MINISTRY OF EDUCATION AND TRAINING
MINISTRY OF HEALTH
NATIONAL INSTITUTE OF HYGIENE AND EPIDEMIOLOGY

TRAN HAI SON

SANDFLIES (DIPTERA: PSYCHODIDAE) AND THE CURRENT SITUATION OF FLAVIVIRUS AND LEISHMANIA INFECTION IN 6 NORTHERN PROVINCES VIETNAM

Specialization: Microbiology
Code: 62 42 01 07

SUMMARY OF DOCTORAL THESIS IN BIOLOGY

Hanoi – 2023

THIS RESEARCH WORK WAS COMPLETED AT CENTRAL INSTITUTE OF HYGIENE AND EPIDEMIOLOGY

Science instructor:

1. Phd. Tran Vu Phong
2. Phd. Associate Professor. Nguyen Le Khanh Hang

Reviewer 1:
Reviewer 2:
Reviewer 3:

The thesis will be defended at the National Institute of Hygiene and Epidemiology.
At the time, date…, month…., year 2024.

The thesis can be found at :

1. National Library
2. Library of the National Institute of Hygiene and Epidemiology

LIST OF PUBLISHED ARTICLES
RELATED TO THE THESIS

1. Vu SN, Tran HS, Tran VP, Tran CT, Tran ND, Dang DA, Nguyen TY, Vu TL, Ngo KP, Nguyen VH, Hoàng NA, Cassan C, Prudhomme J, Depaquit J, Rahola N, Bañuls AL (2021), "axonomical insights and ecology of sandfly (Diptera, Psychodidae) species in six provinces of Northern Vietnam", Parasite. 2021;28:85. doi: 10.1051/parasite/2021080. Epub 2021 Dec 17. PMID: 34928207; PMCID: PMC8686828.
2 . Tran Hai Son, Nguyen Le Khanh Hang, Tran Vu Phong, Tran Cong Tu, Nguyen Viet Hoang, Vu Thi Lieu, Nguyen Thi Yen, Ung Thi Hong Trang, Vu Sinh Nam (2022), "The current status of Leishmania infection in populations of sandfly that was collected in 6 northern provinces of Vietnam, 2016", Journal of Preventive Medicine, volume 32, number 8, 2022.

QUESTION

Sandflies are the main vector of Leishmaniasis, a disease endemic in more than 98 countries with 350 million people at risk and over 2 million new cases every year. In Vietnam, sandfly were first recorded in 1935, and since then, 12 species of sandfly have been recorded distributed from North to South. Most recently in July 2018, Hue General Hospital reported a patient in Quang Binh infected with Leishmaniasis and co-infected with HIV since 2016.

Sandflies in transmitting Flaviviruses is unclear, although there is some evidence of Flaviviruses or Flavivirus RNA related to sandfly such as Saboya virus isolated from sandfly in Senegal (1991-1992), two Flavivirus sequences have been detected in the sandfly *Phlebotomus perniciosus* in Algeria (2007), Ecuador Paraiso Escondido virus (EPEV) in Ecuador (2011) or West Nile virus in Niger (2016). Flavivirus RNA has also been detected in Phlebotomine Sandflies from Portugal. In 2014, Son La province recorded a large-scale outbreak of viral encephalitis (VNVR) lasting from June to September with 164 cases, of which 21 deaths. In recent years, mountainous areas such as Hat Lot and Son La have also continuously recorded Dengue outbreaks ranging from several dozen to several hundred cases.

With the purpose of determining the species composition of sandflies and some of their distribution characteristics in the 6 research provinces as well as describing the infection status of Flavivirus and Leishmaniasis on female sandflies, we conducted the study "Sandflies". (diptera: psychodidae) and the current status of Flavivirus and Leishmaniasis infections in 6 northern provinces of Vietnam". Research with 3 goals:

1) Determine the species composition and some distribution characteristics of Sandflies in 6 northern provinces of Vietnam, 2016-2018.

2) Describe the current status of flavivirus infection in Sandflies at the study site.

3) Describe the current situation of leishmaniasis in Sandflies at the study site.

New scientific points and practical value of the topic :

The study was conducted over a long period of time (2016-2023), so it has high scientific significance when updating the composition of sandfly species at the research site. This is the first study in Vietnam to announce the proportion of female Sandflies carrying Flavivirus and Leishmaniasis.

The results of the study not only supplement the currently lacking knowledge in the medical literature on Sandflies, Flaviviruses, and vector-borne Leishmaniasis, but also help epidemiologists have appropriate strategies in preventing and controlling disease sick.

Research is also especially important at the present time when providing information about two agents Flavivirus and Leishmiania in sandfly vectors, helping to provide more accurate assessments than the trend of research on prevention of Dengue fever or Leishmaniasis.

Structure of the thesis

The thesis includes: 116 pages excluding references and appendices, 12 tables, 39 figures and 1 diagram. Question 2 pages. Overview 45 pages; Research objects and methods 20 pages; Results 26 pages; Discussion 20 pages; Conclusion 2 pages and Recommendations 1 page

Chapter 1. OVERVIEW

1.1. Sandfly and some epidemiological characteristics

Of the more than 800 recognized sandfly species, approximately 464 species was found in the New World and 375 species in the Old World. Phylum: Arthropoda, Class: Insecta, Order: Diptera, Suborder Nematocera, Family: Psychodidae, Subfamily: Phlebotominae (Bigot 1854, K.KertÉsz 1903). The Phlebotominae subfamily is divided into six genera: three from the Old World (Phlebotomus [13 subgenera], Sergentomyia [10 subgenera], and Chinius [4 species]) and three from the New World (Lutzomyia [26 subgenera and group], Brumptomyia [24 species] and Warileya [6 species]). Currently, 78 species of Sandflies have been proven to be vectors of Leishmaniasis. Among the sandfly vectors mentioned above, 7 species are involved in the transmission of *L. major* , 7 species transmit *L. tropica* , 31 species transmit *L. infantum* , and 9 species transmit *L. donovani* .

Sandflies are insects with an complete metamorphosis life cycle. In the development cycle there are 4 distinct phases: egg, larva, pupa and adult. In Vietnam, Sandflies are quite common and can be found in many different habitats. Sandfly have been recognized as vectors of Leishmaniasis in Vietnam since the 1930s.

