

Mosquitoes are the deadliest animals on the planet, due to the many diseases they transmit. In the US, the most common mosquito-borne disease is West Nile virus (WNV), transmitted by Culex mosquitoes. In urban environments, improperly draining storm-water infrastructure, such as catch basins and culverts, which contain polluted stagnant water, provide ideal habitat for these mosquitoes, creating localized foci of WNV transmission risk. Our project focuses on surveying and identifying mosquito habitats in the City of Spokane and the City of Cheney, and understanding the relationship between environmental conditions, mosquito productivity and WNV transmission risk. Faculty from the Department of Biology at Eastern Washington University (Krisztian Magori, Camille McNeely and Jenifer Walke) and from Gonzaga University (Christy Andrade) will collaborate with the Spokane Regional Health District (SRDH), the Washington Department of Health (WADOH) Zoonotic Disease program, and other stakeholders in this project. During the summer, a graduate student from EWU will lead undergraduate students to check catch basins and culverts for standing water, take measurements and water samples, and monitor for mosquito larvae. If mosquito breeding is detected, they will set traps and collect mosquitoes, and send them to WADOH for WNV testing, communicating the risk to SRDH. Water samples and mosquito larvae will be tested for both nutrient and microbial content to identify the combination of nutrients and microbes that support mosquito productivity. The project will also identify specific locations that contribute to significant mosquito production and WNV transmission, allowing SRDH and other stakeholders to monitor and mitigate them.

1. Plan and Acquire

There will be several datasets collected. The unit of collection will be the storm water catch basins and culverts. Researchers will note other potential mosquito breeding sites in the vicinity, but not investigate, especially on private property. Data collected will be the presence/absence of standing water in each storm-water catch basin on the day of survey. If standing water is found, researchers will return once a week to the location and collect the following data each time:

- water level height from the bottom
- water level height from the top
- water temperature (using iButton every 10 seconds)
- the average number of mosquito larvae collected per dip in 10 dippings
- ID number of the sample vial for nutrient analysis
- ID number of the sample vial for microbiome analysis

If any mosquito larvae are present, the following additional data will be collected:

- ID number for the vial with mosquito larvae from this location

In addition, a gravid trap will be placed within 10 m of the catch basin or culvert, and operated once a week, collecting the following additional data:

- ID number of the catch bag containing mosquitoes collected at the site on the
- specific collection date
- the number of mosquitoes of different species (data from WADOH)
- the WNV status of vector mosquitoes (data from WADOH)
- Minimum Infection Rate (calculated from the above 2 information)
- Vector Index (calculated from the above 3 pieces of information)

Once water samples are processed for nutrient analysis, the following data will be obtained for each water sample:

- nitrate concentration
- phosphate concentration

Once water samples are processed for microbial analysis, the following data will be obtained for each water sample:

- overall bacterial load
- species of culturable bacteria
- microbiome species diversity
microbiome composition

Once mosquito larvae are processed, the following data will be obtained for each mosquito larva sample:

- species identity of mosquito larvae collected
- species diversity of mosquito larvae
- species composition of mosquito larvae
- microbial community of mosquito larvae
- species and composition of other macroscopic organisms

We will use existing data on the location of storm water inlets (storm drains) and storm water gravity mains. These are stored as GIS shapefiles at:

<http://data-spokane.opendata.arcgis.com/datasets?t=public%20works>

We will work with the City of Spokane Public Works and Utilities to identify the best dataset to use to identify catch basins and culverts throughout the City. These shapefiles seem to be open to the public.

We will also work with the City of Cheney and potentially Spokane County to find and acquire similar dataset for Cheney.

We will work with the Spokane Regional Health District, the City of Spokane and the City of Cheney, and Spokane County to obtain any necessary permits. We will also apply for a scientific collection permit from the Washington Department of Fish and Wildlife to collect mosquito larvae and adults.

The undergraduate and the graduate student will collect the information initially on paper forms on the presence of standing water and other potential breeding sites. For catch basins containing standing water, a separate form will be completed with information on the water height from the bottom and from the top, the ID of the water temperature logger, and the ID of the water samples for nutrient and microbiome analysis, as well as the presence of mosquito larvae, and their density per dip if present, as well as the ID of the vial containing mosquito larvae. These paper forms will be transferred into Microsoft Excel at the end of the day, both by the undergraduate and the graduate student, duplicated for accuracy. The PI will check at least once a week and consult with the students to resolve discrepancies.

Separate form will be completed for each adult mosquito trapping event, including the ID number of the collection bag, the location and the date of trapping. Upon receipt of information from WADOH, this will be expanded to include the number of adult mosquitoes of each species, the species diversity, and the presence or absence of WNV in the mosquitoes, as well as the MIR at the location and the Vector Index. These forms will be completed in Microsoft Excel.

Another form will contain information on the nutrient levels measured in the water sample. This form will contain the ID number of the vial tested, the location and date of

collection, and the amount of nitrate and phosphate in the water sample. This form will be in Microsoft Excel.

