Study on First Establishments, Expansion Pathways, and Management of Coffee Leaf Rust Disease in Vietnam

January 2023

LE THI MAI CHAM

Study on First Establishments, Expansion Pathways, and Management of Coffee Leaf Rust Disease in Vietnam

A Dissertation Submitted to
the Graduate School of Science and Technology,
University of Tsukuba
in Partial Fulfillment of Requirements
for the Degree of Doctor of Philosophy in Agricultural Science

Doctoral Program in Agricultural Sciences
Degree Programs in Life and Earth Sciences

LE THI MAI CHAM

Table of Contents

List of Figures

List of Tables

Acronyms

AMOVA: Analysis of Molecular Variance
BLAST: Basic Local Alignment Search Tool
BP: Binh Phuoc
C: Chiapas
CB: Coimbra
C-C: Caldas-Chinchiná
CH: Central Highlands
CH-DN: Central Highlands-Dak Nong
CIFC: Centro de Investigação das Ferrugens do Cafeeiro
CLR: Coffee Leaf Rust
CM: Chiang Mai
CP: Capinopolis
CYTB: cytochrome b
DB: Dien Bien
DL: Dak Lak
DMIs: Sterol Demethylation Inhibitors
DNA: Deoxyribonucleic acid
ES: Espírito Santo
G: Guanine
HdT: Hybrido de Timor
LD: Lam Dong
ISSR: Inter Simple Sequence Repeats
ITS: Internal Transcribed Spacer
MCMC: Monte Carlo Markov Chain
MG: Minas Gerais
NCBI: National Center for Biotechnology Information
NJ: Neighbor-joining
NW: Northwest
O: Oaxaca
P: Puebla
PCoA: Principal Coordinate Analysis
PCR: Polymerase Chain Reaction
Q-B: Quindío-Buenavista
QoIs: Quinone Outside Inhibitors
SDHIs: Succinate Dehydrogenase Inhibitors
SE: Southeast
SE-DN: Southeast-Dong Nai
SEM: Scanning Electron Microscope
SL: Son La
SO: Senhora de Oliveira
SSP: São Sebastião do Paraíso
TL: Thailand
TSH-R: The Mycological Herbarium, the Faculty of Life and Environmental Sciences, the University of Tsukuba
USA: United States of America
V: Veracruz
VNM: Venda Nova do Imigrante

ABSTRACT

Leaf rust is a devastating fungal disease in coffee production worldwide. The epidemics of coffee leaf rust (CLR) were not random occurrences in the world but closely related to the historical cultivation of coffee, wind, and human activities in the past. Historically, large numbers of research related to CLR conducted globally indicate the paramount importance of this disease in coffee production. Even though Vietnam, the global second-biggest producer, has a long history of coffee production since 1857, and the rust was first detected in northern Vietnam in 1890; however, coffee leaf rust has only been indicated in several reports. The actual occurrence, and distribution of the disease as well as the causative pathogenic agent have not been investigated. On the other hand, studies on controlling the disease are still being omitted in Vietnam. Fungicides including Qo inhibitors (QoI) were applied in Vietnam and other countries to control this disease. QoIs resistance has occurred in other rust fungi recently reminds us of the necessity to investigate the fungicide-resistant matter in this fungus. Therefore, the purpose of this study is to investigate the current situation of rust, evaluate the genetic differentiation in the genome of pathogenic agents, and then, propose the appropriate approaches to control this disease in Vietnam.

Three regions were chosen for the survey, i.e., the Southeast, Central Highlands, and Northwest. Samples of CLR were collected, evaluated for disease severity and then identified pathogenic agents based on morphological features and nuclear ribosomal DNA internal transcribed spacer (rDNA-ITS) region. After that, based on genetic variation in the ITS region, possible original sources and subsequent expansion pathways of wind-borne and human-aided urediniospores of the CLR fungus were estimated. Furthermore, to investigate the mutation in *CYTB* gene, three polymerase chain reaction (PCR) primer pairs were designed to amplify exons 2 to 4 and applied to dried specimens from Vietnam.

Field surveys conducted at eighty-five plantations across three main coffee planting regions showed coffee leaf samples infected by the rust fungus in forty-one plantations from South to North, and from low altitudes to high mountainous areas. The field observation also indicated that all coffee species and cultivars planted in Vietnam now are susceptible to rust disease. Especially, *Coffea arabica cv. Catimor* varieties exhibited high levels of susceptibility with severe rust symptoms, while *Coffea canephora* varieties had varying degrees of susceptibility. Urediniospores of all isolates had tightly clustered spines on the borderline between echinulate and smooth areas and similarity values in ITS region from 96% to 100% with *Hemileia vastatrix*. Therefore, the pathogenic agent causing CLR disease in Vietnam was identified as *H. vastatrix*. Further analyses of ITS region from 83 samples (41 sequences from Vietnam, 2 from Thailand, and the remaining 40 from American countries) indicated that *H. vastatrix* populations from Vietnam and those from Central and South American countries (Mexico and Brazil) have very close relationships. In addition, there is no clear genetic structure in *H. vastatrix* population with respect to the geographic regions surveyed.The low genetic distance recorded between populations in Vietnam proves that the probability of migration between these regions was high. Moreover, most ancestral haplotypes are present in the Northwest, and the percentage of those ancestral haplotypes in the Northwest is larger than that in other regions. The phylogeographic tree of these isolates also showed that *H. vastatrix* in the Northwest had a close genetic relationship with those from other regions in Vietnam. Therefore, the Northwest was hypothesized as the first establishment of CLR in Vietnam. Besides, based on the haplotype network and neighborNet analysis, other branches of CLR in Vietnam were initiated in the Central Highlands. Analyses of the *CYTB* gene of *H. vastatrix* isolates from Vietnam showed no nucleotide mutations in exons 2 and 3. Meanwhile, two synonymous mutations were

detected in exon 4 in all isolates. Consequently, the application of QoIs to control CLR might be possible because the *H. vastatrix* populations in Vietnam still seem to be sensitive to QoIs.

In conclusion, *H. vastatrix* is the only pathogenic agent causing CLR in Vietnam. The current unstructured populations of *H. vastatrix* with high aggressiveness have been prevailing in all varieties and lines of *C. canephora* and Catimor derivatives in Vietnam. Moreover, some sources of this fungus were established, rapidly adapted to the local environment, and expanded to almost all regions with the aid of human activities. Some strategies to control the disease were proposed with the collaboration of Vietnam's government and local farmers. Especially, free distribution of uncontaminated seeds, seedlings of current/new resistant cultivars to smallholder farmers is strongly preferred. Moreover, plant diet, farmer education, and good cultural practices were also considered in these strategies. Likewise, when applying QoIs in the field, coffee farmers need to combine them with other fungicides or other methods to get rid of fungicide resistance in *H. vastatrix* population.

CHAPTER 1

INTRODUCTION OF COFFEE LEAF RUST DISEASE

1.1 Distribution of coffee in the world and Vietnam

There are over one hundred species listed in the genus *Coffea*, which can be found mostly in Africa and Madagascar, but also in South Asia, Southeast Asia, and Australasia (Davis, 2011; Davis et al., 2019). Among them, two popular commercial species are known as *C. arabica* (arabica coffee) and *C. canephora* (robusta coffee). Arabica coffee tends to grow in cool, dry, and mountainous areas while robusta coffee is distributed in the hot and humid tropical lowlands (McCook, 2006).

Traded coffee started from picked cherries from the wild arabica in southwestern Ethiopian forests and distributed through Yemen. The first coffee cultivation began in the Arabian Peninsula to meet market demand (Tuchscherer, 2003). Subsequently, to meet increased demand coffee in Europe, coffee plantations were opened in British colonies, South India and Ceylon (now Sri Lanka), in Dutch East Indies (now Indonesia), and in Spanish colony, the Philippines (McCook, 2006). Concurrent with the expansion of coffee plantations in South and Southeast Asia, coffee leaf rust disease spread, devastating the coffee cultivation there, as well-known in the epidemics in Ceylon (Sri Lanka) between 1865 and 1889. As a result of the rust epidemic, Dutch coffee planters in Java switched rust-susceptible arabica to robusta coffee, a highly resistant species. This coffee species has rapidly expanded to French and British colonies.

Coffee trees were introduced to the West Indies and northern South America as early as the 1720s, e.g., French colony Martinique and St. Domingo (Dominican Republic) and Dutch Guiana (Srinam). Subsequently, coffee plantations were expanded in Spanish and Portugese colonies in South America, particularly in Brazil, by the early 1800s. Brazil had become the world-largest coffee producer, producing more than 80% of the world’s coffee by the 1880s (McCook, 2017). The coffee had continued to be a profitable commodity crop in Latin America until coffee leaf rust disease was first discovered in Bahia, Brazil in 1970. National and international authorities had responded quickly to survey the disease spread and to develop effective management measure against the rust. The most notable effort was to breed and distribute new rust-resistant cultivars to prevent the rust epidemics. Since then, these hybrid coffee varieties are planted worldwide (McCook, 2006). Currently, *C. arabica* is distributed mostly in East and South Africa, Central and South America, and several areas belongs to Asian Islands. Meanwhile, *C. canephora* covers Asian Islands, Southeast Asia, West Africa and South America (Bunn et al., 2015). The coffee leaf rust had been a manageable disease, particularly in the higher altitudes, i.e., above 1000 to 1400 m asl., untile the pandemic occurred in Colombia in 2008-2011, Mexido and Central America in 2012-2013, and Peru and Ecuador in 2013.

French missionaries first introduced arabica coffee in northern Vietnam in 1857. Afterward, coffee was then expanded to the Northcentral Coast and reached to southern provinces (Central Highlands and Southeast). In 1908, robusta coffee and exelsa coffee (*C. exelsa*) were brought to Vietnam. Other coffee varieties were imported to Central Highlands from Congo. Current two main robusta coffee in Vietnam were derived from Java Island of Indonesia. Vietnam now cultures coffee in the Northwest, the Northcentral Coast, Central Highlands, and Southeast (Anonymous, 2019). After World War II, the introduction of *C. canephora* cultivars enabled Vietnam to become the world’s second-largest coffee producer, after Brazil.

1.2 Coffee leaf rust disease

1.2.1 Disease symptom

Coffee leaf rust is one of the devastating diseases that cause extreme yield loss for global coffee production. Therefore, the disease is included in major tropical plant diseases which were carefully studied and scientifically reported. Until now, CLR disease has appeared only in the *Coffea* family, with no records of CLR in other host plants (Rothfos, 1980; Talhinhas et al., 2017). One of the disease symptoms of CLR is immature defoliation caused by orange spore masses on the lower leaf surface (Talhinhas et al., 2017) and leading to the death of branches and then crop losses (Avelino et al., 2015). It is difficult to distinguish CLR from other leaf-infecting fungal diseases based on the first symptoms-light yellow blotches on the underside of leaves. The older blotches will be covered with a powdery coating of urediniospores. These coatings are firstly in light orange and finally turned to orange-red color. Fungal infection lesions grow larger and may reach each other to construct greater spots. The fungal infection gradually develops into the coffee branches and causes premature leaf drop (Rothfos, 1980).

1.2.2 Rust epidemics and their impacts

Coffee leaf rust first appeared as an epidemic in Ceylon (now Sri Lanka) and southern India in 1869 and caused a yield loss almost by one-third between 1870 and 1877 in Ceylon. By 1913, the Indian Ocean basin and the Pacific downed from almost 30% of contributing to the world’s coffee. Coffee leaf rust also caused severe damage to the economy of the Philippines when making dropping exporting coffee from 16 million pounds to virtually nothing from 1889 to 1892 (McCook, 2006). Applying fungicides preventively, fertilizers, and other cultural techniques, and even planting rust-resistant coffee varieties could control the disease at some degree (McCook, 2006; Avelino et al., 2015). However, outbreaks of CLR were still reported in some areas including Colombia (2008-2011), Mexico and Central America (2012–2013), and Peru and Ecuador (2013) (Avelino et al., 2015). Abnormal weather (temperature, rainfall, and sunshine duration) could be the reasons for the outbreaks of CLR in Central America and Colombia (Avelino et al., 2015). Nonetheless, other factors like wind, anthropogenic elements, and the emergence of new races, particularly in relation to the resistance break-down of the cultivar Lampira in Central America, have to be considered for the outburst of CLR.