1.2. Flavivirus and some epidemiological characteristics

1.2.1. General characteristics of Llavivirus group viruses

1.2.1.1. Classify

The family Flaviviridae includes 4 genera: Flavivirus, Hepacivius, Pestivirus and Pegivirus. This genus has more than 53 members including vector-borne diseases caused by Dengue virus causing Dengue fever, Dengue hemorrhagic fever, Dengue shock syndrome; Japanese encephalitis virus; Yellow fever virus causes yellow fever; Chikungunya virus, Kyasanur Forest disease, Murray Valley encephalitis, Omsk hemorrhagic fever, tick-borne encephalitis, West Nile fever, and Zika.

1.2.1.2. Morphological characteristics and genetic material structure of Flavivirus group

Viruses of the Flavivirus group are spherical, 40 - 60 nm in diameter, inside the core of the virus is the nucleocapsid which is the structure of the virus genome and protein C. The nucleocapsid is surrounded by a membrane (viral shell) which is a lipid double layer, contains glycoproteins and proteins derived from the cell's plasma membrane.

The flavivirus genome is a positive-strand RNA with a CAP structure at the 5' end and specifically lacks a poly-A tail at the 3′ end [55]. The viral genome encodes a single polyprotein that, after transcription by viral and host proteases, form 10 structural proteins (C-prM-E-NS1-NS2A-NS2B-NS3-NS4A- NS4B-NS5)

1.2.2. Virus replication belongs to the Flavivirus group

Flaviviruses replicate in the cytoplasm and virus particle assembly occurs in intracellular vesicles.

1.2.3. Antigenic properties

Flaviviruses all share the same antigenic site. At least eight antigenic complexes have been identified based on neutralization experiments.

1.2.4. Laboratory diagnostics

1.2.4.1. Detection of virus RNA

The reaction performed to detect Flavivirus is RT-PCR or Realtime RT-PCR, in which Realtime RT-PCR gives results faster than traditional RT-PCR. The RT-PCR reactions and processes currently used for detection focus on the target gene segment encoding the virus envelope (E-encoding gene), encoding the envelope membrane (M/E-encoding gene), encoding (pE) and encode proteins NS5, NS3, NS1.

1.2.4.2. Virus isolation method

The virus isolation method usually takes 1-3 weeks, so it does not meet the requirements for quick diagnosis, but the results obtained by this method provide a lot of virus information for virus research. biology, pathology and vaccine development

1.2.4.3. Detection of anti-antibodies

Serological methods will be complicated in places where many viruses belonging to the Flavivirus group (DENV, Japanese B encephalitis, yellow fever) circulate.

1.2.5. Diseases caused by Flaviviruses transmitted by vectors are circulating in Vietnam

Diseases caused by Flaviviruses transmitted by vectors circulating in Vietnam include Dengue, JE, and Zika.

1.3. Leishmania and some epidemiological characteristics

1.3.1. Taxonomic rank of Leishmania

Kingdom	Protista (Haeckel, 1866),
Class	Kinetoplastea (Honigberg, 1963 emend. Vickerman, 1976),
Subclass	Metakinetoplastina (Vickerman, 2004),
Order	Trypanosomatida (Kent, 1880),
Family	Trypanosomatidae (Döflein, 1901),
Subfamily	Leishmaniinae (Maslov and Lukeš 2012) (Ross, 1903).

1.3.2. Leishmania parasite and life cycle

There are about 21 species of the genus Leishmania that cause disease in humans . They can be distinguished on the basis of biological criteria, or laboratory analysis (mainly isoenzyme analysis and DNA analysis), or different clinical and epidemiological symptoms.

1.3.3. Genome characteristics of Leishmania

Leishmania has unique genome organization features compared to eukaryotes, such as intronless genes, polycistrons, and small chromosomes with high gene density. Furthermore, these flagellates possess a single mitochondria called the kinetoplast, which contains a large network of kinetoplast DNA (kDNA).

1.3.3.1. Chromosomal DNA

Ribosomal RNA (rRNA) genes are located mostly on chromosome 27, often existing in multiple copies with a size of approximately 12.5 kb. Among the various components of these genes, the ITS regions are ideal for species identification. The 18S rRNA is a structural RNA of the ribosomal SSU. The high conservation of this gene and its flanking regions makes it suitable for reconstructing phylogenetic relationships.

1.3.3.2. Genes encode proteins

Members of leishmania possess 36 chromosomes, except for the *L. mexicana complex* which has 34 chromosomes. The Viannia genus has 35 chromosomes. The Sauro Leishmania subgenus has 38 chromosomes. The Leishmania genome is compact with a size of 33 Mb.

1.3.3.3. Extra-chromosomal DNA

All kinetoplastid flagellates possess a unique mitochondrial genome called kDNA, which consists of several thousand circular DNA molecules linked together in an interconnected network of thousands of minicircles (approx. 1 kb per ring) and several dozen maxicircles (approximately 23 kb per ring).

1.3.4. Method for diagnosing Leishmaniasis in the laboratory

1.3.4.1. Method for determining Leishmaniasis

Different diagnostic methods and the possibility of detecting, identifying and quantifying Leishmania species, as well as their ability to differentiate at different levels (genus, subgenus, species, species complex, species and population)

1.3.4.2. Method for distinguishing Leishmania species complexes

Identification and differentiation of leishmania species complexes and species can be accomplished through various molecular biological techniques.

Nested PCR and semi-nested PCR can be used to distinguish species with appropriate primers.

The Sanger method is rapidly improving in quality, read length, speed, and cost, and it is widely used for identification of leishmania species complexes and phylogenetic studies.

Next generation sequencing (NGS) technology in the analysis of the Leishmania genome has recently facilitated the discovery of various genetic diversity including single nucleotide polymorphisms (SNPs), variants copy numbers (CNVs), structural variations in detail and provide valuable insights into the complexity of the genome and gene regulation. Genome analysis of Leishmania presents a challenge because of the frequent presence of aneuploidy. This compromises the accuracy of detecting all genetic variations.

