Genomic techniques for identifying viruses in bats

Laura Bergner
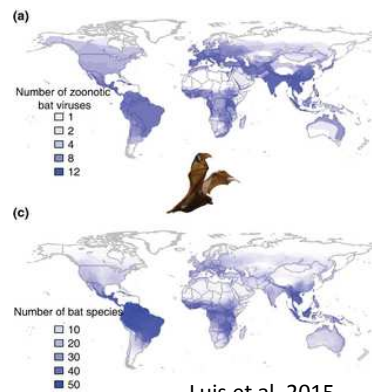
Richard Orton, Daniel Streicker

© Dietmar Nill / naturepl.com

Bats and viruses

- Reservoir of many emerging infectious diseases
  - Influenza, rabies, SARS
- Host over 200 viruses in 27 viral families

Luis et al. 2015

Viral diversity in bats

- What viruses are present in bats?
- How is the diversity of bat viruses distributed geographically and across species?
- What are the ecological and anthropogenic factors influencing viral diversity in bats?
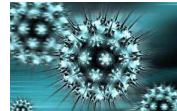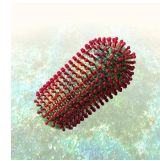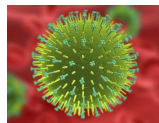
Studying viruses in bats

- Methods for detecting viruses
  - Culture
  - Serology
  - PCR (polymerase chain reaction)

- Only for specific taxa that have been previously characterized
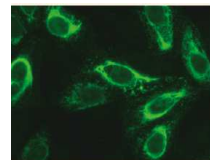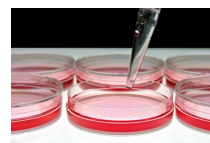
Genomic techniques for studying viral diversity

- Metagenomics
  - Sequence-based characterization of taxa in an environmental or host sample
  - "Virome"
- Illumina sequencing
  - Massively parallel – millions of sequences per run
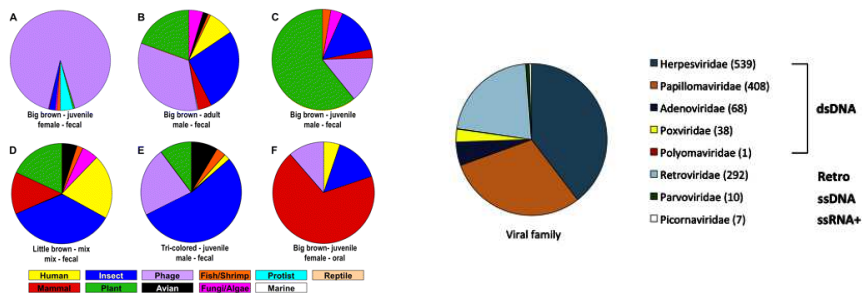 - Shotgun sequencing
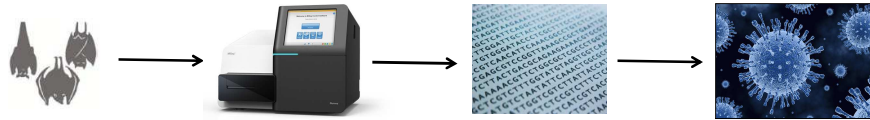
Genomic techniques for studying viral diversity

- Advantages
  - Viral discovery
  - Less biased than other methods
  - Possible data on diet, parasites, bacteria, host genetics?
- Disadvantages
  - Many sequences thrown away
  - Dark matter