1.2.3 Pathogenic agents: Taxonomy, life cycle, and infection process

The rust fungi belong to *Puccinials* (*Pucciniomycetes*, *Basidiomycota*). They are obligate parasites of vascular plants and cause rust diseases in a large number of important crops worldwide. Management of rust disease is difficult because the life cycle of most rust fungi are complicated with up to five different spore stages including basidiospore, spermatium, aeciospore, urediniospore, and teliospore, which may occur on the same host taxonomic group (autoecious life cycle) or on two taxonomically unrelated hosts (heteroecious life cycle) (Cummins & Hiratsuka, 2003; Kolmer, Ordonez, & Groth, 2009). Still others, like some species in the genus *Hemileia* (*Zaghouaniaceae* P. Syd. & Syd.) (Aime & McTaggart, 2020), are not known for their complete life cycle. *Hemileia* is characterized by its unique uredinial anamorph (Ritschel, 2005) and comprised of 42 species found mainly on the family *Rubiaceae* and *Apocynaceae* plants.

Hemileia vastatrix and *H. coffeicola* Maubl. & Roger have been known as causative agents of the coffee leaf rust disease to date (Thirumalachar & Narasimhan, 1947; Ritschel, 2005; Talhinhas et al., 2017). *Hemileia vastatrix* is characterized by the tightly clustered spines at the boundary between echinulate and smooth areas, while *H. coffeicola* lacks this feature (Ritschel, 2005). Moreover, *H. vastatrix* is the major agent causing yeild loss on coffee production worldwide. On the other hand, *H. coffeicola* occurs only in West and Central Africa on *C. canephora* (Waller, 1982). The global outbreaks of CLR previously have always been related to *H. vastatrix* (Talhinhas et al., 2017).

Hemileia vastatrix Berk. & Broome was described and named in 1869 as an obligate pathogenic fungus of the CLR (Ritschel, 2005; Talhinhas et al., 2017). The life cycle of *H. vastatrix* relies on its urediniospores (Ritschel, 2005). Teliospores of *H. vastatrix* were described by Ward (1882) and confirmed by Fernandes et al. (2009), while its spermogonial and aecial stages have not been known until now (Talhinhas et al., 2017).

The infection process starts with the germination of urediniospores. These spores need water (wet conditions) and optimal temperature (21–25ºC) to germinate. Together with temperatures below 15ºC and above 28 ºC, the strong light also inhibit spore germination. After penetrating coffee leaves through stomata, appressoria are produced and the infection is started (Waller, 1982) by forming penetration hypha. More intercellular hyphae will be generated to form mycelial networks after that. Uredinia are superstomata in rounded to irregular orange pustules after 3 weeks (Talhinhas et al., 2017). The sporulation usually starts 2 weeks to several months after infection depending on environmental conditions and the remaining leaves on the tree (Waller, 1982). When being mature, urediniospores can be dispersed by water drops (raindrops), wind (Waller, 1982), or human activities (McCook, 2006).

In Vietnam, CLR was indicated in some reports. However, the scientific information including pathogenic agent(s) and infection cycle, occurrence and outbreak, its impacts on livelihood of farmers was limited. Annually, the disease begins from March to April and from July to September in the northern part of Vietnam. By contrast, the rust disease symptoms become apparent in the wet season between April to September in the southern part.

1.2.4. The expansion of coffee leaf rust in the world

Coffee leaf rust epidemics were not random occurrences in the world but closely related to the historical cultivation of coffee, wind, and activities of humans in the past. This disease was first detected in Ethiopian wild coffees between 1500-1869 but did not cause much harm to the plants (McCook, 2006). Coffee cultivation began in Arabian Peninsula with seedlings from Africa. After that, Indian, Dutch, French, and British obtained coffee seeds or seedlings from this area to expand coffee planting in Ceylon, India, Africa, and Asia (McCook, 2006). Early 17th century, arabica coffee arrived in French and then spread to French colonies after that (Ferreira et al., 2019). Rust was first epidemic in Ceylon and southern India from 1869-1885 intimately related to spores transported from Ethiopia's source by monsoon winds and human activities. After that, from 1875 to 1920, CLR spread to the Indian Ocean basin and the Pacific with reasons connected to the movement of coffee planters and other routes including trade, communication, and wind. This pandemic made European coffee planters on Java Island import robusta coffee (*C. canephora*) with natural rust-resistant ability in 1901. At this time, this coffee cultivar divided the world's coffee into two areas with arabica coffee production in Latin America while Africa, Asia, and the Pacific areas planted robusta coffee. However, the rust endemic was still remaining in the second area between 1920 and 1950. After 1950, CLR spread to western Africa and burned in the Americas (New World's coffee area) between 1970 and 1985. Robusta and some new rust-resistant hybrid coffee varieties were introduced to replace traditional arabica varieties in Latin American countries to control CLR. Since then, these hybrid coffee varieties are planted worldwide (McCook, 2006, 2019). Coffee leaf rust totally changed the global distribution of coffee varieties in the world.

1.3 Coffee rust disease control approach in the world and Vietnam

1.3.1 Resistant coffee varieties

Hemileia vastatrix infects different pathogenicity in different varieties of coffee (Mayne, 1932). Twelve races of *H. vastatrix* were established based on Flor's gene-for-gene hypothesis about four dominant genes related to plant resistance and corresponding virulence factors in the pathogen (Noronha-Wagner and Bettencourt, 1967). Thereafter, 30 physiologic races were identified based on S_H factors in Coffea (D'Oliveira, 1971). *Coffea canephora* can resist naturally while *C. arabica* is susceptible *H. vastatrix*. Hybrido de Timor (HdT)– a natural hybrid between *C. arabica* and *C. canephora* was discovered in Timor in 1927. This hybrid showed a great resistant level against all races of *H. vastatrix* (Rodrigues et al., 1975). The ability of resistance in HdT plants is deliberate by genes originating from Robusta (S_H6–S_H9 and others) and Arabica (S_H1, S_H2, S_H4, and S_H5) (Noronha-Wagner and Bettencourt, 1967). After HdT discovery, breeding programs for the purpose of transferring resistant genes to main Arabica cultivars were conducted by a research and advanced training center of coffee (Centro de Investigação das Ferrugens do Cafeeiro - CIFC), Colombia, and Brazil. The hybrids Catimor and Sarchimor were from cultivars Caturra and HdT CIFC832/1, Villa Sarchi and HdT CIFC832/2, respectively. The hybrids between Caturra and HdT CIFC1343, and Catua_ı and HdT CIFC2570 together with Catiomor, Sarchimor are sources of the majority of current rust-resistant varieties (Talhinhas et al., 2017).

Currently, *C. canephora* is planted in the Central Highlands and Southeast, whereas *C. arabica cv. Catimor* covers most of northern Vietnam (Anonymous, 2019). All species were expected to be resistant to CLR to some degree. It can be concluded that Vietnam is currently planting varieties resistant to CLR.

1.3.2 Chemical and biological controls

Besides planting rust-resistant varieties, the chemical application also showed effective control. Preventative treatment is conducted using coper-base fungicides while curative treatment is using systemic fungicides including quinone outside inhibitors (QoIs), and sterol demethylation inhibitors (DMIs) (Talhinhas et al., 2017). QoI fungicides are applied alone or in mixtures with other systemic fungicides like triazoles (DMI group) (Souza et al., 2011; Honorato et al., 2015) or copper-based fungicides (copper oxychloride and hydroxide, cuprous oxide, and Bordeaux mixture) (Zambolim, 2016; Costa et al., 2019) to cure CLR and showed efficiently control as well as essentially increasing yield. Combining or alternating using different fungicides is an effective strategy to prevent fungicide resistance (Zambolim, 2016) which happened in other rusts like *Puccinia horiana*, the causative agent of chrysanthemum white rust, (Cook, 2001; Matsuura, 2019; Matsuzaki et al., 2020; 2021) and soybean rust *Phakopsora pachyrhizi* (Klosowski, 2016; Muller et al., 2021). In addition, potassium silicate (Lopes et al., 2013), essential oils (Pereira et al., 2012), benzothiadiazole (BTH) group (acibenzolar-S-methyl) (Fernandes et al., 2013), and plant extract were applied to control CLR (Talhinhas et al., 2017).

Entomopathogenic fungus *Lecanicillium lecanii*, a soilborne fungus, and bacteria are applied as biological approaches in controlling CLR. *Lecanicillium lecani* can parasite not only the urediniospores of *H. vastatrix* (Vandermeer et al., 2010) but also the green coffee scale *Coccus viridis*. Another study showed that endophytic bacteria including *Brevibacillus choshinensis*, *Salmonella enterica*, *Pectobacterium carotovorum*, *Cedecea davisae* significantly reduced rust disease on coffee leaf discs and seedlings experiments (Silva et al., 2012). Meanwhile, rust disease incidence and severity were reduced when spraying *Bacillus thuringiensis* isolate B157 and *Pseudomonas putida* isolate P286. Especially, the isolate B157 was a potential biocontrol agent for CLR because the effective control was equal to copper hydroxide (Haddad et al., 2009). Subsequently, the mechanism of CLR control of these bacteria was proved to produce antifungal compounds (Haddad et al., 2013).

In Vietnam, farmers have been applying copper-based fungicides to control CLR.

Moreover, many chemical formulations containing QoIs and DMIs are available on the market (Circular No.19/2021/TT-BNNPTNT dated 28/12/2021), therefore, QoIs and DMIs are used to cure coffee plants from CLR.

1.4 Objectives of the research

Vietnam has a long history of coffee cultivation since 1857 and is now ranked as the global second-biggest producer. In Brazil, the global biggest producer, and other countries in the Americas, research on CLR from outbreaks, and pathogenic agents, to approaches to control CLR have been conducted. However, studying CLR in Vietnam has not been given enough attention. Therefore, this study was conducted in order to evaluate the current situation of CLR, evaluate the genetic differentiation in the genome of the present population of pathogenic agent, and then propose the appropriate approaches to control this disease in Vietnam. Field survey and experiments were conducted to (i) investigate the occurrence of CLR in three main coffee production regions, (ii) identify pathogenic agents causing CLR in Vietnam, (iii) evaluate the current genetic diversity and population structure of the coffee rust fungus based on sequencing of the ribosomal DNA internal transcribed spacer (rDNA-ITS) region, (iv) hypothesize the geographic region where *H. vastatrix* first established in Vietnam, (v) establish appropriate PCR method with newly designed specific primers to amplify fragments of exon 2, 3, and 4 in the *CYTB* gene of *H. vastatix,* and (vi) examine F129L and G143A mutations in *CYTB* gene of *H. vastatrix* in Vietnam. These works were partitioned into 3 chapters with the first chapter is field investigation and identification of coffee leaf rust agents in Vietnam; second, evaluating the current genetic diversity and population structure of coffee leaf rust based on ribosomal DNA internal transcribed spacer (rDNA-ITS) sequences; and lastly, sequence analysis of *CYTB* gene in *H. vastatrix* in relation to potential QoI fungicide resistance.

CHAPTER 2

FIELD INVESTIGATION AND IDENTIFICATION OF COFFEE LEAF RUST AGENTS IN VIETNAM

2.1 Introduction

Coffee trees do not die from early defoliation by heavy rust infections (primary loss), but the fungal infection will cause a severe decrease in berry yield and quality in a subsequent year (secondary loss). The severe disease occurrence will cause rapid decline in the trees' vigor and make the trees more vulnerable to other pathogenic fungi and pests in successive years. Thus, yield losses by the coffee rust have been recorded up to 35%, and one to two billion US dollars have been lost annually (Talhinhas et al., 2017). It is, therefore, extremely important to frequently update the disease status in the field as well as identify the pathogenic agents to propose timely measures to control the CLR disease.