1.3.5. Leishmaniasis and some epidemiological characteristics

1.3.5.1. Epidemiology of Leishmaniasis in the world

Leishmaniasis affects mostly poor people in developing countries; 350 million people are considered at risk of Leishmaniasis, and about 2 million new cases occur every year. Approximately 95% of CL cases occur in the Americas, Mediterranean basin, Middle East and Central Asia with 132,568 cases reported in this regions in 2012. In its report on the health burden of infectious and parasitic diseases worldwide, Hotez ranked Leishmaniasis 9th with 2,357,000 cases annually, mainly occurring in Africa, Southeast Asia, and East Asia. Mediterranean, Western Pacific, America and Europe, of which Southeast Asia is the hardest hit with about 67.3% of the total number of cases worldwide.

1.3.5.2. Epidemiology of Leishmaniasis in Vietnam

From 1978 to 2018, cases of the disease were reported sporadically, mainly in the Northern region of Vietnam. A total of 6 patients were reported to be related to the leishmaniasis parasite, including 4 cases with confirmed co-infection with HIV.

1.3.6. Clinical features, treatment and prevention

According to clinical characteristics, the disease can be divided into the following 3 types:

1.3.6.1. CL-cutaneous Leishmaniasis

Leishmaniasis CL resides in the skin, with an incubation period of several weeks to several months. The causative agent is infection with *L. major* or *L. tropica*

1.3.6.2. Visceral Leishmaniasis (VL- Visceral Leishmaniasis)

Visceral Leishmaniasis resides in internal organs, the incubation period in most cases is 3-6 months, in some cases several weeks to several years. The main symptoms are fever, spleen swelling, increased blood Iggamma, acute anemia, and leukopenia. Without treatment, the patient will die within two years from complications of wasting and secondary infection.

1.3.6.3. Mucocutaneous Leishmaniasis (MCL)

Leishmaniasis of the skin and mucous membranes is less common.

1.3.6.4. Treatment and prevention

Treatment of VL is usually performed with agents belonging to the group of pentavalent antimonials (meglumin antimonate, sodium stibogluconate) pentamidine, or amphotericin B.

Chapter 2 . RESEARCH AND METHODS

2.1. Objective research methods 1

Content: Determine the species composition and some distribution characteristics of Sandflies in 6 northern provinces of Vietnam, 2016-2018.

2.1.1. Research location

The six provinces selected for the study are: Quang Ninh, Ninh Binh, Lang Son, Lao Cai, Ha Giang and Son La. (Figure 3.1)

2.1.2. Research object: Sandfly, phylum Arthropoda, class: Insecta, order: Diptera, Family: Psychodidae, Genus: Phlebotominae.

2.1.3. Sample collection time: From 30th May, 2016 to 13rd October, 2016

2.1.4. Research design: Cross-sectional descriptive survey

2.1.5. Sample size: Use the total sampling method.

2.1.6. Methods of collecting Sandflies in different habitats

2.1.6.1. Sample collection method: Use CDC miniature light traps (CDC miniature light traps, John W. Hock Co. FL, USA) to collect Sandflies.

2.1.6.2. Method for screening Sandflies in the field: Distinguishing Sandflies from other types of mosquitoes based on typical morphological characteristics. Sample storage: Male sandfly: store in 1.5ml tubes containing 70% alcohol, female sandfly: store in 1.5ml tubes in liquid nitrogen.

2.1.7. Method of making specimens: Performed at the Department of Entomology and Medical Animals - National Institute of Hygiene and Epidemiology

2.1.8. Sandfly identification method: Identification based on the identification key of Lewis (1978, 1987) and Killick Kendrick et al. (1991) adds comparisons with the descriptions of Newstead (1911), Raynal (1936), Abonnenc E. 1972, Johnson H. 1991 and Lewis 1982 [11, 25, 181-184]. Sample images were observed with the camera system on a nilkon E600 electron microscope and analyzed with NIS-Elements software. The images of species identification results were sent and confirmed at the Montpellier Research Institute, France (IRD).

2.1.9. Output indicators in research

Abundance: RA = (Total number of species/Total number of individuals) x 100

Density: D = Total number of individuals/Total number of traps set/Number of nights of setting traps

Significance level: Mean = Total number of individuals collected/Total number of traps set

Number of species: SR = Number of species in the collected habitat (including *Se. sp2* and *Se. sp3*)

2.1.10. Data entry and analysis

Data were entered using Excel and analyzed using Stata ver 14 and Excel software. Images were taken and measured using NIS-Elements software. Use Kruskal–Wallis test statistical analysis to compare the distribution of Sandflies by province and habitat.

2.2. Objective research method 2

Content: Describe the current status of flavivirus infection in Sandflies at the research site.

2.2.1. Research location: Department of Entomology and Animal Medicine

2.2.2. Research object: Viruses in the genus Flavivirus, family Flaviviridae in specimens collected at target 1.

2.2.3. Study design: Cross-sectional description with laboratory analysis

2.2.4. Sample size: All thorax and abdomen samples of female Sandflies were collected: 1009 samples.

2.2.5. Biological products and equipment: DNA/RNA extraction products: Proteinase K; β-mercaptoethanol 48.7%; Chloroform: isoamyl alcohol (24:1): 2-propanol; d CTAB solution. PCR biological products: QIAGEN Onestep RT-PCR; Primer pair for RT-PCR reaction -Standard control: + Positive control (Positive control – POS): DEN 1-4 (Arbo Virus Laboratory, Department of Virology, National Institute of Hygiene and Epidemiology); Negative control (No Template Control - NTC): use distilled water to check the process of mixing chemical biological products; Negative Extraction Control (NEC): use distilled water to check the extraction process

2.2.6. Identify/identify Flaviviruses using RT-PCR technique

Judging results: results are accepted when: Positive control: has a specific band equivalent to the size of the designed primer pair (250 bp); Negative reaction control (NTC), negative extraction control (NEC): negative. Negative for Flavivirus group: presence of PCR products in non-specific positions or no presence of PCR products. Positive for Flavivirus group: specific PCR product has a size of 250 bp. Samples positive for the Flavivirus group will be subjected to Sanger gene sequencing

2.2.7. Gene sequencing using the Sanger method

Purification of PCR products: Purification of PCR products. Use biological product: ExoSAP-IT™ PCR Product Cleanup Reagent. The procedure was performed according to the manufacturer's instructions.

Synthesis of PCR products for gene sequencing (PCR sequencing reaction):

Each reaction gives one primer, cFD2 and MAMD, respectively

Purification of gene sequencing PCR products: Gene sequencing PCR products were purified using the Dye Ex 2.0 Spin kit.