Bat viromes

- Insectivorous bats – mostly dietary plant and insect viruses
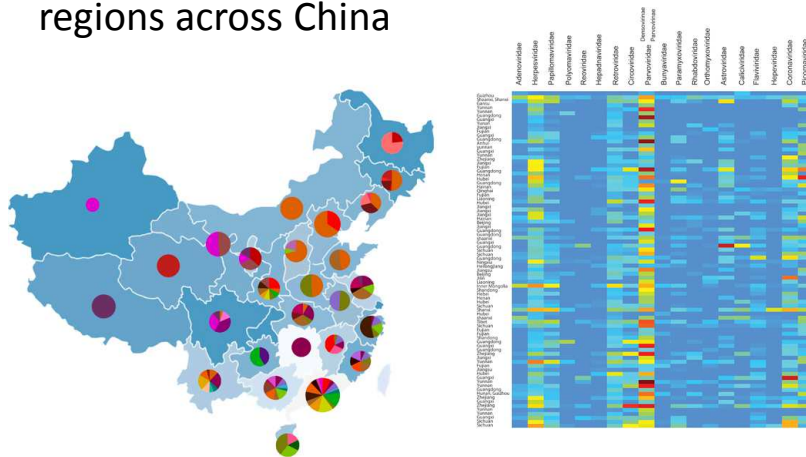- Frugivorous bats – new mammalian viruses
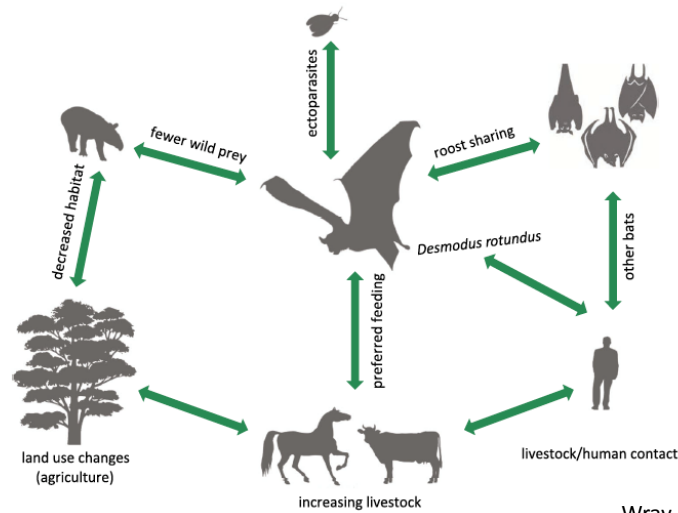

Donaldson et al. 2010

Baker et al. 2013

Bat viromes

- 40 bat species from different ecological regions across China


Wu et al. 2016

Why are vampire bats (*Desmodus rotundus*) interesting?

Wray et al. 2016

Objectives and expectations

1. Characterize viral diversity in *D. rotundus* using metagenomics

2. Investigate ecological and anthropogenic factors influencing viral diversity in *D. rotundus*

- Known viruses in *D. rotundus*: rabies virus, paramyxovirus, coronavirus, polyomavirus, adenovirus
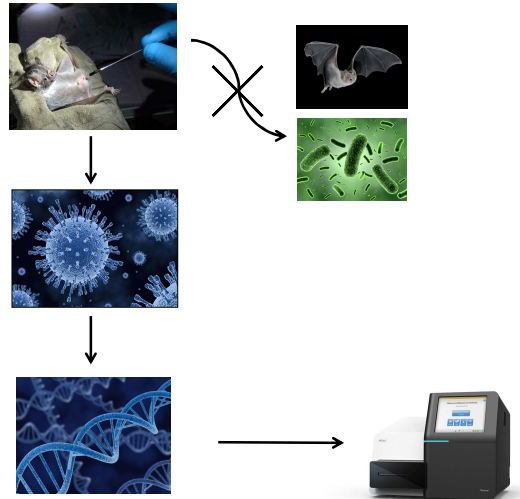- Known in other Neotropical bat species: influenza virus