Hemileia vastatrix and *H. coffeicola* Maubl. & Roger have been known as causative agents of the coffee leaf rust disease to date (Thirumalachar & Narasimhan, 1947; Ritschel, 2005; Talhinhas et al., 2017). *Hemileia vastatrix* is the major agent causing yeild loss on coffee production worldwide. On the other hand, *H. coffeicola* occurs only in West and Central Africa on *C. canephora* (Waller, 1982). Previously, the global outbreaks of CLR have always been related to *H. vastatrix* (Talhinhas et al., 2017).

There are three main coffee-growing regions in Vietnam—the Southeast, the Central highlands, and the Northwest. Topographic and climatic differences affect seasonal disease occurrences and their severity in the three major coffee-growing regions. Coffee leaf rust has been reported periodically in Vietnam. However, the actual occurrence and distribution of the disease in Vietnam and the causative agents of the coffee leaf rust have not been investigated. Therefore, the objectives of this chapter were to (i) investigate the occurrence of CLR in three main coffee production regions and (ii) identify pathogenic agents causing CLR in Vietnam.

2.2 Materials and methods

Field investigation and sample collection

Coffee species were identified based on morphology (flower, bract, apex, leaf, fruit, etc.), according to the descriptions of Backer and Bakhuizen Van Den Brink (1965). Three coffee regions, including the Southeast (Dong Nai and Binh Phuoc provinces), the Central Highlands (Dak Lak, Dak Nong, and Lam Dong provinces), and the Northwest (Son La and Dien Bien provinces), were investigated for the occurrence of CLR. The accessed plantations and small-holder farms were along the street and were more than 1 km apart (Figure 2.1). Two rows of trees around the edges were examined for CLR fungi in each visited location. Ten coffee leaves with orange-colored rust sori were collected from three trees at each collection site, packed in a plastic bag, and referred to as one sample.

The CLR disease incidence was calculated based on the number of accessed plantations/small-holder farms infected by rust fungi in each region. Photographs of the rust-infected leaves were also taken, and the severity of the coffee leaf rust was estimated using the standard area diagram (Figure 2.2) of Belan et al. (2020). Based on the percentage of rust pustules covering the lower surface of an infected leaf, disease severity can be estimated from level 0 to level 7 (0%: level 0; 0.1–0.99%: level 1; 1.0–2.0%: level 2; 2.01–4.0%: level 3; 4.01–8.0%: level 4; 8.01–16.0%: level 5; 16.01–20.0%: level 6; 20.01–50.0%: level 7).

After sampling, the leaves were then dried by pressing them between dry blotting papers and maintained at room temperature (25 ± 2 ºC) at the microbiology laboratory of the Biotechnology Center of Ho Chi Minh City. Dried specimens were imported to Yokohama Plant Protection Station, Japan with Import Permit Number 1Y904r and then deposited in the

Mycological Herbarium of the University of Tsukuba (TSH). All samples were collected between July and November in 2019 (Table 2.1).

Morphological observation

Morphological observation of the coffee leaf rust fungus was conducted to determine the taxonomic identity. The urediniospores were scraped from sori on infected leaves using a sterile surgical knife, then mounted in a drop of lactophenol aniline blue solution on a microscope slide and covered by a glass cover slip. The slides were then observed under an Olympus BX51 light microscope (Olympus, Tokyo, Japan), and length, width, and wall thickness of the spores ($n = 20$/specimen) were measured by ImageJ software (Collins, 2007), following the measuring method for dorsoventrally asymmetric (flattened) urediniospores described by Ritschel (2005). The width was measured at the largest part in the lateral view where the smooth and echinulate side of the urediniospore was seeable. Meanwhile, the length of urediniospore was consistent with the greatest value of the measure. The surface structure of urediniospores was examined by a Hitachi S-4200 scanning electron microscope (SEM) (Hitachi, Tokyo, Japan) after coated with platinum-palladium to 25 nm thicknesses with a Hitachi E-1030 ion sputter (Hitachi, Tokyo, Japan).

DNA extraction, ITS region amplification, and sequencing

Deoxyribonucleic acid (DNA) extraction from herbarium specimens requires special methods, in which only a small amount of urediniospores must be picked to avoid possible cross-contamination. The thermal-shock method given by Fraczek et al. (2019) was used for DNA extraction, and BPS buffer was substituted with buffer-1 (10 mM Tris-Cl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl). The protocol can be summarized as follows: A coffee leaf with fully matured, seemingly clean sori on the lower surface was chosen for DNA extraction from each sample (one sample consisted of ten leaves from three different trees). Urediniospores were collected from several pustules on a single leaf and immersed in a 200 µl tube containing 30 µl buffer-1. The tubes were then vortexed and incubated at 95 ºC for 15 min and immediately transferred to a deep freezer (-80 ºC) for 10 min. Subsequently, the spore suspension was defrosted at room temperature (25 ± 2 ºC) and centrifugated at 15,000 rpm for 1 min. Finally, the supernatant was transferred to a new 200 µl tube. The nucleic acid concentration and 260/280 and 260/230 ratios were estimated using a DS-11 spectrophotometer (DeNovix, Delaware, USA).

The ITS region was amplified using ITS5 and ITS4 primers (White et al., 1990). A 25 µl reaction mixture containing 1–2 µl DNA template (30 – 50 ng/µl), 0.2 mM of each primer (2.5 µl), 12.5 µl Gene RED PCR Mix Plus (Nippon Gene, Tokyo, Japan), and 5.5–6.5 µl autoclaved distilled water, were prepared. Polymerase chain reaction (PCR) was conducted using a TaKaRa PCR Thermal Cycler Dice® Touch (TaKaRa, Tokyo, Japan). The PCR process was as follows: initial denaturation at 95 ºC for 5 min, followed by thirty-five cycles of denaturation at 95 ºC for 1 min, annealing at 54 ºC for 30 s, and extension at 72 ºC for 1 min, followed by a final extension at 72 ºC for 10 min.

The amplification was confirmed using electrophoresis. The amplified products were purified using FastGene™ Gel/PCR extraction kit (Nippon Genetics, Tokyo, Japan) and consigned to Eurofins Genomics (Tokyo, Japan) for sequencing. The obtained sequences were edited and assembled using ATGC software (Genetyx Co., Tokyo, Japan). Finally, the sequences were aligned and compared with other sequences in the National Center for Biotechnology Information database using the BLASTn program (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.3 Results

Field investigation

Field investigations showed that *C. canephora* is commonly planted in southern Vietnam, while *C. arabica cv. Catimor* is planted in the Northwest. In the Southeast of Vietnam, thirty-three plantations in Dong Nai and Binh Phuoc provinces were investigated, of which 26.7% (4/15) plantations in Binh Phuoc and 38.9% (7/18) plantations in Dong Nai were infected by the rust fungus. Meanwhile, thirty-five plantations in the Central Highlands were investigated. The incidence of coffee rust disease in the three provinces was variable, i.e., 41.7% (5/12) in Dak Lak, 77.8% (7/9) in Dak Nong, and 50.0% (7/14) in Lam Dong province. Our investigation at 17 plantations in the mountainous area of North Vietnam (Northwest) revealed that coffee rust disease incidence in this region was high—63.6% (7/11) and 66.7% (4/6) in Son La and Dien Bien provinces, respectively (Figure 2.3). After thorough field investigations at eighty-five coffee plantations in Vietnam, a total of forty-one CLR samples were collected from forty-one plantations in the three regions (Table 2.1).

The severity of coffee leaf rust on *C. arabica* in the Northwest and *C. canephora* in the Central Highlands and Southeast of Vietnam was assessed using individual survey sites and the method described by Balen et al. (2020) (Figure 2.2 and Table 2.1). Overall, the severity of the disease in Vietnam ranged from 1–7. In which, the Southeast had the lowest severity level, while the highest value of disease severity belongs to the Northwest (Table 2.1). In Southeast Vietnam, farmers usually intercrop coffee with black pepper. In addition, most of accessed locations in this region were managed by smallholder farmers. In seven CLR-infected smallholder farms in Dong Nai province (Southeast), the disease severity was highest in farm numbers S1-14 and S1-18 (levels 5–6). Next, farm numbers S1-9, S1-11, and S1-15 had moderate disease severity with disease levels ranging from 2 to 4. Lastly, smallholder farm S1-5 had the lowest value of disease severity (level 1–2) with only two or three rust pustules scattered on the lower surface of coffee leaves (Appendix 1). Four smallholder farms in Binh Phuoc province (Southeast) suffered from CLR disease with modest severity (Table 2.1) with only several large rust pustules were sprinkled on the lower surface of coffee leaves.

Central Highlands is the main coffee-producing area in southern Vietnam. In three surveyed provinces, Dak Lak province had a moderate value of disease severity with levels from 3 to 5. The symptoms of CLR disease were clear and there was no significant difference in disease severity between survey sites (Table 2.1). In the meantime, provinces Dak Nong and Lam Dong displayed variations in the rust severity between surveyed sites. Indeed, plantations H2-4 (Dak Nong) and H3-3 (Lam Dong) showed the lowest disease severity (level 1). In these two plantations, the former was carefully managed with an adequate fertilizer regime, while the latter is located in a high mountainous area. Plantations H2-7 and H2-9 (Dak Nong) together with H3-8 and H3-14 (Lam Dong) exhibited high disease levels. Most of them had been abandoned prior to our sampling. The remaining plantations/farms had moderate disease severity (Table 2.1).

The Northwest exhibited the highest disease incidence and severity compared with the remaining regions. In Northwest, *C. arabica cv. Catimor*, a rust-resistant variety of coffee, accounted for 95%. Figure 2.3 displayed severe rust lesions on the lower surface of coffee leaves. The disease severity ranged from 4 to 7 in all plantations surveyed (Figure 2.4 and Table 2.1).

Coffee plantations in the Northwest had the highest rust severity (up to level 7). While in southern Vietnam the disease appeared less severe than that in northern Vietnam. The severity of rust disease in Southeast Vietnam was found to be lower than in the other areas. Arabica varieties at all sites in the Northwest exhibited high susceptibility as the number and size of the rust lesions were greater when compared with other regions, while robusta varieties in the Central Highlands and the Southeast exhibited varying degrees of susceptibility (Figure 2.4 and Appendix 1). It was noticeable that those in Dak Nong and Dong Nai were highly variable among survey sites, ranging from highly resistant to highly susceptible (Table 2.1).

Morphological and molecular characteristic of the coffee leaf rust fungus in Vietnam

The telial stage of pathogenic agents was not detected in all specimens collected in Vietnam. The uredinia were suprastomatal on the abaxial surface of coffee leaves. Urediniospores were reniform in shape, 21–45 µm in length, and 14–31 µm in width. The wall was 0.8–2.6 µm thick (Table 2.2), smooth ventrally and coarse echinulate dorsally. The urediniospores were characteristically ornamented with a rim of thorough bundle of spines on the borderline between echinulate and smooth areas (Figure 2.5 and Appendix 2). All DNA sequences of ITS regions ranged from 850 bp to 886 bp and had similarity values from 98% to 100% with *H. vastatrix,* except sample TSH-R30071 and TSH-R30045 with 96.34% and 97.72% similarity with *H. vastatrix,* respectively (Table 2.3).

2.4 Discussion and conclusion

The size of urediniospores of *H. vastatrix* and *H. coffeicola* are equivalent. However, *H. vastatrix* has the backbone of tightly clustered spines on the borderline between echinulate and smooth areas while *H. coffeicola* lacks this feature (Ritschel, 2005). Table 2.2, Figure 2.5, and appendix 2 exhibited that all isolates in Vietnam are *H. vastatrix* because of owning all mentioned characters. Table 2.3 also supports morphological identification when all isolates had similarity values from 96% to 100% with *H. vastatrix*. This result is in agreement with a previous study when *H. coffeicola* could be found only in wet areas in West and Central Africa (Waller, 1982).