Run the ABI 3100-AvantTM Genetic Analyzer .

2.2.8. Data entry and analysis

Data were entered using Excel and analyzed using Stata ver 14 and Excel software. Gene sequence analysis using the BLAST function on NCBI, identification on the Web Flavivirus Genotyping Tool Version 0.0.

2.3. Objective research method 3

Content: Describe the current situation of Leishmaniasis in Sandflies at the research site.

2.3.1. Research location: Department of Entomology and Animal Medicine - Department of Virology - National Institute of Hygiene and Epidemiology

2.3.2. Research object: Leishmania parasites on female Sandflies collected in target 1.

2.3.3. Study design: Cross-sectional description with laboratory analysis

2.3.4. Sample size: All thorax and abdomen samples of female Sandflies were collected: 1009 samples.

2.3.5. Biological products and equipment

- DNA/RNA extraction biological products: similar to section 2.2.5
- Nested PCR biological products: GoTaq; Primer pair for Nested-PCR reaction

Standard certification:

+ Positive control (POS): *Leishmania infantum* 680 bp.
+ Negative control (No Template Control - NTC): use water that does not contain DNase-RNase to check the process of mixing chemical products
+ Negative Extraction Control (NEC): use water that does not contain Dnase-Rnase to check the extraction process
- Equipment: similar to section 2.2.5.

2.3.6. Nested PCR reaction

Prepare biological products for Nested-PCR reaction: according to instructions of Gotaq biological kit (Promega), outer primer pair CBS2XF-CBS1XR, internal primer pair LIR-13Z

Judging results: Results are accepted when:

+ Positive control: has a specific band equivalent to the size of the designed primer pair of 680 bp.
+ Negative reaction control (NTC), negative extraction control (NEC): negative.
- Negative for leishmania group: appearance of PCR products in non-specific positions or no presence of PCR products.
- Positive for Leishmania group: specific PCR product with size over 500 bp. Corresponding band size: *L. amazonensis*

MHOM/BR/73/LV78 (517bp); *L. major* MHOM/ET/95/FV1 (560-570 bp); *L. infantum* (680 bp); *L. tropica* (750 bp) [130].

- Samples positive by Nested-PCR are subjected to genetic sequencing (NGS).

2.3.7. Next generation sequencing (NGS) gene sequencing method

Use biological products and chemicals from the Nextera XT DNA Library Prep kit. Equalize the sample library with the ISEQ 100 machine using the Standard Normalization - Illumina method.

2.3.8. Blood meal analysis

Sample size: DNA from female Sandflies recorded as blood-fed and positive for Leishmania will be used to determine blood meals. This is to determine whether this Leishmania strain related to the sandfly's blood meal is human or animal blood. However, in this study, we did not conduct blood meal analysis because the samples positive for Leishmaniasis in the study did not record blood.

2.3.9. Data entry and analysis

Data were entered using Excel and analyzed using Stata ver 14 and Excel software. Data after sequencing, use FastQ file for analysis.

Chapter 3. RESULTS

3.1. Species composition and some distribution characteristics of Sandflies in 6 northern provinces of Vietnam, 2016-2018

3.1.1. Sandfly species composition by breed, density and abundance

Table 3.1. Number, sex, density and abundance of Sandflies by species in 6 northern mountainous provinces of Vietnam, 2016-2018

Loài	Số lượng	Cái/Đực[*n]	Mật độ	Độ phong phú
Sergentomyia (Neophlebotomus) sylvatica	249	87/161 ([*1])	0,0253	9,632
Sergentomyia (Parrotomyia) brevicaulis group	66	49/17	0,0067	2,553
Segentomyia (Parrotomyia) barraudi group	324	303/21	0,0329	12,534
Sergentomyia (Sergentomyia) bailyi	55	44/11	0,0056	2,128
Sergentomyia (Neophlebotomus) hivernus	49	46/3	0,0050	1,896
Sergentomyia (Neophlebotomus) perturbans	11	5/6	0,0011	0,426
Sergentomyia (Neophlebotomus) khawi	25	7/18	0,0025	0,967
Sergentomyia sp2	201	140/58 ([*3])	0,0204	7,776
Sergentomyia sp3	10	8/2	0,0010	0,387
Sergentomyia und_sp	83	65/18	0,0084	3,211
Sergentomyia (Neophlebotomus) sp.	4	4/0	0,0004	0,155
Sergentomyia sp.	990	50/928 ([*12])	0,1006	38,298
Grassomyia indica	6	1/5	0,0006	0,232
Phlebotomus (Anaphlebotomus) stantoni	102	46/55 ([*1])	0,0104	3,946
Phlebotomus (Euphlebotomus) yunshengensis	87	28/59	0,0088	3,366
Phlebotomus (Larroussius) betisi	50	3/47	0,0051	1,934
Phlebotomus (Euphlebotomus) mascomai	35	17/18	0,0036	1,354
Phlebotomus (Euphlebotomus) sp.	21	5/16	0,0021	0,812
Phlebotomus (Larroussius) sp.	12	4/8	0,0012	0,464
Phlebotomus sp.	33	25/8	0,0034	1,277
Idiophlebotomus sp.	14	6/8	0,0014	0,542
Chinius junlianensis	31	27/4	0,0031	1,199
NA	127	79/40 ([*8])	0,0129	4,913
Tổng số	**2585**	**1049/1511 ([*25])**	**0,2626**	**100**

[*n] số lượng không thể phân biệt cái/đực, NA: số lượng không thể định loài.

Identification results show that there are 5 genera of Sandflies: Sergentomyia (n=2067, 79.96%) accounts for the highest proportion, followed by Phlebotomus (n=340, 13.15%), Chinus (n=31, 1.2%), Idiophlebotomus (n=14, 0.54%) and Grassomyia (n=6, 0.23%) (Table 3.1).

A total of 13 species have been identified, of which Sergentomyia has 7 ; Phlebotomus species has 4 species and Chinius has 1 species.

3.1.2. Distribution of Sandflies by province

The Sergentomyia is the most dominant in the 6 provinces. Statistical analysis shows that the distribution by breed between provinces is not significantly different. But the species distribution, including *Se. sp2* and *Se. sp3* , is very different between provinces (p-value =0.002, α=0.05).