Study sites
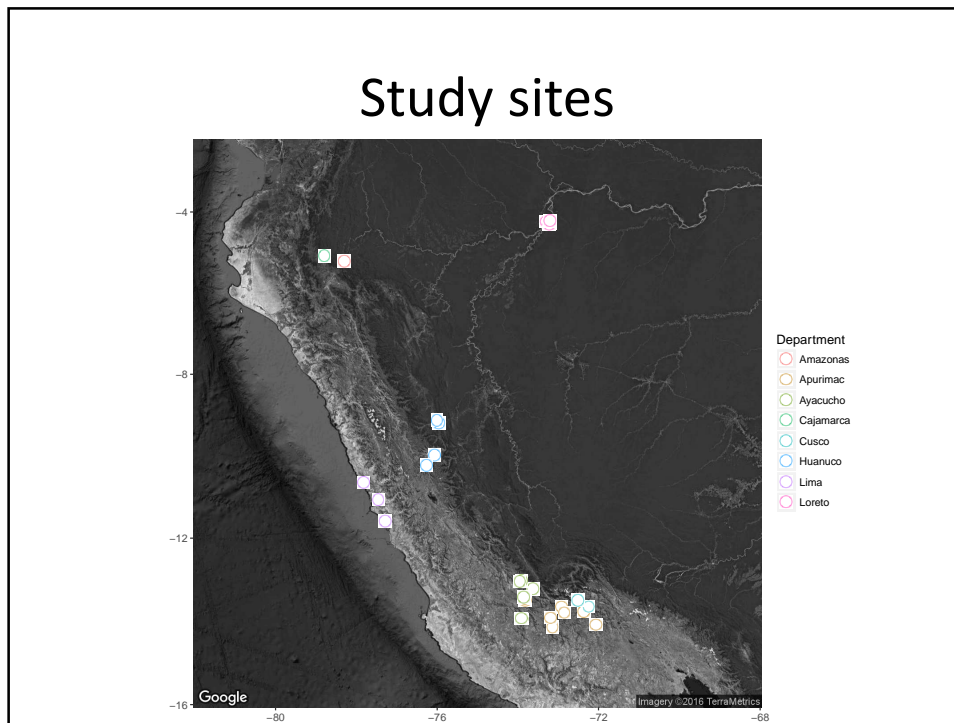
Sampling *D. rotundus*

- Trap bats at known colonies
- Saliva, fecal, whole blood swabs from up to 30 bats per site
- Urine swabs from individuals as possible
- Pooled urine samples from some sites