The results of field investigations show that the leaf rust disease is present in all coffee growing regions in Vietnam, from South to North, and from low altitudes to high mountainous areas. These observations suggest that all coffee species and cultivars planted in Vietnam now are susceptible to rust disease. Previously, *Coffea arabica* cv. Catimor and HdT were known to be rust-resistant (Talhinhas et al., 2017). However, our survey shows that this cultivar can be infected by rust fungus. The number and size of the rust lesions was large on the coffee tree leaves in the Northwest plantations, where Catimor varieties are predominately planted (Table 2.1, Figure 2.4, and Appendix 1). The detection of rust-infected leaves in regions planting rust-resistant coffee varieties is not unique to this research; numerous previous researches have revealed that the rust disease existed in HdT growing areas of Brazil and Colombia (Maia et al., 2013; Cabral et al., 2016; Talhinhas et al., 2017). This raises concerns about whether all rust-resistant coffee varieties have become susceptible to *H. vastatrix* or whether the situation is only present in a few countries or areas. The most probable reasons for the loss of rust resistance are mutations in the genome of *H. vastatrix* or the appearance of new pathogen races (Talhinhas et al., 2017). However, other factors that could affect the rust resistance of these coffee varieties must also be studied because climatic conditions and cultural practices can also promote disease outbreaks (Avelino et al., 2015). For example, there were high levels of rust severity found in the abandoned coffee plantations of the Central Highlands.

In conclusion, the pathogenic agent causing CLR disease in Vietnam is *H. vastatrix*. This fungus infected not only *C. canephora* but also hybrid coffee *C. arabica cv. Catimor.* Therefore, further research on factors including climatic conditions and culture practices is required to control CLR disease in Vietnam.

- Parts of this chapter have already been published in the following paper: Le CTM, Okane I, Ono Y, Tsuda Y, Yamaoka Y. Incidence of Coffee Leaf Rust in Vietnam, Possible Original Sources and Subsequent Pathways of Migration. Front Plant Sci. 2022 Apr 5;13:872877. doi: 10.3389/fpls.2022.872877. PMID: 35449887; PMCID: PMC9016365. https://doi.org/10.3389/fpls.2022.872877.

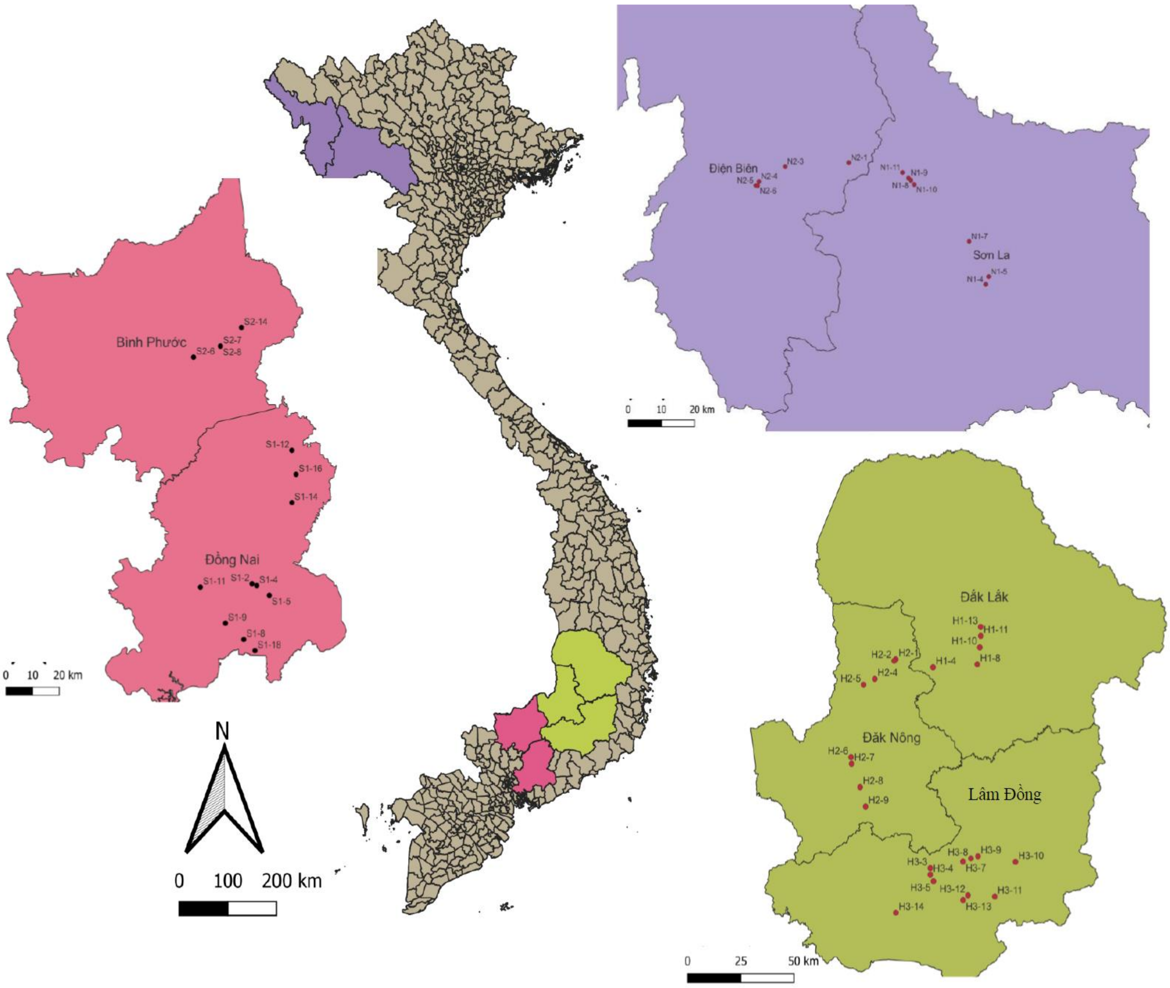


Figure 2.1. Geographic distribution of plantations where coffee leaf rust samples were collected in Vietnam. Different colors correspond to distinctive main coffee planting regions: purple color corresponds to the Northwest region, rubine red corresponds to the Southeast region, and olive green corresponds to Central Highlands. Red and black dots represent the location of plantations in each region.

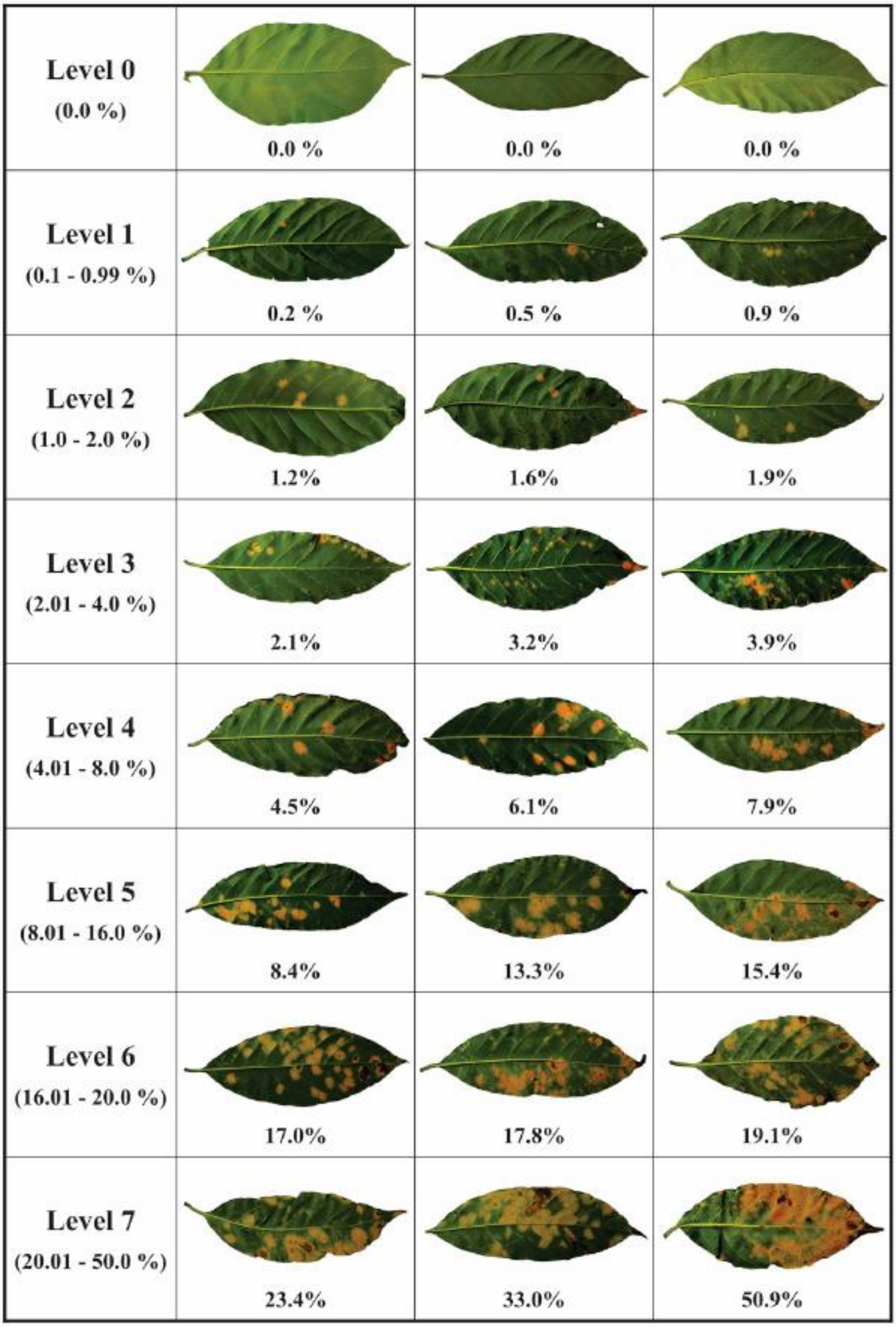


Figure 2.2. A standard area diagram was constructed based on the method of Belan et al. (2020) to estimate the severity of *Hemileia vastatrix* in *Coffea canephora*. The numbers express the percentage of infected leaf areas holding urediniospores of this fungus.