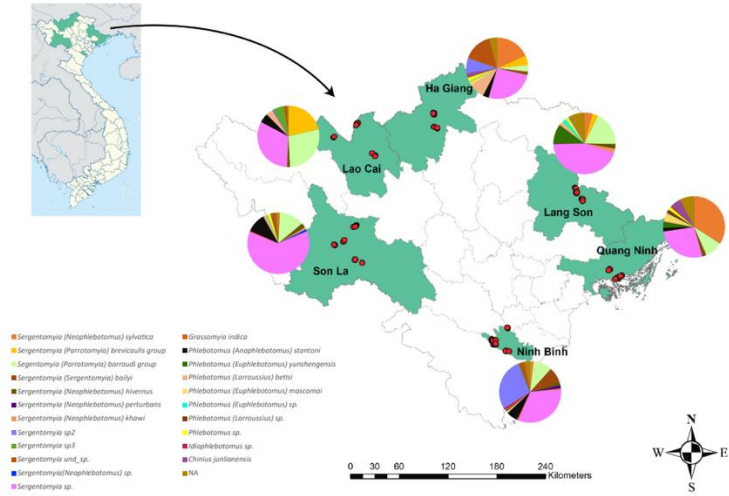

Figure 3.1. Collection points and composition of sandfly species in 6 northern provinces of Vietnam, 2016

3.1.3. Distribution of Sandflies according to habitat

Sandflies were most commonly collected in caves (n = 1431, relative abundance RA = 55.36 and $_{\text{cave}}D = 0.79$). The highest species richness was in caves ($_{\text{cave}}SR = 15$, including *Se. sp2* and *Se. sp3*). Regarding relative abundance, we also collected many outdoor Sandflies, with 936 specimens corresponding to RA=36.21 and species richness SR=15 (including Se. *sp2 and Se. sp3*). However, the density of Sandflies collected by traps placed in dog cages was higher than that placed outside the house ($D_{\text{cages). dog}} = 0.36$, $D_{\text{outside the house}} = 0.23$). The density of Sandflies indoors is low and similar to the density in chicken/poultry/duck coops and lower than the density in buffalo/cow/goat barns ($_{\text{Indoor}}D = 0.08$; $D_{\text{poultry barn}} = 0.10$; $D_{\text{barn}} = 0.12$). This distribution according to habitat is not significantly different among the 6 provinces. In caves, the number of species is the highest, all varieties and species have been found (Table 3.3 and figure 3.1). Statistical analyzes show that the distribution of species is different depending on habitat (p-value <0.01, $\alpha = 0.01$).

Table 3.3. Number of Sandflies, abundance, density and number of species according to habitat

Tên loài	Chuồng gia súc	Hang	Chuồng gia cầm	Chuồng chó	Trong nhà	Ngoài nhà
Sergentomyia (Neophlebotomus) sylvatica	1	141	14	1		92
Sergentomyia (Parrotomyia) brevicaulis group		45				21
Segentomyia (Parrotomyia) barraudi group	4	172	4		1	138
Sergentomyia (Sergentomyia) bailyi	15	10	5		3	17
Sergentomyia (Neophlebotomus) hivernus	3	10	5	1		28
Sergentomyia (Neophlebotomus) perturbans		6			1	4
Sergentomyia (Neophlebotomus) khawi	3	16			1	5
Sergentomyia sp2	5	135	2		3	55
Sergentomyia sp3		9				1
Sergentomyia und_sp		20	1		2	60
Sergentomyia (Neophlebotomus) sp.		4				
Sergentomyia sp.	12	566	12	2	6	378
Grassomyia indica		4				2
Phlebotomus (Anaphlebotomus) stantoni	11	32	8	7	11	24
Phlebotomus (Euphlebotomus) yunshengensis		80		1		6
Phlebotomus (Larroussius) betisi		19	1			30
Phlebotomus (Euphlebotomus) mascomai	2	24		1		8
Phlebotomus (Euphlebotomus) sp.		14	2			4
Phlebotomus (Larroussius) sp.		10	1	1		
Phlebotomus sp.	3	15	2	1		11
Idiophlebotomus sp.		7				7
Chinius junlianensis	1	24				6
NA	5	68	9	1	2	39
Tổng	65	1431	66	16	30	936
Độ phong phú (RA)	2,51	55,36	2,55	0,62	1,16	36,21
Mật độ (D)	0,12	0,79	0,10	0,36	0,08	0,23
Mức ý nghĩa (Mean)	1,44	15,66	1,57	1,78	1,07	5,29
Số loài (SR)*	9	15	7	5	6	15

NA: số lượng không thể định loài; * bao gồm cả *Se. sp2* và *Se. sp3*

3.1.4. Distribution of female Sandflies in 6 mountainous provinces in Northern Vietnam, 2016

Female Sandflies were collected in 6/6 provinces in the study. The highest number of female Sandflies collected was in Ninh Binh and Lang Son with 298 individuals (28.41%) and 294 individuals (28.03%), followed by Ha Giang with 159 individuals (15, 16%), Quang Ninh 125 individuals (11.92%), Son La with 98 individuals (9.34%) and Lao Cai 75 individuals (7.15%).

3.2. Actual Flavivirus infection in Sandflies at the study site

3.2.1. Screening for Flaviviruses in female Sandflies

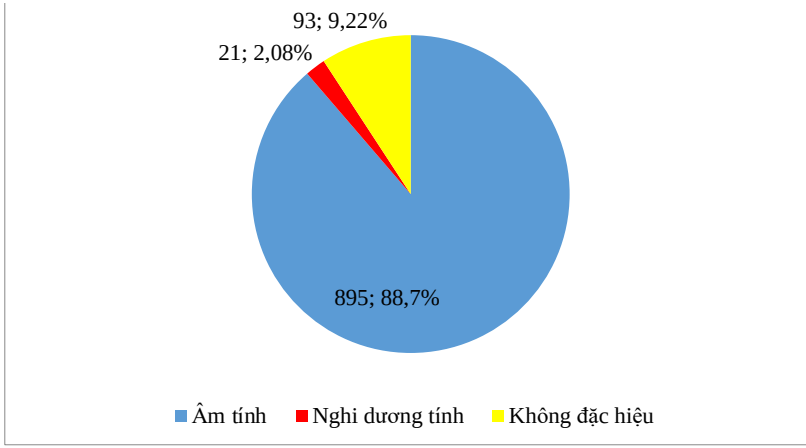

Figure 3.4. Results of Flavivirus screening on female Sandflies

Results of Flavivirus screening on Sandflies showed that 21/1009 samples were suspected to be positive for Flavivirus or Flavivirus RNA related to Sandflies (2.08%). 895 samples were determined to be negative for Flavivirus, accounting for 88.7%, the remaining 93 samples had PCR products, but were not specific to flavivirus, these samples accounted for 9.22% (Figure 3.4). All 21 samples positive for Flavivirus by RT-PCR method will be subjected to gene sequencing to confirm the presence of Flavivirus in the collected sandfly population.