Laboratory pipeline

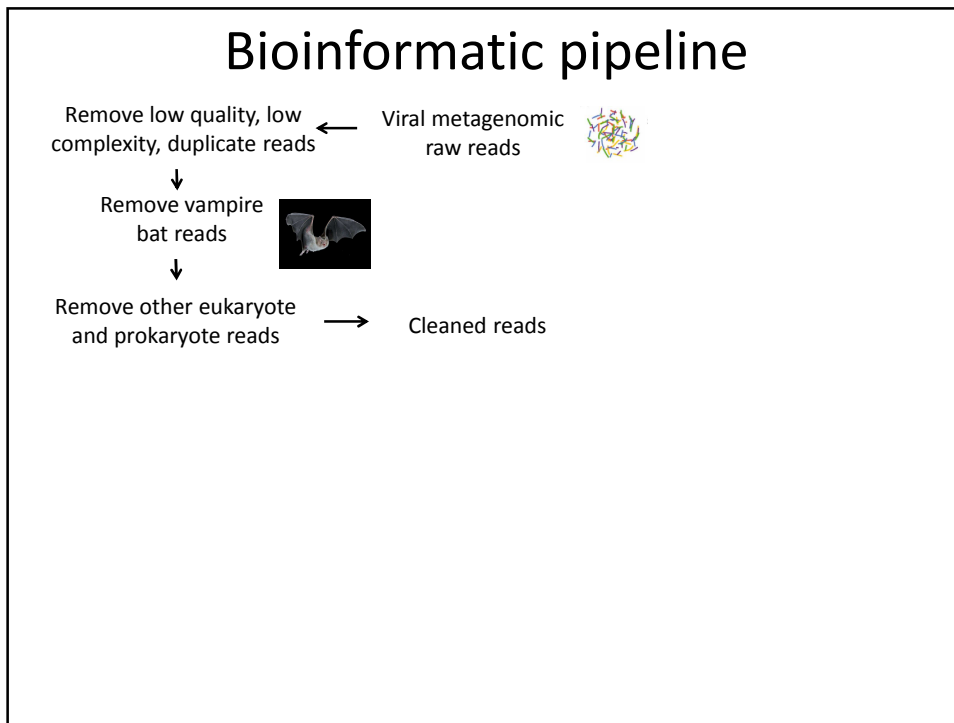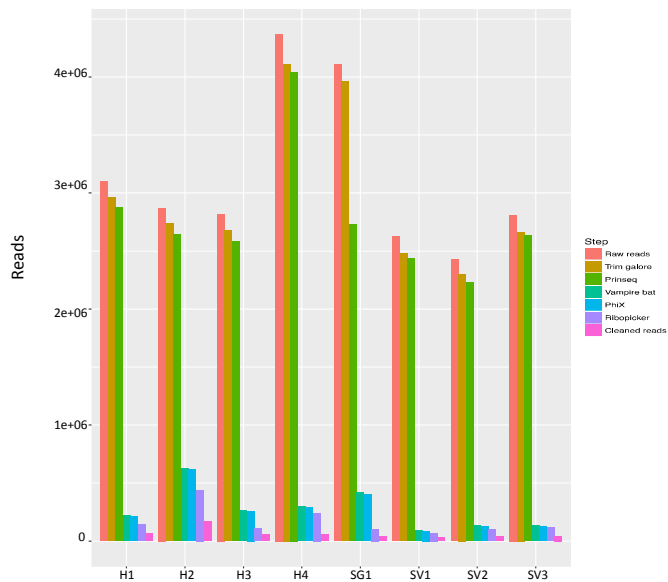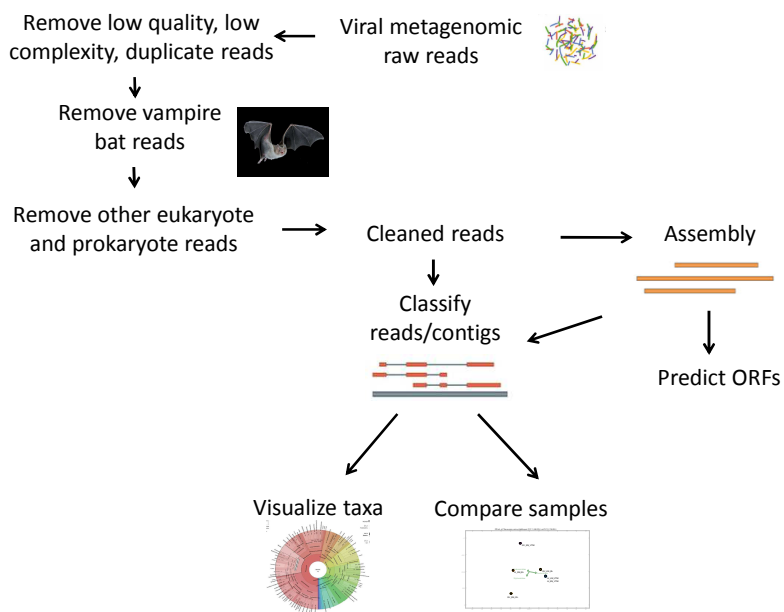# Illumina sequence data

Bioinformatic pipeline

Remove low quality, low complexity, duplicate reads ← ← Viral metagenomic raw reads

↓

Remove vampire bat reads

↓

Remove other eukaryote and prokaryote reads → Cleaned reads

Cleaning reads - NextSeq

Cleaning reads - MiSeq

Bioinformatic pipeline

Viral metagenomic raw reads → Remove low quality, low complexity, duplicate reads → Remove vampire bat reads → Remove other eukaryote and prokaryote reads → Cleaned reads

Cleaned reads → Classify reads/contigs

Cleaned reads → Assembly → Predict ORFs

Classify reads/contigs → Visualize taxa

Classify reads/contigs → Compare samples

Total sequencing reads

Microbiome?