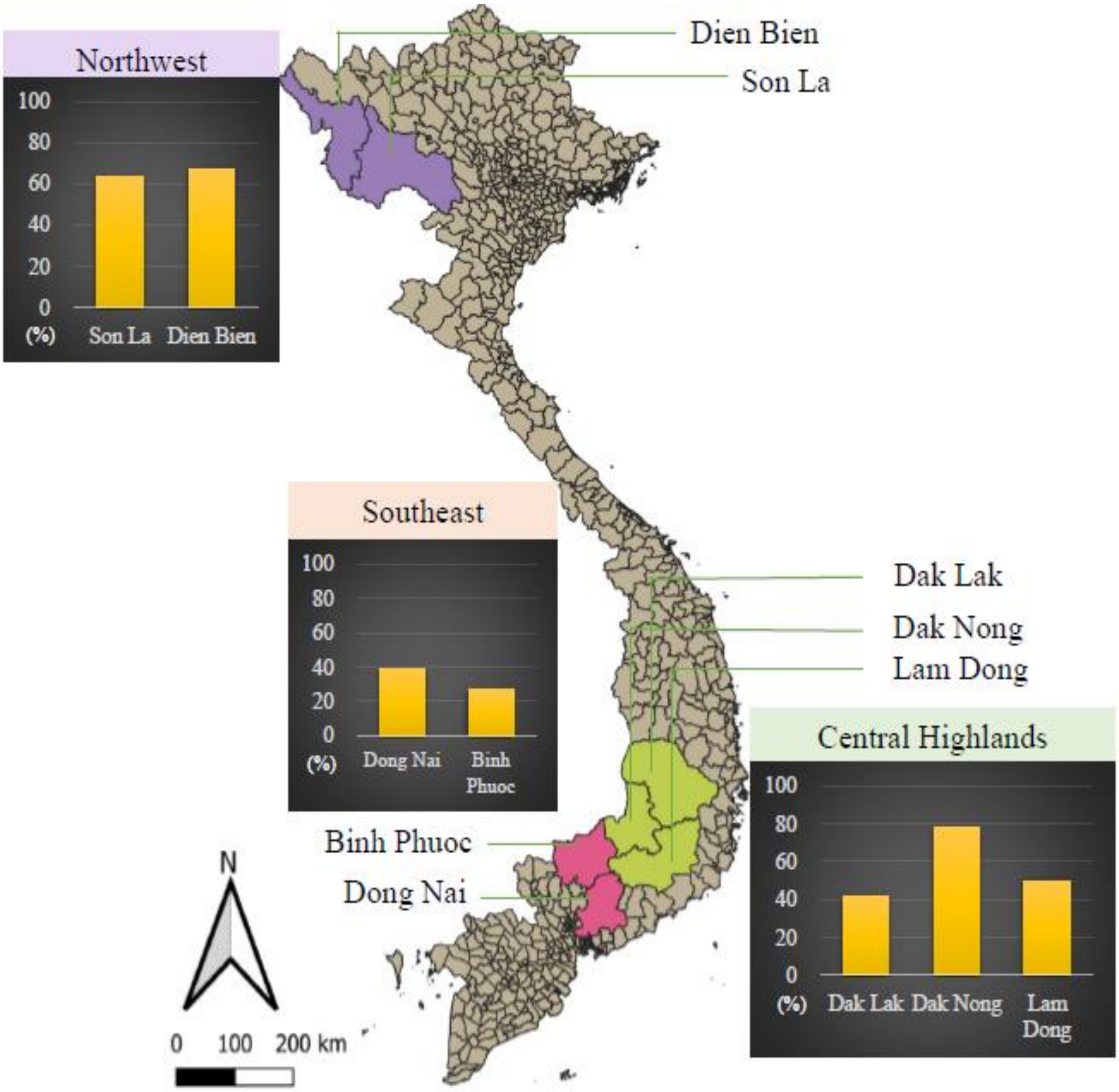


Figure 2.3. Rust disease incidence in different geographical sample regions of Vietnam. Colored areas on the map correspond to sampling regions: purple color indicates the Northwest, greenish-yellow signifies the Central Highlands, and reddish pink shows the Southeast.

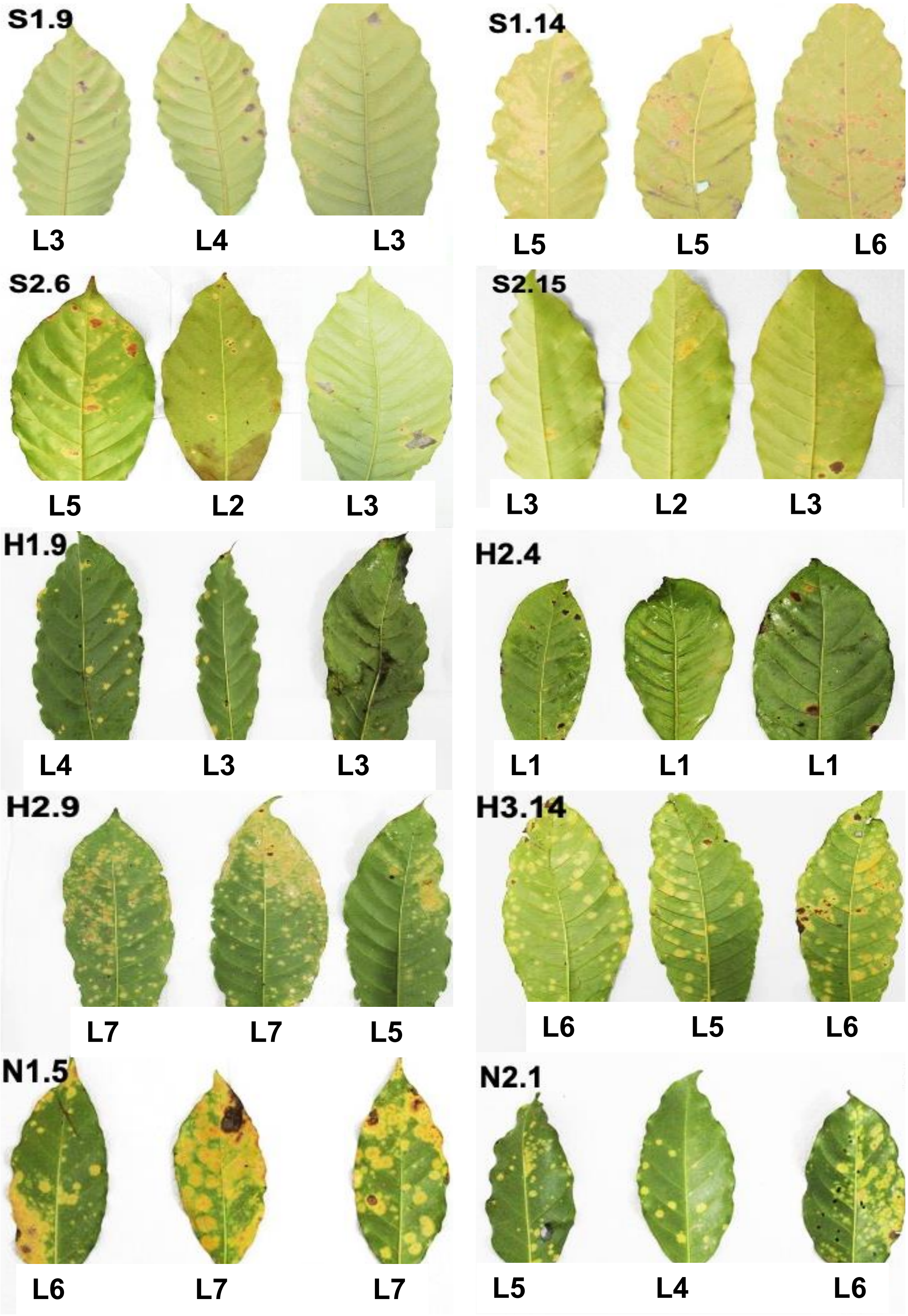


Figure 2.4. Coffee leaf samples infected by rust fungus with different severity level collected from three main coffee planting regions in Vietnam. S1: Southeast-Dong Nai, S2: Southeast-Binh Phuoc, H1: Central Highlands-Dak Lak, H2: Central Highlands-Dak Nong, H3: Central Highlands-Lam Dong, N1: Northwest-Son La, N2: Northwest-Dien Bien, and L: Level of disease severity.

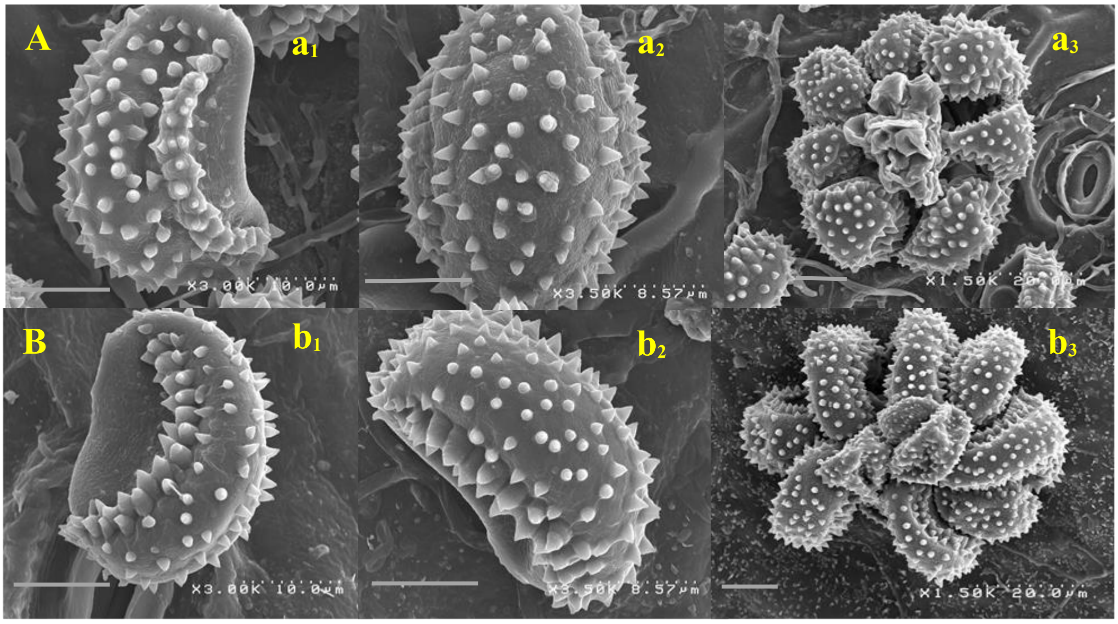


Figure 2.5. *H. vastatrix* on *C. canephora* (A- specimen TSH-R30055 from the Central Highlands) and *C. arabica cv. Catimor* (B- specimen TSH-R30155 from the Northwest): (a1, b1) urediniospore characterized with gathered spines on the borderline between echinulate and smooth areas, (a2, b2) upper surface structure of an urediniospore and (a3, b3) a suprastomatal uredinium of *H. vastatrix* collected in Vietnam. *Bar= 10 µm*.

Table 2.1. Coffee leaf rust samples collected in Vietnam

Area	Province	Date of collection	Latitude	Longitude	Specimen ID	TSH No. *	Severity % **			Level of severity **
							Sample 1	Sample 2	Sample 3	
The Southeast	Dong Nai	12-Jul-19	10.98369	107.2536	S1_2	TSH-R30018	15.4	15.4	13.3	5
			10.94672	107.3131	S1_5	TSH-R30022	1.2	0.5	0.9	1–2
			10.85789	107.1626	S1_9	TSH-R30024	2.1	6.1	3.3	3–4
			10.97286	107.0765	S1_11	TSH-R30028	3.2	2.1	1.9	2–3
			11.24236	107.3906	S1_14	TSH-R30031	15.4	15.4	17	5–6
			11.26664	107.3794	S1_15	TSH-R30034	1.2	2.1	1.2	2–3
			10.77028	107.2639	S1_18	TSH-R30036	17	8.4	13.3	5–6
	Binh Phuoc	15-Jun-19	11.70625	107.053	S2_6	TSH-R30042	8.4	1.2	2.1	2–5
			11.74158	107.1456	S2_7	TSH-R30043	8.4	4.5	2.1	3–5
			11.80078	107.2175	S2_14	TSH-R30045	8.4	15.4	7.9	4–5
			11.79633	107.2173	S2_15	TSH-R30046	2.1	1.2	2.1	2–3
The Central Highlands	Dak Lak	7-Aug-19	12.55333	108.181	H1_8	TSH-R30050	13.3	8.4	7.9	4–5
			12.55117	108.1911	H1_9	TSH-R30051	6.1	2.1	2.1	3–4
			12.62142	108.192	H1_10	TSH-R30052	7.9	13.3	13.3	4–5
			12.66742	108.196	H1-11	TSH-R30055	8.4	7.9	4.5	4–5
			12.7035	108.1957	H1_12	TSH-R30059	6.1	7.9	7.9	4
	Dak Nong	6-Aug-19	12.56772	107.8235	H2_2	TSH-R30063	13.3	15.4	8.4	5
			12.49522	107.7401	H2_4	TSH-R30067	0.9	0.5	0.5	1
			12.47192	107.6916	H2_5	TSH-R30069	7.9	3.2	3.2	3–4

* TSH: the Mycological herbarium of the University of Tsukuba, Japan, ** Belan et al 2020

(Continued)

