools. NCBI provides viral and viroid genome sequence data, related information and tools. I also attended the meeting of virus group of NCBI. Internship at NCBI gave me the best chance to gain understanding of bioinformatics related jobs, knowledge, and networking.

Schedule of activities

Periods	Activities
June 27 - 29	- I did the paper work to get ID badge. - I started browsing general information from the viral genome webpage of NCBI and got some idea about what NCBI staff scientists do. Link of webpage: http://www ncbi.nlm.nih.gov/genome/viruses/
June 30 - July 1	- Meeting with Dr. Yiming Bao every morning - Learning about “Explore Viral Genome Sequences” section and “Download Viral Genome Data” section
July 5 - 8	- Meeting with Dr. Yiming Bao every morning

	- Learning about "Explore Viral Genome Sequences" section (continued) - Learning about " Resource Tools" section and practice about use of tools in this section - Meeting with Virus group of NCBI on Friday 8.
July 11 -15	- Meeting with Dr. Yiming Bao every morning - Learning about "Virus Variation Resource" section - Meeting with Virus group of NCBI on Friday 15.
July 18 -22	- Meeting with Dr. Yiming Bao every morning - Learning about "Virus Variation Resource" section (continued) and practice about use of tools in this section - I concluded the details about my duty and report to Dr. Yiming Bao. - I did the paper work to return ID badge.

Resource of viral genomes in NCBI

Explore Viral Genome Sequences

I learnt about Reference Sequences (RefSeq) project. RefSeq is the most complete genome for each virus. It is usually derived from the first sequences of new virus submitted in Genbank. The RefSeq of new virus can be replaced with another sequence of new virus to be new RefSeq of the new virus if that sequence is more complete than the previous RefSeq. I also learnt how NCBI staffs select the sequence to determine the RefSeq. The complete or nearly complete sequences submitted later and show high similarity with RefSeq will be determined as neighbors.

Resource Tools

This section provides information about retroviruses, a link of human interaction database from other sources. This resource also contains Pairwise Sequence Comparison Tool (PASC) and genotyping tool. For PASC, this tool can analyze the pairs of sequences and classify the viruses to Family, Genus, and Species level of taxonomy. For example, we can merge or separate the species of virus with percent identity in our criteria. For genotyping tool, we can use this tool to do genotyping between other viruses and their reference sequences. We can try to test genotyping with example of viral sequence (as sequence model for typing) and many options to study the use of genotyping tool.

Virus Variation Resource

This section is the section of databases of NCBI. There are a lot of virus databases following Influenza Dengue, West Nile, MERS, Ebola, Rota, and Zika virus databases. NCBI staffs constructed the databases to contribute the user access to target data by using keywords for searching such as name of countries, years, and hosts. They retrieved the data from GenBank then classified the data to easy to access. We can access and download a lot of data via viral genome webpage of NCBI following this link of this webpage.

In the case of influenza virus database that I downloaded the sequence data to do my research, NCBI retrieve protein and nucleotide sequences from GenBank to use to construct flu database. We can download and select the sequences by options of search set and additional filters, which are very useful when we want to download the sequences of different segments from a same isolate. This database also provides many tools that are interesting for research such as alignment and phylogenetic tree building tool, and BLAST.

Meeting with the members of virus group of NCBI

I attended the weekly meeting of the virus group. Main detail of meeting is for development of the databases and webpage that including report some problem they found. The members of the virus group are composed of biologist, biochemist, and computer scientist. They work together. There are a lot of duties they have to do. Conclusions of their duties usually are:

- Develop, maintain, and create the tools such as Pairwise Sequence Comparison Tool (PASC), genotyping, alignment, and tree construction tools.
- Develop, maintain, and construct the virus databases
- Curation the sequences for RefSeq project