Diet?

Parasites?

Viruses 0.3 %

Pilot run questions

- In what buffer should we collect metagenomic samples in the field?
    (VTM or RNALater?)
- Will VTM work for vampire bat samples?
    – Fetal bovine serum (FBS) may contain cow viruses
- Can we extract samples to enrich for viruses relative to other taxa?
- What viruses do we find?

Pilot run design

Sample type
Fecal swab
Saliva swab

VTM → RNALater

Enriched   Unenriched   Enriched

H_EN_VTM   H_UN_VTM   H_UN_RL   H_EN_RL
SV_EN_VTM   SV_UN_VTM   SV_UN_RL   SV_EN_RL

Same sample, reads vs contigs

Reads

Contigs

Fecal swabs – same individuals, different storage medium

VTM

RNALater

Diverse viruses in FBS

Bovine viral diarrhea virus – a common cell culture contaminant

Viruses inconsistent between sample types, buffer, extraction method

Bacteriophages: Adenoviridae, Myoviridae, Podoviridae, Siphoviridae, unclassified Caudovirales

Caudovirales
dsDNA viruses, no RNA stage
Viruses
Herpesviridae
Iridoviridae
Mimiviridae
Papillomaviridae
dsRNA viruses
Partitiviridae
unclassified dsRNA viruses
Retroviridae
ssDNA viruses
Anelloviridae
Parvoviridae
ssRNA positive-strand viruses, no DNA stage
Flaviviridae
Virgaviridae
unclassified viruses

Legend:
H_EN_VTM   H_EN_RL   H_UN_VTM   SV_EN_RL   SV_EN_VTM   SV_UN_RL

Sample (non)clustering based on viral diversity

Sample type clustering based on viral diversity
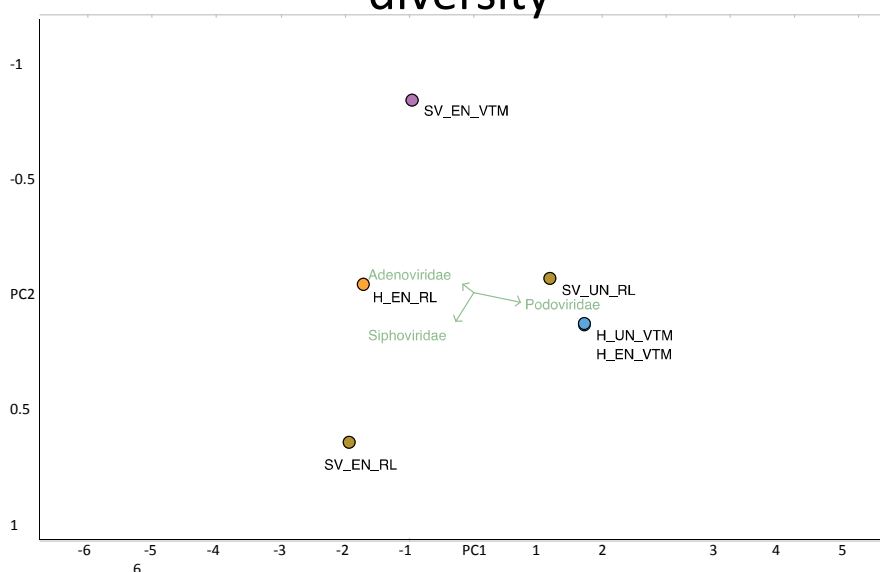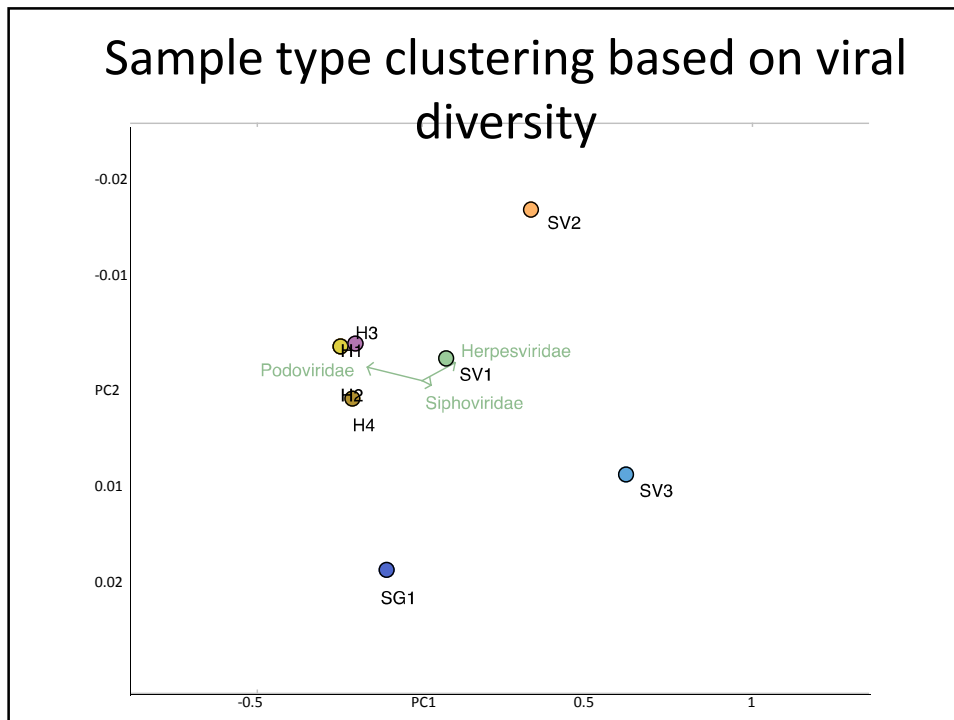
A few viruses in Peruvian vampire bats (so far…)