Area	Province	Date of collection	Latitude	Longitude	Specimen ID	TSH No. *	Severity % **			Level of severity **
							Sample 1	Sample 2	Sample 3	
The Central Highlands	Dak Nong	6-Aug-19	12.18189	107.6379	H2_6	TSH-R30071	8.4	6.1	13.3	4–5
			12.18189	107.6379	H2_6	TSH-R30071	8.4	6.1	13.3	4–5
			12.15553	107.6402	H2_7	TSH-R30073	33	50.9	50.9	7
			12.06183	107.6771	H2_8	TSH-R30080	8.4	8.4	8.4	5
			11.98397	107.7004	H2_9	TSH-R30085	33	33	13.3	5–7
	Lam Dong	8-Aug-19	11.73731	107.9789	H3_3	TSH-R30088	0.2	0.5	0.5	1
			11.77639	108.1537	H3_8	TSH-R30098	17	17	17.8	6
			11.78475	108.1842	H3_9	TSH-R30101	6.1	7.9	17	4–6
			11.76256	108.3454	H3_10	TSH-R30104	8.4	8.4	13.3	5
			11.62358	108.2564	H3_11	TSH-R30107	6.1	17	8.4	4–6
			11.62833	108.1409	H3_12	TSH-R30109	13.3	7.9	7.9	4–5
			11.55897	107.8309	H3_14	TSH-R30118	17	15.4	19.1	5–6
The Northwest	Son La	17-Aug-19	21.23894	103.8599	N1_4	TSH-R30120	13.3	50.9	15.4	5–7
			21.25931	103.8687	N1_5	TSH-R30125	19.1	50.9	33	6–7
			21.35375	103.8124	N1_7	TSH-R30129	23.4	50.9	8.4	5–7
			21.51772	103.6437	N1_8	TSH-R30133	15.4	19.1	15.4	5–6
			21.52444	103.6375	N1_9	TSH-R30137	19.1	33	19.1	6–7
			21.50633	103.6534	N1_10	TSH-R30145	13.3	4.5	23.4	4–7
			21.53867	103.6202	N1_11	TSH-R30148	7.9	19.1	33	4–7

* TSH: the Mycological herbarium of the University of Tsukuba, Japan, ** Belan et al 2020

(Continued)

Area	Province	Date of collection	Latitude	Longitude	Specimen ID	TSH No. *	Severity % **			Level of severity **
							Sample 1	Sample 2	Sample 3	
The Northwest	Dien Bien	18-Aug-19	21.56531	103.4651	N2_1	TSH-R30155	13.3	6.1	19.1	4–6
			21.51497	103.206	N2_4	TSH-R30160	17	8.4	13.3	5–6
			21.50392	103.1969	N2_5	TSH-R30167	13.3	6.1	33	4–7
			21.50336	103.2022	N2_6	TSH-R30172	23.4	4.5	13.3	4–7

* TSH: the Mycological herbarium of the University of Tsukuba, Japan, ** Belan et al 2020

Table 2.2. The length, width, and wall thickness of urediniospores of the coffee leaf rust fungus collected in Vietnam

Sample ID.	Length (μm)	Width (μm)	Wall Thickness (μm)
TSH-R30018	28–36	17–29	0.9–1.8
TSH-R30022	26–36	18–21	1.0–2.6
TSH-R30024	24–33	16–25	0.9–1.5
TSH-R30028	34– 38	19–27	1.0–2.2
TSH-R30031	36–39	22–25	1.1–2.1
TSH-R30034	20–34	17–28	0.9–1.6
TSH-R30036	30–47	21–29	1.1–1.9
TSH-R30042	29–41	18–26	1.2–2.5
TSH-R30043	29–31	17–21	1.5–2.2
TSH-R30045	34–40	22–29	0.9–2.0
TSH-R30046	37–41	27–29	1.2–2.1
TSH-R30050	29–38	17–26	1.3–1.9
TSH-R30051	27–38	20–26	0.9–2.0
TSH-R30052	31–45	25–30	1.1–1.9
TSH-R30055	32–37	19–24	1.1–2.1
TSH-R30059	21–37	17–26	1.2–2.5
TSH-R30063	35–41	27–30	1.0–2.2
TSH-R30067	29–39	23–27	0.9–1.8
TSH-R30069	25–36	14–27	0.9–2.0
TSH-R30071	21–36	20–25	1.2–2.1
TSH-R30073	29–35	23–29	1.1–1.9
TSH-R30080	27–39	17–28	1.5–2.3
TSH-R30085	29–38	20–28	1.2–1.9
TSH-R30088	24–32	15–23	0.9–1.8
TSH-R30098	29–33	19–24	1.3–2.1
TSH-R30101	30–37	21–28	1.2–2.0
TSH-R30104	28–32	20–25	0.8–2.2
TSH-R30107	29–39	24–31	1.2–1.6
TSH-R30109	28–38	19–28	1.2–1.8
TSH-R30118	31–36	21–27	1.1–2.0
TSH-R30120	26–39	20–24	0.9–1.5
TSH-R30125	30–38	19–30	1.2–2.2
TSH-R30129	27–32	16–24	1.2–1.5
TSH-R30133	33–39	22–31	1.3–1.4
TSH-R30137	32–41	23–29	1.2–2.1
TSH-R30145	24–36	15–21	1.1–2.0
TSH-R30148	32–42	22–30	1.5–2.1
TSH-R30155	21–40	16–24	1.1–1.9
TSH-R30160	30–36	21–25	1.1–1.4
TSH-R30167	30–38	19–24	0.8–1.8
TSH-R30172	31–45	21–30	1.2–1.4

Table 2.3. Results of BLAST search on NCBI for Vietnamese coffee leaf rust fungus sequences

Sample ID.	**Results**	**Query cover**	**Identity (%)**	**Accession No.**
TSH-R30018	*Hemileia vastatrix* voucher CMPH-14	100	99.65	KX260246
TSH-R30022	*H. vastatrix* voucher CMPH-14	100	99.88	KX260246
TSH-R30024	*H. vastatrix* voucher CMPH-95	100	99.19	KX260248
TSH-R30028	*H. vastatrix* voucher CMPH-234	100	99.77	KX260253
TSH-R30031	*H. vastatrix* isolate CT_28_v	100	99.31	EF394131
TSH-R30034	*H. vastatrix* voucher CMPH-95	100	98.73	KX260248
TSH-R30036	*H. vastatrix* voucher CMPH-95	100	99.65	KX260248
TSH-R30042	*H. vastatrix* isolate Coimbra_5	100	99.31	KX260248
TSH-R30043	*H. vastatrix* isolate CTInNa3	100	99.55	EF394118
TSH-R30045	*H. vastatrix* voucher CMPH-234	100	97.72	KX260253
TSH-R30046	*H. vastatrix* voucher CMPH-14	100	99.41	KX260246
TSH-R30050	*H. vastatrix* voucher CMPH-95	100	99.77	KX260248
TSH-R30051	*H. vastatrix* isolate CTInNa2	100	99.32	EF394129
TSH-R30052	*H. vastatrix* voucher CMPH-234	100	99.88	KX260253
TSH-R30055	*H. vastatrix voucher CMPH-165*	100	99.55	KX260250
TSH-R30059	*H. vastatrix* voucher CMPH-95	100	99.55	KX260248
TSH-R30063	*H. vastatrix* voucher CMPH-95	100	99.32	KX260248
TSH-R30067	*H. vastatrix* voucher CMPH-95	100	98.30	KX260248
TSH-R30069	*H. vastatrix* voucher CMPH-234	100	99.43	KX260253
TSH-R30071	*H. vastatrix* voucher CMPH-95	100	96.34	KX260248
TSH-R30073	*H. vastatrix* voucher CMPH-14	100	99.88	KX260246
TSH-R30080	*H. vastatrix* voucher CMPH-142	100	98.96	KX260249
TSH-R30085	*H. vastatrix* voucher CMPH-142	100	98.73	KX260249
TSH-R30088	*H. vastatrix* voucher CMPH-142	100	98.96	KX260249
TSH-R30098	*H. vastatrix* voucher CMPH-14	100	99.54	KX260246
TSH-R30101	*H. vastatrix* voucher CMPH-95	100	98.50	KX260248
TSH-R30104	*H. vastatrix* voucher CMPH-14	100	99.31	KX260246
TSH-R30107	*H. vastatrix* voucher CMPH-14	100	99.88	KX260246
TSH-R30109	*H. vastatrix* voucher CMPH-95	100	98.85	KX260248
TSH-R30118	*H. vastatrix* voucher CMPH-95	100	98.27	KX260248
TSH-R30120	*H. vastatrix* voucher CMPH-14	100	99.65	KX260246
TSH-R30125	*H. vastatrix* voucher CMPH-14	100	99.77	KX260246
TSH-R30129	*H. vastatrix* voucher CMPH-14	100	99.88	KX260246
TSH-R30133	*H. vastatrix* voucher CMPH-14	100	99.77	KX260246
TSH-R30137	*H. vastatrix* voucher CMPH-95	100	99.77	KX260248

(Continued)

Sample ID.	**Results**	**Query cover**	**Identity (%)**	**Accession No.**
TSH-R30145	*H. vastatrix* voucher CMPH-95	100	99.54	KX260248
TSH-R30148	*H. vastatrix* voucher CMPH-95	100	99.77	KX260248
TSH-R30155	*H. vastatrix* voucher CMPH-14	100	99.42	KX260246
TSH-R30160	*H. vastatrix* voucher CMPH-14	100	99.88	KX260246
TSH-R30167	*H. vastatrix* voucher CMPH-14	100	99.88	KX260246
TSH-R30172	*H. vastatrix* voucher CMPH-14	100	100.00	KX260246

CHAPTER 3

EVALUATING THE CURRENT GENETIC DIVERSITY AND POPULATION STRUCTURE OF COFFEE LEAF RUST IN VIETNAM BASED ON RIBOSOMAL DNA INTERNAL TRANSCRIBED SPACER (rDNA-ITS) SEQUENCES

3.1 Introduction

In disease management, the best strategy to control pathogenic agents is by using resistant varieties (Silva et al., 2006). Hibrido de Timor (HdT) is a natural hybrid between *Coffea arabica* and *C. canephora*, which was resistant to all races of the coffee rust fungus, *H. vastatrix*, in the 1950s. The hybrids Catimor and Sachimor have also been considered resistant to *H. vastatrix* (Talhinhas et al., 2017). Planting disease-resistant coffee varieties was considered an effective strategy for preventing or decreasing the infection of *H. vastatrix* for a long time. However, in recent years, there have been reports of the outbreak of CLR in most coffee-growing areas in northern South America, Mexico, and Centeral America, including regions planting rust-resistant varieties (Maia et al., 2013; Cabral et al., 2016; Talhinhas et al., 2017). Thus, to develop effective and practical management programs, we must know the current pathogenic diversity of the rust populations, understand the underlying genetic variations, and predict the tendency for variations in future.

The native range of *H. vastatrix* is suggested to be a geographic range spanning the Great Lake region, Ethiopia, and the eastern half of the Congo River basin in East Africa (McCook 2019). After the domestication of *C. arabica* in Yemen in the fifteenth century, the fungus and its cultivated host had spread through Ceylon and southern India in 1869. The reasons for the first explosion of CLR were monsoon, shipment of infected coffee seedlings or goods, and the attached rust urediniospores on the clothes of British soldiers returning from Ethiopia. Thereafter, the CLR pandemic exploded in the Indian Ocean basin and the Pacific from 1875−1920. Besides other routes related to wind, trade, and communication, the migration of coffee planters from infected areas to the Indian Ocean basin and the Pacific could also be considered the causal factor. Wind, infected materials, and anthropogenic activities also explain the CLR outburst in West Africa during 1950−1970 and in America during 1970−1989 (McCook, 2006; McCook, 2019).