3.2.2. Identification of Flaviviruses by Sanger sequencing method

Results obtained after sequencing of 21 samples: 16 samples in which we did not find any information when comparing the sequences of those samples to the genetic database of NBCI. There are 3 samples with sequences showing similarities to flavivirus, but the length of the homologous segment is too short (<31bp) to determine. There are 2 samples number 4 (M2.25.56) and 17 (M3.57.07) whose sequences were determined to be related and similar to Dengue type 2 (DEN2).

Table 3.4. Location of RNA fragments obtained in the NS5 gene region of Flavivirus

Mẫu	Vị trí bắt đầu	Độ dài	Vị trí cuối cùng	Kết quả BLAST	Kết quả cluster
M2.25.56	8928	223	9159	Flaviviridae Flavivirus Dengue virus	Dengue virus 2
M3.53.07	8952	237	9189	Flaviviridae Flavivirus Dengue virus	Dengue virus 2

3.2.3. Flavivirus characteristics in female Sandflies

Among the 6 investigated provinces, screening results recorded the appearance of ARN fragments of Flaviviruses in general and DEN2 in female Sandflies in 2 provinces, Ninh Binh and Lang Son. Ha Giang, Lao Cai and Quang Ninh, Son La have not found traces of flavivirus in sandfly populations. The proportion of female Sandflies carrying ARN of flavivirus (DEN2) in the female sandfly population in our study is 0.198% (n=2/1009).

Table 3.5. Information on samples suspected of being infected with f lavivirus in sandfly populations

Tỉnh	Vĩ độ	Kinh độ	Muỗi cát	Sinh cảnh	Flavivirus
Ninh Bình	20°13.983’	105°42.704’	*Sergentomyia sp2*	Hang	DENV2
Lạng Sơn	21°56.069’	106°41.061’	*Sergentomyia barraudi*	Hang	DENV2

3.3. Current status of leishmaniasis in Sandflies

3.3.1. Screening for leishmaniasis using the Nested-PCR method

Of the 1009 samples screened, 20 were suspected of being infected with leishmaniasis and 989 were negative.

Table 3.6. 20 samples suspected of being infected with Leishmaniasis by Nested-PCR method

Nhóm mẫu	Số lượng	Ký hiệu mẫu	Kích thước sản phẩm (bp)
1	02	Mẫu 68 (Hình 3.7A-L1) Mẫu 895 (Hình 3.7B-L6)	300
2	06	Mẫu 79 (hình 3.7A-L4) Mẫu 99 (hình 3.7A -L2) Mẫu 906, 907, 909, 910 (hình 3.8B -L7)	500
3	05	Mẫu 96 (hình 3.7A -L1) Mẫu 181(hình 3.7B -L5) Mẫu 900, 897 (hình 3.7B -L6) và 912 (hình 3.7B -L7)	600-750
4	07	Mẫu 69 (hình 3.7A -L2), Mẫu 75, 92 (hình 3.7A -L3) Mẫu 904, 905, 899 (hình 3.7B - L6) và 663 (hình 3.7B – L8)	>=800

All 20 samples will be genetically sequenced to confirm the presence of Leishmania in the collected sandfly population.

3.3.2. Identify Leishmania by NGS gene sequencing method

Gene sequencing using NGS method with Nextera XT DNA Library Prep kit. Equalize the sample library with the ISEQ 100 machine using the Standard Normalization - Illumina method.

The results obtained after sequencing 20 suspect samples resulted in a total of 121,651 sequence fragments (10GB) with sizes ranging from 20 bp to 1728 bp (average 73,945 bp). We removed sequences that were too small in size (under 200bp) to obtain 71 gene sequences for analysis. Of the 71 sequences obtained, 40 sequences had no information found when Blasted to NBCI's gene database, 21 sequences showed similar results but the species information obtained was not related to Leishmania. There are only 10 sequences identified as Leishmaniasis belonging to 3 samples: 2, 4, 20. Samples 2 and 20 have gene sequences similar to classes 13, 16, 18, 25, 31, 38, 39 and 50 on the minicircle

kinetoplas gene of *Leishmania infantum*, while sample number 4 has a sequence similar to chromosome 27 of leishmania, from position 251340 to position 268948.

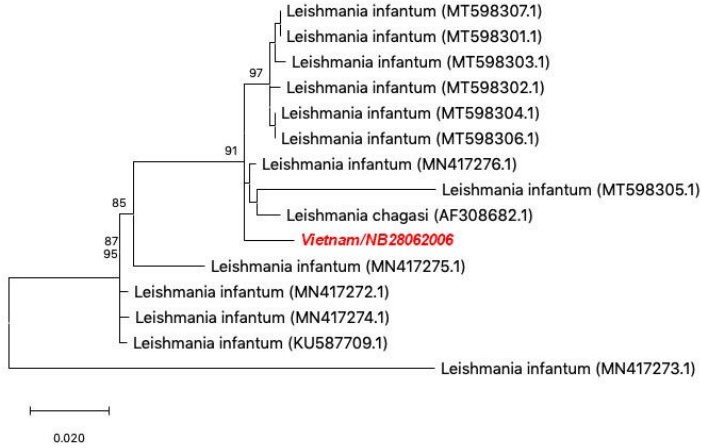

Figure 3.7. Phylogenetic tree of Leishmania in the female sandfly population in Vietnam (Maximum likelihood method, using kDNA genes)