Another form will contain information on the microbiome of the water sample. This form will include the ID number of the vial tested, the location and date of collection, and the bacterial load, culturable microbe, and the microbiome diversity and composition in the water sample. This form will be in Microsoft Excel.

Finally, another form will contain information on the mosquitoes in the collection of mosquito larvae. This form will include the ID number of the vial of mosquito larvae, the location and date of collection, the number of mosquito larvae of different species, the microbial community of mosquito larvae, as well as a list of other macroscopic organisms found in the water sample. This will also be in Microsoft Excel.

I believe the amount of data generated will not be more than 1 GB, and it shouldn't pose a challenge in terms of storing or sharing. The only exception might be raw sequence data from the microbiome portion of the project, but that will just be used for generating the microbiome composition.

Data will be updated as collected, but static afterwards.

The PI and the graduate student will be responsible for data management as part of their regular duties.

2. Describe/Metadata and Manage Quality

Catch basin/culvert dataset

Dataset on the presence/absence of standing water at each catch basin and culvert in the City of Spokane and the City of Cheney.

Breeding site dataset

Dataset on the characteristics of each catch basin and culvert in the City of Spokane and the City of Cheney that contained standing water, with the amount and characteristics of standing water (volume, residence time, temperature, nutrient levels, microbial load, microbiome composition, presence/absence of mosquito larvae and other macroscopic organisms).

Mosquito larvae dataset

Dataset on the characteristics of each catch basin and culvert in the City of Spokane and the City of Cheney that were inhabited by mosquito larvae at any point in the season, including mosquito larval density, species of mosquitoes, length of the season when containing mosquito larvae, microbial characteristics in mosquitoes, in addition to the characteristics above.

Adult mosquito dataset

Dataset on adult mosquitoes collected, associated with the catch basins and culverts in the City of Spokane and the City of Cheney inhabited by mosquito larvae at any point in the season, including the number of adult mosquitoes of various species collected, the West Nile virus status of the vector mosquitoes, and the Minimum Infection Rate and Vector Index throughout the season.

All data will be collected as CSV (comma-separated value) files, in order to easily read in R. These will be then merged with shapefiles of the catch basins/culverts to

create maps. Potentially MS Access will be used to cross-reference datasets.

R will be used for data analysis. Quality assurance will be conducted by cross-referencing ID numbers between different datasets using R, potentially MS Access. R will be used to test hypotheses for the presence/absence of mosquito larvae, adult mosquitoes, and West Nile virus infection, in relation to nutrient levels and microbiome characteristics.

Metadata will be created and updated by the graduate student and the PI, in consultation with the co-PIs and other collaborators. Metadata will be created at the time of data collection. As the CSV files are created, each column will be identified and described in the metadata, following metadata standards.

FGDC-CSDGM will be used.

QA/QC will be provided by duplicate recording from the undergraduate and the graduate student, and checking for discrepancies by the PI, and resolving discrepancies. QA/QC will also be conducted using R to check linkages between datasets by cross-referencing using ID numbers.

3. Backup/Secure and Preserve

Data will be stored on paper forms in the field, which will be transcribed into digital CSV spreadsheets at the end of the day. Paper forms will be stored in a filing cabinet at EWU. CSV files will be saved in a Dropbox folder archived and privately shared between the project participants.

Data will be backed up using Dropbox cloud storage. The PI has a Professional Dropbox account which allows 1 TB of storage. While there are no human participants in the project, there are potential access restrictions regarding exact location information.

Collaborators, such as SRDH, the City of Spokane and the City of Cheney will be consulted before sharing location-specific information with third parties. Restrictions will be imposed by removing spatial identifying information from CSV files.

The final format of the data will be both CSV and ArcGIS shapefiles.

Data and metadata will be deposited in DataOne.org. Identifying information to specific locations will be subject to approval of collaborators and stakeholders such as SRDH, City of Spokane and City of Cheney. The point of contact will be the PI, Krisztian Magori.

There are no associated costs.

4. Publish and Share

All data will be shared with the collaborators (SRDH, WADOH, City of Spokane and City of Cheney). A decision will be made after consultation with them about the granularity of the data that can be made available to the public, particularly in terms of spatial resolution. Data deemed to be releasable to the public will be then shared as either CSV files (and possibly ArcGIS shapefiles) on DataOne.org.

Restrictions might apply in terms of specific geographic locations.

Access restrictions can be overcome after discussion and special authorization by SRDH and the City of Spokane and the City of Cheney.

Peer-reviewed article will be written based on the data collected. The graduate student

will use the data collected for his thesis project, which will be published as a peer-reviewed article. In addition, the results of the study will be informally described as blog posts and web pages.

Metadata will be stored along with the data on DataOne.org, made available for harvest by the USGS Science Data Catalog. We will create a DOI for the data using the Digital Object Identifier Creation Tool or the Open Science Framework.