Coffea arabica was first cultivated in northern Vietnam in 1857 by French missionaries (D’haeze et al., 2004; Anonymous, 2019). Some decades after that, in October 1890, *H. vastatrix* was first collected on *C. arabica* in some plantations (Hennings, 1895). Several years later, *C. arabica* spread to Southern Vietnam. However, when diseases, including coffee rust, caused yield losses in the coffee plantations, *C. canephora* was introduced to replace *C. arabica*, and it rapidly spread in the southern parts of Vietnam in the early 1900s (D’haeze et al., 2003). In the meantime, the Northwest of Vietnam also switched to Catimor derivatives due to its ability to resist CLR (Anonymous, 2019). The occurrence of CLR and the pathogenic agent was revealed in our study. *Hemileia vastatrix,* the only pathogenic agent causing CLR disease in Vetnam, infected both *C. arabica* and *C. canephora* in three surveyed regions. However, the current genetic diversity and population structure of the coffee rust fungus in Vietnam has not been revealed. Therefore, the objectives of this chapter were to (i) evaluate the current genetic diversity and population structure of the coffee rust fungus based on sequencing of the ribosomal DNA internal transcribed spacer (rDNA-ITS) region and (ii) hypothesize the geographic region where the coffee rust fungus first established in Vietnam. The understanding first establishment of CLR fungus and the expansion pathway of this fungus in three main coffee-growing regions in Vietnam will contribute to proposing appropriate strategies to control this disease in Vietnam.

3.2 Materials and methods

Besides sequences from Vietnam, two sequences from Thailand, and forty ITS sequences of *H. vastatrix* from a few Central and South American countries (Mexico, Brazil, and

Colombia), and those from Portugal's Coffee Rust Research Centre (CIFC: Centro de Investigação de Ferrugens do Cafeeiro) were downloaded from the NCBI Genbank (Table 3.1) and added to the data, for identifying the original source(s) of CLR in Vietnam. Rust fungus samples from the same coffee-growing area (province or city) in different coffee-growing regions (Northwest, Central Highlands, and Southeast in Vietnam, and Thailand) were designated as a single population. Rust samples represented by ITS sequences downloaded from GenBank were treated similarly. Totally, 20 populations of *H.vastatrix* were examined in this study, including Southeast-Dong Nai, Southeast-Binh Phuoc, Central Highlands-Dak Lak, Central Highlands-Dak Nong, Central Highlands-Lam Dong, Northwest-Son La, and Northwest-Dien Bien populations from Vietnam; Chiang Mai population from Thailand; Veracruz, Chiapas, Oaxaca, and Puebla populations from Mexico; Minas Gerais-Coimbra, Minas Gerais-Capinopolis, Minas Gerais-São Sebastião do Paraíso, Minas Gerais-Senhora de Oliveira, and Espírito Santo-Venda Nova do Imigrante populations from Brazil; Quindío-Buenavista and Caldas-Chinchiná populations from Colombia; and the CIFC population from Portugal.

The obtained sequences were aligned using Clustal X version 2.0 (Larkin, 2007). The number of haplotypes, haplotype diversity (Hd) (Nei and Li, 1979), and nucleotide diversity (Pi) of each population were calculated using the Dnasp program version 6 (Librado and Rozas, 2009). To visualize the correlation between nucleotide diversity and geographical pattern, the regression statistics between nucleotide diversity and latitude was analyzed using PAST v4.03 (Hammer et al., 2001), and Microsoft Excel was used to generate a scatter diagram. The demographic pattern in each population and in the whole *H. vastatrix* population was evaluated by conducting Tajima's neutrality test (Tajima, 1989) using Arlequin version 3.5 (Excoffier and Lischer, 2010).

To evaluate the genetic structure of *H. vastatrix* populations, analyses of molecular variance (AMOVA) within each population and among different populations were calculated using GenAlEx 6.51b2 (Peakall and Smouse, 2012). Next, the matrices of pairwise genetic distance (Φ_{PT}) (Weir and Cockerham, 1984) between populations and individuals based on the AMOVA were calculated using GenAlEx 6.51b2. These matrices were used as input data for principal coordinate analysis (PCoA) to visualize the relationships between populations and the genetic structure of the whole *H. vastatrix* population.

To evaluate the genetic relationship among the haplotypes and to predict the ancestors of the populations in Vietnam, parsimony and NeighborNet network approaches were employed (Bryant and Moulton, 2013). Firstly, Dnasp program version 6 was used to generate a raw nexus file that was manually prepared to form an input trait file for PopART software version 1.7 (Leigh and Bryant, 2015). Next, a Median-joining haplotype network was constructed using PopART version 1.7. Meanwhile, the NeighborNet network (Bryant and Moulton, 2013) was generated using Splitstree 5 software (Huson and Bryant, 2006). Moreover, a map-based phylogenetic tree was generated to estimate the migration routes and the influence of geographic patterns on the similarities or differences among the *H. vastatrix* populations. First, a map figure was generated using GeoMapApp 3.6.12 (www.geomapapp.org) (Ryan et al., 2009). Next, a phylogenetic tree of the *H. vastatrix* populations was built using the program MEGA X (Kumar et al., 2018), based on the maximum likelihood statistical method with 1000 bootstrap replications, Generalized time-reversible model (Tavare, 1986) with the rate among sites Gamma distributed with invariant sites (G + I) was used. Lastly, a 3D graphical interface of the phylogeographic tree was generated by combining data from the map, phylogenic tree, and locations of collected samples using GenGIS2 software (Parks et al., 2013).

3.3 Results

The DNA sequences of the ITS region from all samples were 863 – 869 bp and had 96.34%–100% similarity with *H. vastatrix* (Table 2.3) (Genbank accession numbers: LC682363 – LC682405). From the eighty-six sequences (forty-one sequences from Vietnam), 123 polymorphic (segregating) sites (seventy-seven singleton variable sites and forty-six parsimony informative sites) were detected. In total, fifty-two haplotypes (thirty-six haplotypes from Vietnam) were identified (Table 3.2), with 0.952 ± 0.00022 haplotype diversity (Hd) and 0.00779 nucleotide diversity (Pi) (Table 3.3). The nucleotide diversity was highest (0.02651) in the Southeast (Binh Phuoc population), and the genetic variation was absent in the Veracruz (Mexico) population. Populations from Dak Nong (Central Highlands), Lam Dong (Central Highlands), and Dong Nai (Southeast) in Vietnam, and Chiang Mai in Thailand had higher diversity (0.02527, 0.01452, 0.01175, and 0.01153, respectively) (Table 3.3) than the remaining populations. A significant negative correlation between latitude and nucleotide diversity was detected when focusing on the populations in Vietnam and Thailand ($p < 0.05$) population (Figure 3.1). The value of Tajima's neutrality test (D) for the whole population of *H. vastatrix* was –0.19518. Although almost all populations in Vietnam had a negative value of Tajima's D (except for the Central Highlands-Lam Dong population); however, the values were not significant (Table 3.4).

Haplotype diversity of most of the populations was high (> 0.900) and did not show any geographic patterns, except the Caldas-Chinchiná population from Colombia, which showed a much lower value (0.500) than others (Table 3.3). When we focused on the haplotype distribution, the most common and major haplotype was Hap 8 (18.1%; detected in all populations except for the Thailand population), followed by Haps 17 [10.8%; Mexico (Veracruz, Chiapas, and Puebla), CIFC-Portugal, Brazil (Minas Gerais-São Sebastião do Paraíso)], Hap 32 [8.4%; Colombia and Brazil (Minas Gerais-Coimbra)], and three haplotypes, namely, Hap 9 (the Northwest and the Central Highlands of Vietnam), Hap 24 [Mexico (Chiapas and Oaxaca)], and Hap 27 (Colombia and the Southeast of Vietnam). The remaining haplotypes were singleton (Table 3.2 and Figure 3.2).

The Analysis of molecular variance (AMOVA) showed that genetic differentiation among the twenty *H. vastatrix* populations from the six countries (Vietnam, Thailand, Brazil, Colombia, Mexico, and Portugal) was low ($F_{ST} = 0.175$) but significant ($p < 0.01$). Only 17% of the total variation was found among populations, while 83% of the differences were among individuals within a population (Table 3.5). The genetic differentiation (Φ_{PT}) between the Vietnamese–Thai and Central and South American populations, and the Brazilian and Caldas-Chinchiná population (Colombia) was high (Table 3.6). Meanwhile, the genetic differentiation between the Mexican and Vietnamese–Thai populations was lower than that observed between other populations (Table 3.6). The PCoA chart indicated that the Vietnamese–Thai and the Mexican populations (Puebla, Chiapas, Veracruz, and Oaxaca) had closer genetic distance than combinations with the Brazilian–Colombian populations (Figure 3.3).

Analysis of the genetic structure of the whole *H. vastatrix* population (by studying the genetic distance between individuals) revealed that most individuals were closely gathered, while some from the Central Highlands and Southeast were distantly scattered (TSH-R30034, 30045, 30046, 30051, 30067, 30071, 30080, 30085, 30088, 30101, and 30118) (Figure 3.4). Haplotype (Figure 3.2) and NeighborNet (Figure 3.5) networks showed three groups in the *H. vastatrix* populations. The first group included Haps 13, 14, 19, 20, 21, 22, 23, 25, 28, 29, 30, 31, 32, 33, and 34, which were derived from Hap 8, and were mostly from Central and South American countries and the Northwest of Vietnam. There were 19 haplotypes in group 2, namely, Haps 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 15, 16, 17, 18, 38, 40, 41, and 45 from Vietnam, Thailand, Brazil, and CIFC-Portugal. The remaining haplotypes (Haps 26, 27, 35, 36, 37, 39, 42, 43, 44, 46, 47, 48, 49, 50, 51, and 52) were included in the third group and were mainly from the Central Highlands and the Southeast of Vietnam.

The map-based phylogenetic tree highlighted the close genetic relationship of some populations in Mexico and Brazil with the Northwest (Vietnam) populations. Populations from

Oaxaca, Chiapas, and Puebla (Mexico) were directly related to populations from Son La and Dien Bien (Northwest of Vietnam) (Figure 3.6A). Besides that, some individuals from Minas Gerais and Espírito Santo (Brazil) had close relationships with some individuals from the Northwest of Vietnam (Figure 3.6A). On the contrary, a population from Oaxaca had a closer relationship with some individuals from Dong Nai (Southeast of Vietnam). Figure 3.6B shows the close relationship of the Chiang Mai (Thailand) population with populations in the Central Highlands and the Southeast (Dong Nai) of Vietnam. Moreover, all populations in the Central Highlands and the Southeast were connected with the Northwest populations (Figure 3.6B).

3.4 Discussion and conclusion

Hemileia vastatrix in Vietnam has low genetic divergence and unstructured population but high haplotype diversity

Genetic diversity and population structure analyses disclosed the real picture of the *H. vastatrix* populations in Vietnam. Firstly, the genetic differentiation of *H. vastatrix* in Vietnam is low, as shown in the calculated AMOVA (Table 3.5) and Table 3.6. The percentage of variation within populations surpasses that among populations (Table 3.5), indicating that the genetic divergence among populations is small. Moreover, the nucleotide diversity values of populations in Vietnam are low (less than 0.1; Table 3.3). Secondly, *H. vastatrix* populations in Vietnam are unstructured. The PCoA diagram (Figure 3.4) illustrates that some individuals are scattered from others. Dispersed haplotypes are also well illustrated as they are divided into groups distinguished from two others (Figures 3.2 and 3.5); however, these groups are divided without relation to the regions where CLR samples were collected. Lastly, the haplotype diversity of *H. vastatrix* in Vietnam is high. This is not only because of their high values showed in Table 3.3 but also because many groups of haplotype are divided into networks (Figures 3.2 and 3.5). Besides that, the linear regression (Figure 3.1) shows that the nucleotide diversity has a negative relationship with latitude, the higher the latitude, the lower the nucleotide diversity value. Most haplotypes in Vietnam are unique. Mutations in the *H. vastatrix* populations of Vietnam are likely to be spontaneous without being affected by geographical factors or the host's resistance genes. Nonetheless, random alterations due to environmental changes, such as warmer climate or increased annual rainfall, could result in the formation of new virulent races (Ali et al., 2014). Furthermore, the unstructured population of *H. vastatrix* could result from asexual reproduction (Ramírez-Camejo et al., 2021). The results of this study are similar to those from studies in the coffee-growing regions of Peru and Brazil (Cinthia et al., 2017; Cinthia et al., 2021; Maia et al., 2013; Cabral et al., 2016).

