
Plan Overview

A Data Management Plan created using DMPTool

Title: Mosquito evolution in response to a dengue envelope protein-guided insecticidal peptide

Creator: Christopher Kearney

Affiliation: Baylor University (baylor.edu)

Funder: National Institutes of Health (nih.gov)

Funding opportunity number: PAR-21-155

Grant: <https://grants.nih.gov/grants/guide/pa-files/PAR-21-155.html>

Template: NIH-GDS: Genomic Data Sharing

Project abstract:

The long-term objective of this project is to provide a solution to the alarming increase in insecticide resistance in mosquito vectors of human pathogens. Insecticides target specific proteins in the mosquito and resistance mutations in these proteins prevent the binding and toxicity of the insecticide. In the proposed research, an orally toxic peptide insecticide which specifically targets the mosquito gut receptor of dengue virus-2 will be fed to *Aedes aegypti*, the mosquito vector of dengue virus, to force the evolution of resistance in the mosquito. It is hypothesized that mosquitoes that have lost the ability to bind the insecticide will have also lost the ability to bind and vector dengue virus itself. Thus, mosquito populations would evolve away from vectoring dengue if resistance were to develop. This peptide insecticide is modular, fusing the commercially successful insecticidal peptide, Hv1a, to a guide peptide comprising 15 amino acids from the envelope protein of dengue virus-2 which specifically bind mosquito cells but not human cells. Guide peptides for future mosquitocidal peptides could be sourced from other vector-borne pathogens to control other mosquito genera or other vectors attracted to sugar baits. Specific Aim 1 is to evolve populations of mosquitoes which are resistant to the guided Hv1a peptide insecticide. Three populations will be passaged, being fed the dose in 10% sucrose which kills 80% of the population (LD80). Survivors will be mated and the resulting eggs provide the next passage, with dosage increased with increased survivorship, until the entire population can survive the original LD100. Control populations fed unguided Hv1a and plain 10% sucrose will be in parallel passages. Another set of populations will be fed deltamethrin insecticide at the LD80 to evolve resistance, providing an in-house comparison to standard insecticides. In Specific Aim 2, eggs from guided Hv1a-evolved and control populations will be sent to our collaborator, Berlin Londono of Kansas State University, to assay for the ability of dengue virus-2, provided in an oral blood meal, to infect the adult mosquitoes. RNA-Seq and specific sequencing of the dengue virus receptor protein gene (prohibitin) will track the evolution of gene expression and prohibitin through the passaging. If vectorial capacity is lost, the project hypothesis is proven and an inexpensive means of modulating vectorial capacity will have been discovered. If vectorial capacity is maintained over the entire 3-year term of the passaging, then an insecticide uniquely impervious to resistance will have been discovered instead.

Start date: 04-01-2022

End date: 03-31-2025

Last modified: 06-18-2021

Grant number / URL: <https://grants.nih.gov/grants/guide/pa-files/PAR-21-155.html>

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Mosquito evolution in response to a dengue envelope protein-guided insecticidal peptide

Data type: human genomic data and non-human genomic data

The GDS Policy applies to all NIH-funded research that generates large-scale human or non-human genomic data as well as the use of these data for subsequent research. Large-scale data include genome-wide association studies (GWAS), single nucleotide polymorphisms (SNP) arrays, and genome sequence, transcriptomic, metagenomic, epigenomic, and gene expression data, irrespective of funding level and funding mechanisms (e.g. grant, contract, cooperative agreement, or intramural support). NIH Institute or Centers (IC) may expect submission of data from smaller scale research projects based on the state of the science, the programmatic priorities of the IC funding the research, and the utility of the data for the research community.

Gene expression data from RNA-Seq analysis will be collected in this project from mosquito populations at various time points throughout the passaging to track changes in gene expression as the mosquito populations evolve in response to increasing dosages of guided Hv1a peptide insecticide offered in 10% sucrose. Other populations monitored in similar fashion will be populations evolving in response to increasing dosages of deltamethrin insecticide and also two types of control populations, namely populations given the nontoxic Hv1a peptide and populations supplied only with plain 10% sucrose.

All of this RNA-Seq data will be posted on Gene Expression Omnibus at <http://www.ncbi.nlm.nih.gov/geo/>

PacBio sequencing will be performed, at various time points in the mosquito passaging, on the prohibitin receptor in the mosquito gut, which naturally binds dengue virus but is also targeted by the guided Hv1a protein insecticide. These sequences will trace the evolution of the prohibitin gene in response to guided Hv1a during passaging. Sequencing will also be done on the two control populations offered Hv1a or plain 10% sucrose.

All of these full-gene-length reads of prohibitin gene evolution will be posted on the Sequence Read Archive at <http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi>

Genomic research advances our understanding of factors that influence health and disease, and sharing genomic data provides opportunities to accelerate that research through the power of combining large and information-rich datasets. <https://osp.od.nih.gov/scientific-sharing/policies/>

The RNA-Seq and prohibitin gene sequencing data mentioned above will be loaded onto these public servers within a few months of analyzing the data and all data will be completely loaded before publication of the research in a journal.

Data repositories

Identify the data repositories to which the data will be submitted, and for human data, whether the data will be available through unrestricted or controlled-access.

Investigators should register all studies with human genomic data that fall within the scope of the GDS Policy in dbGaP by the time that data cleaning and quality control measures begin. After registration in dbGaP, investigators should submit the data to the relevant NIH-designated data repository (e.g., dbGaP, GEO, SRA, the Cancer Genomics Hub). NIH-designated data repositories need not be the exclusive source for facilitating the sharing of genomic data, that is, investigators may also elect to submit data to a non-NIH-designated data repository in addition to an NIH-designated data repository. However, investigators should ensure that appropriate data security measures are in place, and that confidentiality, privacy, and data use measures are consistent with the GDS Policy.