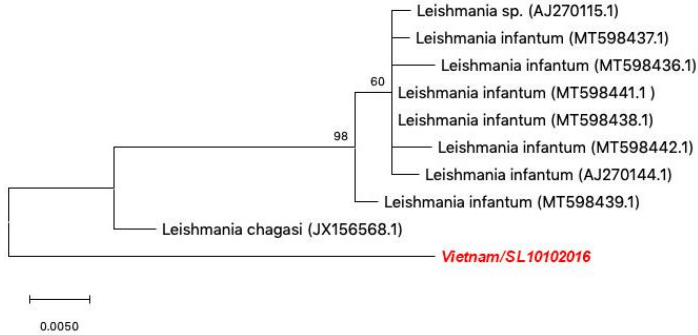

Figure 3.8. Phylogenetic tree of Leishmania in the female sandfly population in Vietnam (Maximum likelihood method, using kDNA genes)

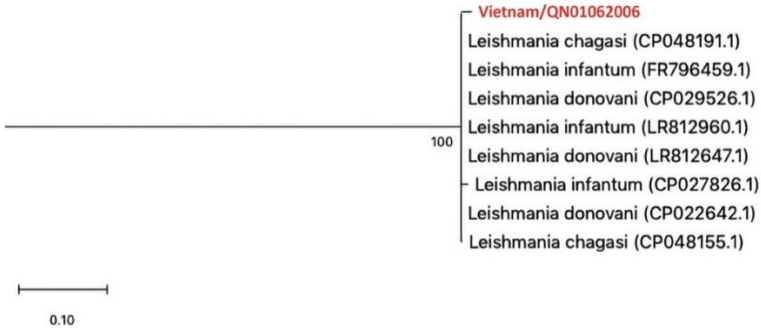

Figure 3.9. Phylogenetic tree of Leishmania in the female sandfly population in Vietnam (Maximum likelihood method, using genes on chromosome 27)

Thus, the gene sequencing results showed that of the 3 samples with similarities to leishmania, 2 samples were identified as *Leishmania infantum* and 1 sample was identified as *Leishmania sp.*

3.3.3. Some characteristics of leishmaniasis in sandfly populations

Among the 6 provinces investigated, screening results recorded the presence of leishmaniasis in 3 provinces: Son La, Quang Ninh and Ninh Binh. The remaining provinces of Ha Giang, Lao Cai and Lang Son have not found traces of Leishmania in sandfly populations. The rate of leishmaniasis infection in the female sandfly population in our study was 0.297% (n=3/1009).

Table 3.9. Leishmania information on sandfly populations

nh	Vĩ độ	Kinh độ	Muỗi cát	Sinh cảnh	Leishmania
ı La	21°11.063’	104°03.277’	*Phlebotomus sp.*	Chuồng lợn	*Leishmania infantum*
ʒ Ninh	20°59.615’	107°12.592’	*NA*	Hang	*Leishmania sp.*
Bình	20°14.457’	105°41.371’	*Sergentomyia sp2*	Ngoài nhà	*Leishmania infantum*

Chapter 4. DISCUSSION

4.1. Species composition and some biological characteristics of Sandflies in 6 mountainous provinces in northern Vietnam, 2016-2018

4.1.1. Identification of sandfly species in Vietnam by morphological characteristics

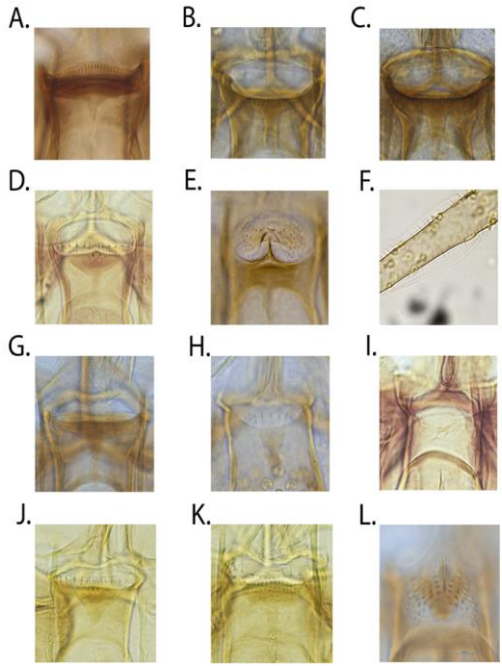

Figure 4.3. Cibarium and ascoid morphology of females *Grassomyia indica* (A), *Sergentomyia barraudi group* (B, C), *Se. khawi* (D), *and Se. anodontis group (E), ascoid on Antennae 3 of Se. sp.3* (F) , *Se. brevicaulis group* (G), *Se. sylvatica* (H) , *Se. bailyi* (I), cibarium *of Se. hivernus* (J), cibarium *of Se. perturbans group* (K), *cibarium of Idiophlebotomus sp.* (L).

4.1.2. Ecology and habitat of Sandflies in the investigated provinces

The highest species richness values were found in caves and outdoors (SR=15 including *Se. sp2* and *Se. sp3*). In barns and homes, species richness ranged from 5 to 9, suggesting more anthropophilic behavior or attraction to domestic animals for some species such as *Ph. stantoni*. In fact, although this species was found in all environments, it was the main species collected mainly in homes (11/30 specimens; 36.67%) and in dog houses (7/16 samples; 43.75%). In caves, this species accounts for only 2.24% (32/1431 specimens). Further studies on the feeding preferences of this species are needed as it has been described as a cave species in Thailand (SRhang=26) and Malaysia (SRhang=18, n= 1548 **).**

4.2. Current status of Flavivirus infection in Sandflies at the research site

4.2.1. Identification of Flavivirus in Sandflies

In our study, the patient sample after having the RT-PCR product amplified with the cFD2/MAMD primer pair (250 bp) was sequenced to confirm the gene fragment.

The results of the study identified two sequences similar to the DEN2 virus. The difference in the study is that the sample we screened was the mash of female Sandflies, while in Scaramozzino's study human samples were used. This shows that the DEN2 blister strains carried by female Sandflies are related to the DEN2 patterns in patients. Although this does not prove the role of Sandflies in transmitting DEN2 in particular and flavivirus in general, however, through this study we also see that it is entirely possible for Sandflies to carry flaviviruses and especially DEN2. In the context that natural ecological zones are shrinking due to urbanization impacts, it is possible for viruses to also mutate to adapt to new conditions. In the future, we hope to have further studies on the disease transmission role of these virus strains.