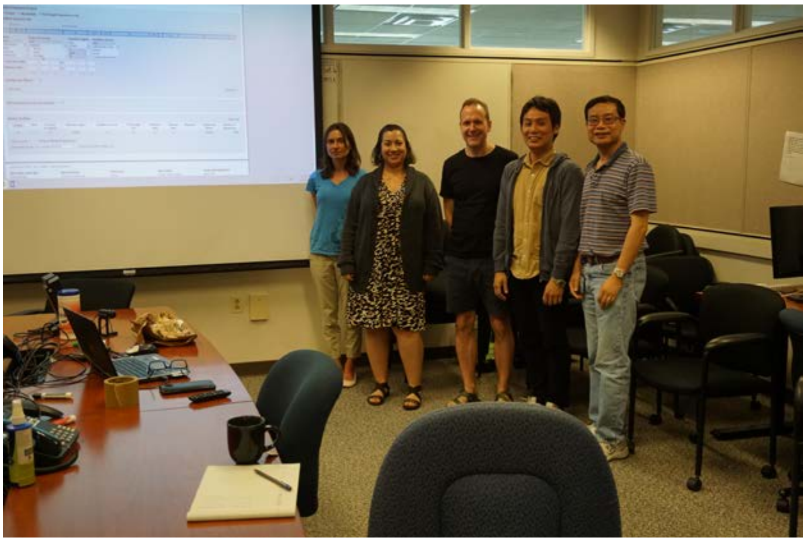
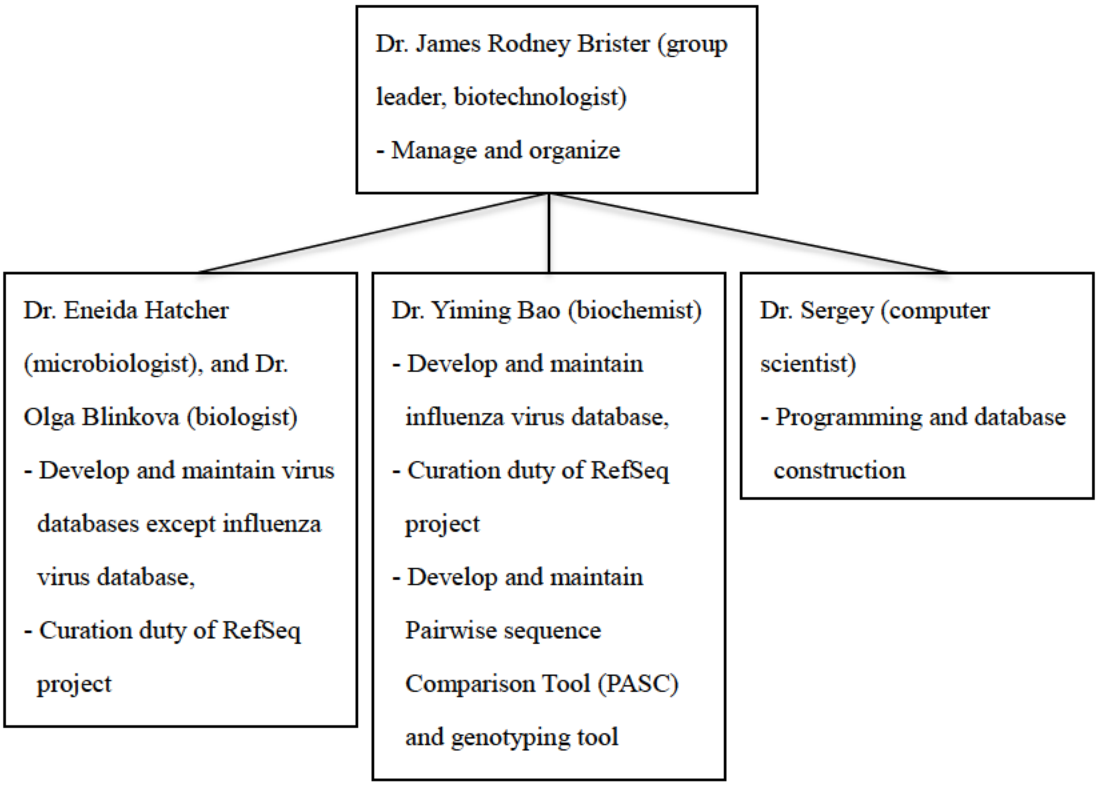


Figure 1. My photograph with the members of virus group, Dr. Olga and Dr. Eneida (from left side) are the staff scientistes, Dr. Rodney (middle) is the group leader of virus genome group. Dr. Bao is a staff scientists and my host supervisor or host (right)



Dr. James Rodney Brister (group leader, biotechnologist)
- Manage and organize

Dr. Eneida Hatcher (microbiologist), and Dr. Olga Blinkova (biologist)
- Develop and maintain virus databases except influenza virus database,
- Curation duty of RefSeq project

Dr. Yiming Bao (biochemist)
- Develop and maintain influenza virus database,
- Curation duty of RefSeq project
- Develop and maintain Pairwise sequence Comparison Tool (PASC) and genotyping tool

Dr. Sergey (computer scientist)
- Programming and database construction

Figure 2. Organization and their duties for virus genome group in NCBI

Figure 3. My desk at NCBI

Figure 4. NCBI building

- What do you think the positive impact of the activity will have on your further career path?

I am looking for a job position, at university in my country (Thailand) and international organization. An internship at NCBI gave me valuable experience and insight knowledge of the careers as a bioinformatician, allowing me to make more informed choices when looking for employment opportunities. I can bring the knowledge and experience I gained in this internship back with me not only to CZC, but also to Japan and Thailand. By the experience of NCBI internship, I believe I can improve not only bioinformatics research project but also system of management and administration at the university or organization that I am interested in. The internship experience from NCBI that is a world-class organization might be one of important qualification helps me to pass the good job selection among high competition. Moreover, this internship also let I known what kind of the necessary knowledge and proficiency I should have or improve when I work in bioinformatics field.

- Advice for your junior fellows

Review about principle of bioinformatics, molecular biology, and virology will be useful for the persons who are interested in the internship at virus group of NCBI.

Approval of supervisor	Institution・Official title・Name Division of Bioinformatics, Research Center for Zoonosis Control Prof. Kimihito Ito 印

※1 Send the electronic file to the Leading School section, International Affairs Office, also submit the original print out with seal of supervisor to the Leading School section, International Affairs Office.
※2 Attach a copy certificate of the content of internship activity that is prepared by the counterpart at the internship institution (any form with a signature of the counterpart).
※3 The Steering Committee of the Leading Program will first confirm the content of this report and report will be forwarded to the Educational Affairs Committee for credits evaluation.

Submit to：Leading School section, International Affairs Office

Ext: 9545　e-mail: leading@vetmed.hokudai.ac.jp