Non-human data may be made available through any widely used data repository, whether NIH-funded or not, such as GEO, SRA, Trace Archive, Array Express, Mouse Genome Informatics, WormBase, the Zebrafish Model Organism Database, GenBank, European Nucleotide Archive, or DNA Data Bank of Japan.

All of this RNA-Seq data will be posted on Gene Expression Omnibus at <http://www.ncbi.nlm.nih.gov/geo/>

All of these full-gene-length reads of prohibitin gene evolution will be posted on the Sequence Read Archive at <http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi>

All data submitted will be available with unrestricted access to the public.

Data submission expectations and timeline

Investigators should submit large-scale genomic data as well as relevant associated data (e.g. phenotype and exposure data) to an NIH-designated data repository in a timely manner. Investigators should also submit any information necessary to interpret the submitted genomic data, such as study protocols, data instruments and survey tools. Genomic data undergo different levels of data processing, which provides the basis for NIH's expectations for data submission and timelines for the release of the data for access by investigators. These expectations and timelines are provided in the Supplemental Information. In general, NIH will release data submitted to NIH-designated data repositories no later than six months after the initial data submissions begins, or at the time of acceptance of the first publication, whichever occurs first, without restrictions on publication or

other dissemination.

In addition to the sequencing data, submissions to these servers will include metadata on the sequencing devices used and the sequencing conditions, the extraction of the RNA and DNA, library preparation, and details on the passaging experiments (including explanations of treatment labels) to allow for full interpretation of the sequencing data by those that download the data.

The RNA-Seq and prohibitin gene sequencing data mentioned above will be loaded onto these public servers within a few months of analyzing the data and all data will be completely loaded before publication of the research in a journal.

Informed consent and institutional certification

Respect for, and protection of the interests of, research participants are fundamental to NIH's stewardship of human genomic data. The informed consent under which the data or samples were collected is the basis for the submitting institution to determine the appropriateness of data submission to NIH-designated data repositories, and whether the data should be available through unrestricted or controlled access.

For research that falls within the scope of the GDS Policy, submitting institutions, through their Institutional Review Boards (IRB's), privacy boards, or equivalent bodies, are to review the informed consent materials to determine whether it is appropriate for data to be shared for secondary research use. Specific considerations may vary with the type of study and whether the data are obtained through prospective or retrospective data collections. NIH provides additional information on issues related to the respect for research participant interests its "*Points to Consider for IRB's and Institutions in their Review of Data Submission Plans for Institutional Certifications*" (updated in 2016 to "*Points to Consider for Institutions and Institutional Review Boards in Submission and Secondary Use of Human Genomic Data under the National Institutes of Health Genomic Data Sharing Policy*").

Human subject data will not be collected in these experiments.

Exceptions to data submission expectations

In cases where data submission to an NIH-designated data repository is not appropriate, that is, the Institutional Certification criteria cannot be met, investigators should provide a justification for any data submission exceptions requested in the funding application or proposal. The funding IC may grant an exception to submitting relevant data to NIH, and the investigator would be expected to develop an alternate plan to share data through other mechanisms.

All data will be made publicly available with no exceptions.

Intellectual Property

NIH encourages patenting of technology suitable for subsequent private investment that may lead to the development of products that address public needs without impeding research. However, it is important to note that naturally occurring DNA sequences are not patentable in the U.S. Therefore, basic sequence data and certain related information (e.g. genotypes, haplotypes, *p*-values, allele frequencies) are pre-competitive. Such data made available through NIH-designated data repositories, and all conclusions derived directly from them, should remain freely available, without any licensing requirements.

NIH encourages broad use of NIH-funded genomic data that is consistent with a responsible approach to management of intellectual property derived from downstream discoveries, as outlined in the NIH *Best Practices for the Licensing of Genomic Inventions* and Section 8.2.3. Sharing Research Resources, of the NIH Grants Policy Statement. NIH discourages the use of patents to prevent the use of or to block access to genomic or genotype-phenotype data developed with NIH support.

No intellectual property will be pursued with this data. This data will remain freely available, without any licensing requirements.

Planned Research Outputs

Dataset - "RNA-Seq data for gene expression changes in *Aedes aegypti* in response to treatment with guided Hv1a peptide insecticide "

Intended Repository: Gene Expression Omnibus at <http://www.ncbi.nlm.nih.gov/geo/>

This dataset is RNA-Seq data from *Aedes aegypti* mosquito populations in serial passaging with increasing dosages of guided Hv1a, which is the peptide insecticide omega-hexatoxin Hv1a from the spider *Hadronyche versuta*, fused to a guide peptide derived from the binding site in the envelope protein of dengue virus-2 for the mosquito gut receptor prohibitin.

Dataset - "Prohibitin full-gene-length sequences for evolution in *Aedes aegypti* in response to guided Hv1a treatment"

Intended Repository: Sequence Read Archive at <http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi>

Prohibitin full-gene-length PacBio sequences from *Aedes aegypti* mosquito populations in serial passaging with increasing dosages of guided Hv1a, which is the peptide insecticide omega-hexatoxin Hv1a from the spider *Hadronyche versuta*, fused to a guide peptide derived from the binding site in the envelope protein of dengue virus-2 for the mosquito gut receptor prohibitin.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
RNA-Seq data for gene expression changes in <i>Aedes</i> ...	Dataset	2025-04-18	Open	None specified	1 GB	None specified	None specified	No	No
Prohibitin full-gene-length sequences for evolutio ...	Dataset	2025-04-18	Open	None specified	1 GB	None specified	None specified	No	No