4.2.2. Prevalence of Flavivirus infection in Sandflies

Another study by Gregory Moureau in 2010 also announced the discovery of Flavivirus RNA in Sandflies. In this study, 1508 Sandflies collected in France and Algeria, from August 2006 to July 2007, 2 pools of males out of 67 pools of this species *Phlebotomus perniciosus* in Algeria were positive for Flavivirus. The results of these two pools have sequences similar to Flavivirus, related to insect vectors of the genus Culex. This is the first description of insect-specific flaviviruses of the Culicidae family (including Aedes, Culex, Mansonia, Anopheles mosquito genera), found in Sandflies of the Psychodidae family. The difference between this study and ours is the discovery of Flaviviruses in male sandflies. Therefore, in the future, when screening for Flaviviruses on Sandflies, we will recommend conducting both on males and females to better understand the nature of Flavivirus viruses.

4.2.3. Some characteristics of Flaviviruses in female sandflies

Our detection of DEN2 traces on female sandflies of the genus Sergentomyia in Vietnam is not enough to confirm whether they transmit this virus or not. However, in the future, the genus Sergentomyia in general and the two species *Segentomyia (Parrotomyia) barraudi group* and *Segentomyia sp2* need to be studied further to understand their role in transmitting Flaviviruses, especially Dengue type 2.

4.3. Current status of Leishmaniasis in sandflies at the research site and the risk of transmission to humans

4.3.1. Leishmania identified in Quang Ninh

Visceral leishmaniasis in humans was first reported in 2000 in Quang Ninh province. These cases raise questions about local transmission of leishmaniasis in Vietnam. Samples collected from a patient in Quang Ninh were identified by the Queensland International Institute, Brisbane, Australia as *Leishmania infantum* or *L. donovani*.

In this study, we screened 1 sample of *Leishmania sp.* from a sandfly caught in a cave near the patient's home in 2000. Genetic sequencing results of the leishmania strain in the study determined that the sequence was 99.89% similar to *L. infantum* or *L. donovani* (Figure 3.9). This is completely similar to the results isolated from patients of the Queensland Inter-National Institute, Brisbane, Australia.

4.3.2. Leishmania identified in Son La

In the positive leishmania sample from Son La, we found 7 gene sequences similar to 8 classes (class 13, 16, 18, 25, 31, 38, 39 and 50) on the minicircle kinetoplas of the published *Leishmania infantum* strains. This shows the complexity of the genetic structure of this group of parasites. The circular genes on kinetoplasms are highly repetitive and multi-layered. Without using the NSG gene sequencing technique, it is very difficult to analyze and separate these repeats (Table 3.9).

4.3.3. Leishmania identified in Ninh Binh

Similar to the strain in Son La, the Leishmania sample we collected in Ninh Binh showed two sequences similar to the sequences of *Leishmania infantum* strains published on class 16 and 25 of the minicircle kinetoplas.

Although these data are not enough to conclude *Se. sp2* is the vector that transmits *L. infantum*, but this discovery confirms for the first time the presence of Leishmania in the sandfly population in Vietnam, and is a strain that causes a very dangerous visceral disease. Sandflies of the genus Sergentomyia have never been recorded to transmit Leishmaniasis, so knowledge about them is still limited. With the results of the project, we think the role of disease vectors of the Sergentomyia genus should be increased and should be included in surveillance to clarify the disease transmission role of the Sergentomyia genus in general and the *Se. sp2* in particular

Limitations of the study

In our research results, we discovered two new sandfly species, *Se. sp2* and *Se. sp3*. Currently, we are collecting more related information and collaborating with domestic and foreign sandfly experts to confirm.

Finding traces of Leishmania in female sandfly populations in Vietnam is a very new discovery. However, this still does not prove the role of Sandflies in transmitting the leishmania parasite in Vietnam. There is still a lot of evidence that needs to be supplemented, such as the ability to transmit and the multiplication of the virus in the body those Sandflies. In the future, when we can raise and clone sandfly species and isolate disease-causing Leishmania strains, we hope to provide more complete information about the disease transmission role of this vector in Vietnam.

Compared with other studies announcing the discovery of Flaviviruses in Sandflies, those studies have isolated virus strains, allowing a more complete sequence of virus gene segments to be obtained. In this study, we did not conduct isolation, so the results were limited to recording traces of flavivirus on collected female sandfly samples. On the other hand, some studies around the world show that male sandflies are also capable of being infected with Flavivirus. In this study, we have just performed screening on female sandflies. Therefore, with future research we will be able to conduct on male sandfly samples.

CONCLUDE

1. Species composition and some distribution characteristics of Sandflies in 6 northern provinces of Vietnam, 2016-2018

In 6 mountainous provinces in Northern Vietnam, 2585 Sandflies were collected, including 1511 males (58.5%) and 1049 females (40.6%), recording 5 sandfly breeds: Sergentomyia (n=2067, 79.96%) accounted for the highest proportion, Phlebotomus (n=340, 13.15%), Chinus (n=31, 1.2%), Idiophlebotomus (n=14, 0.54%) and Grassomyia (n=6, 0.23%).

A total of 15 sandfly species have been recorded, of which 13 sandfly species have been identified. Unidentified Sandflies have two species: *Sergentomyia sp2. and Sergentomyia sp3.*

Sandflies were recorded to be present in all habitats including: indoors (n=30), outdoors (n=936), bufalo cages (n=65), chicken cages (n=66), Pigs cages (n=41), dog cages (n=16) and caves (n=1431). Sandflies are mainly collected in caves. The highest species richness was in caves and outdoor habitats SR=15 (including *Se. sp2 and Se. sp3*).

2. Current status of flavivirus infection in Sandflies at the research site

Proportion of female sandflies carrying RNA of DEN2 virus in the study was 0.198% (n=2).

3. Current status of leishmaniasis in Sandflies at the research site

Leishmania is determined on female sandfly populations in three provinces: Son La, Quang Ninh and Ninh Binh. Leishmania infection rate infection in the studied female sandfly population was 3/1009 individuals (0.297%).

REQUEST

1. Increase surveys to collect Sandflies in other provinces to have more data to compare with the results of the study.
2. Continue and expand studies on morphology and molecular biology to accurately identify mosquito vectors that transmit diseases and further clarify the role of pathogens such as flaviviruses, especially DEN2 and leishmaniasis on Sandflies for a more comprehensive and comprehensive picture.
3. Expand research on the behavior and ecology of Sandflies in habitats near humans to have proactive prevention measures in the context of finding DEN2 and leishmania RNA in female Sandflies.