- Bat hepevirus – 2 sites
- Circovirus – 2 sites
- Bat mastadenovirus
- Bat adeno-associated virus
- Bat anellovirus
- Herpesvirus – all sites, saliva

Summary

- Metagenomic sequencing method works but inconsistent viruses detected between buffers and enrichment methods
- Bacteriophages predominate
- Possibly due to pooled samples and low viral load
- FBS viruses are present but consistent, so we can exclude them from VTM samples

Future plans

- Exclude bacteriophages
- More optimization needed for extraction/pooling
- Replicates
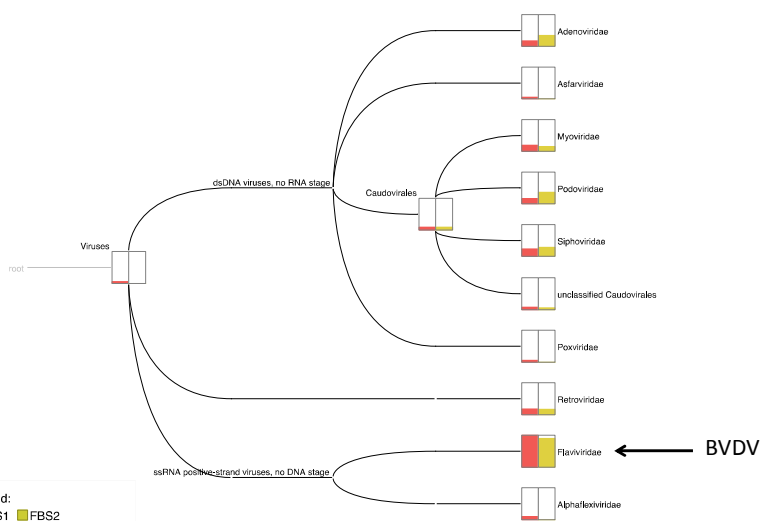- Process samples by colony for larger set of sites across Peru

Thank you!

Diverse viruses in FBS

Viral families in similarly treated samples