The hypothesized original sources and the migration direction of *Hemileia vastatrix* in Vietnam

Based on the population genetic analyses, a hypothetical interpretation of the first establishment of CLR disease and the successive spore transportation pathways through coffee-growing areas in Vietnam is possible. The frequency of Haps 5, 8, 9, 11, 17, 27, and 32 in the whole *H. vastatrix* population and their positions in the phylogenetic analyses (Figure 3.7 and Figures 3.2, 3.5) indicate that these haplotypes are shared in most populations and steady in the reproductive process. Thus, these haplotypes are referred to as the 'early established haplotypes'. Among them, Hap 32 originated from Hap 8; and Haps 11 and 5 are descendants of Hap 9. Haps 8, 17, 27, and 32 are shared between Brazil, Colombia, CIFC-Portugal, and Mexico. In addition, the pairwise genetic distance matrices (Table 3.6), PCoA diagram (Figure 3.3), and phylogeographic tree (Figure 3.6A) highlight the close genetic distances and genetic similarities between some populations in Mexico and Brazil and the Northwest population in Vietnam. These results prove that *H. vastatrix* populations from Vietnam (mostly from the Northwest) and those from Central and South American countries (Mexico and Brazil) have very close relationships. However, I cannot ascertain that the *H. vastatrix* population in Vietnam spread from these countries just because the first rust detection in Vietnam was in 1890, while the first confirmation of rust establishment in Brazil and Mexico was in the 1970s and 1980s, respectively (McCook, 2006; McCook, 2019). On second thought, *H. vastatrix* might have migrated to Vietnam from African countries because Vietnam and these countries were colonized by the French government during World War I. *Hemileia vastatrix* was first identified in Vietnam in 1890, the exact timeline of the global CLR pandemic (1875−1920). At that time, rust-infected coffee seedlings or materials from infected areas in Africa (where *H. vastatrix* is native) might have been carried to Vietnam by the French. However, confirming the origin of CLR in Vietnam necessitates further studies with larger rust samples from other countries, especially African countries. For these reasons, this discussion only focused on the first establishment of CLR in the three main coffee growing regions in Vietnam.

Three main inferences regarding the origin and migration route of *H. vastatrix* can be gathered based on the present study. Firstly, because Φ_{PT} was previously measured only based on the genetic drift (Hedrick, 2005), the low genetic distance recorded between populations in Vietnam (Table 3.6 and Figure 3.3) proves that the probability of migration between these regions was high. Secondly, most ancestral haplotypes (Haps 5, 8, 9, 11) are present in the Northwest, and the percentage of those ancestral haplotypes in the Northwest is larger than that in other areas (Figure 3.7). Thirdly, the phylogeographic tree shows that *H. vastatrix* from Brazil and Mexico have genetic similarity with populations in the Northwest of Vietnam (Figure 3.6A); all populations from southern Vietnam had a close genetic relationship with the Northwest population (Figure 3.6B). Therefore, I can infer that *H. vastatrix* existed in the north of Vietnam in the early years of introducing coffee, then spread to the Southeast and the Central Highlands as coffee cultivation was extended to these areas. Although *C. arabica* was cut down for Catimor coffee in northern Vietnam (Anonymous, 2019), urediniospores of the rust fungus could still exist by parasitizing coffee plants in other plantations (the replacement of coffee varieties was not synchronized well and did not occur at the same time in all plantations of a region) or wild coffee plants and infected Catimor coffee, making this coffee variety susceptible to coffee rust. Hap 8 is one of the first ancestors that spread in northern Vietnam. After that, this haplotype became widespread in coffee plantations in the Northwest and formed Haps 13, 14, and 25. Later, it was transported to the South via human activities or infected materials and formed Haps 20, 27, and 34. Hap 27 is one of the ancestral haplotypes in the Southeast. Meanwhile, another ancestor, Hap 9, is widespread and formed Haps 5, 10, 11, and 12. This ancestor and some haplotypes (Haps 5 and 11) were then carried to the Central Highlands of Vietnam and formed Haps 1, 3, and 15 (Figures 3.2 and 3.5). Thus, Haps 5, 8, 9, 11, and 27 are the ancestral haplotypes of the Vietnamese *H. vastatrix*.

On the other hand, data from the population genetic analyses (Figures 3.2, 3.4, 3.5) suggest that coffee rust in Vietnam has another origin besides spreading from the Northwest. This is indicated in the PCoA diagram (Figure 3.4): some individuals from the Central Highlands and the Southeast are scattered distantly from others. Figures 3.2 and 3.5 also pointed out that some haplotypes (Haps 46–52) are genetically distant from others and form a different group. Information from the history of coffee in Vietnam consolidates this hypothesis. *Coffea canephora* was introduced in the Central Highlands of Vietnam to replace *C. arabica* in the early 1900s (D'haeze et al., 2003; Anonymous, 2019). When *C. canephora* spread to southern Vietnam, they might have also brought *H. vastatrix* to this region, broadened, and formed other branches of CLR in Vietnam. The appearance of absent haplotype (black node) in the haplotype network and the NeighborNet (Figures 3.2, 3.5) suggest that other haplotypes are yet to be discovered, and some of them are likely to be ancestral for haplotypes found in this study.

◆ This chapter have already been published in the following paper: Le CTM, Okane I, Ono Y, Tsuda Y, Yamaoka Y. Incidence of Coffee Leaf Rust in Vietnam, Possible Original Sources and Subsequent Pathways of Migration. Front Plant Sci. 2022 Apr 5;13:872877. doi: 10.3389/fpls.2022.872877. PMID: 35449887; PMCID: PMC9016365. https://doi.org/10.3389/fpls.2022.872877.

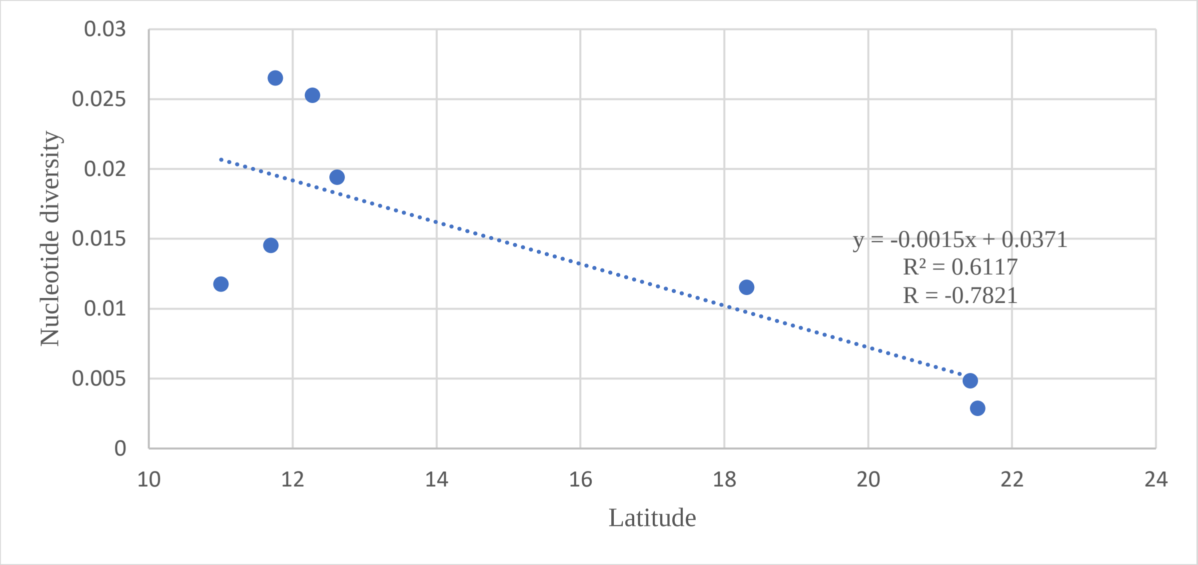


Figure 3.1. Linear regression model for the latitude and nucleotide diversity of *Hemileia vastatrix* populations from Vietnam and Thailand. The dotted line represents the regression line with equation y = 0.037 – 0.0015x.

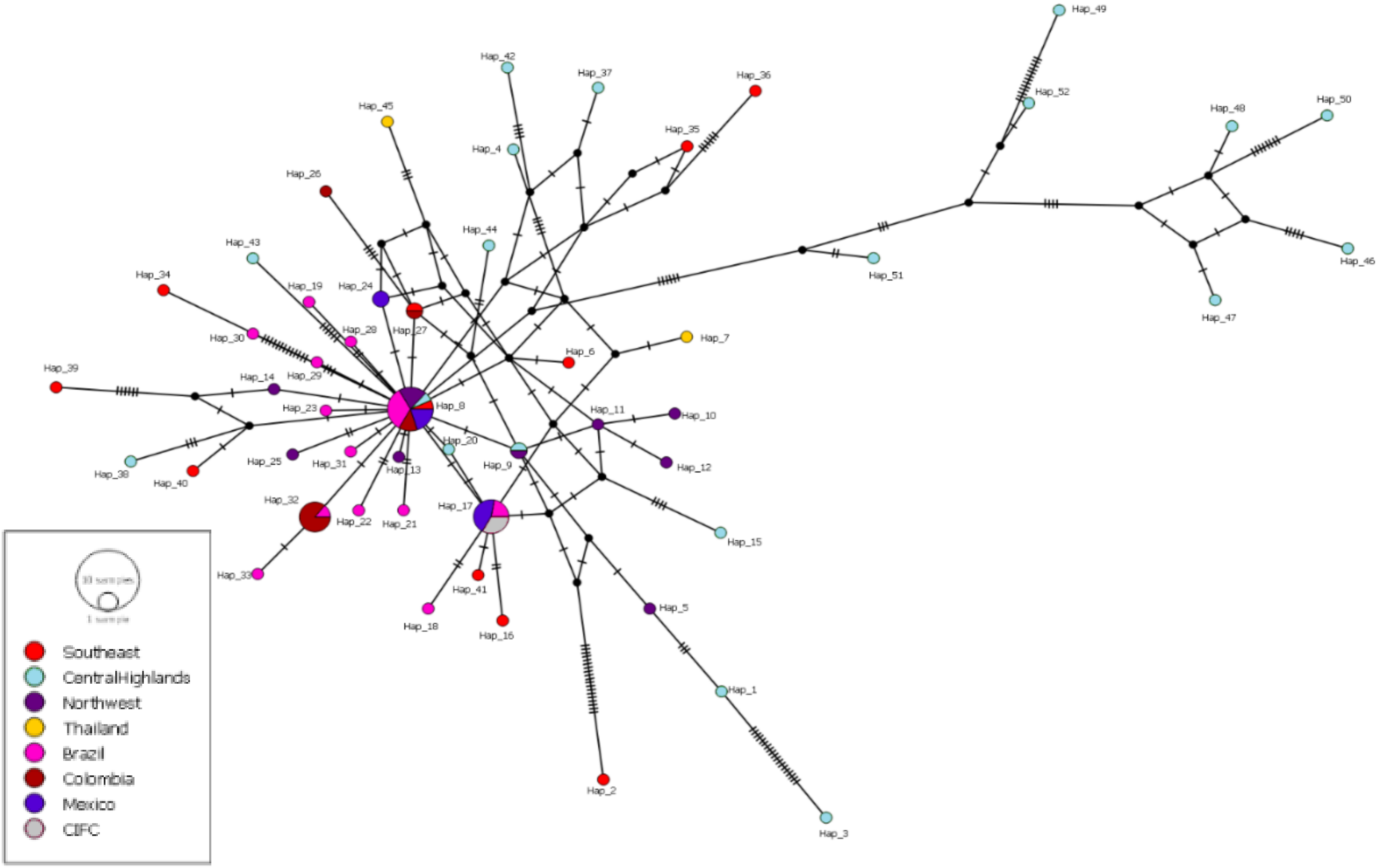


Figure 3.2. Median-joining haplotype network inferred from the ITS regions of *Hemileia vastatrix* sampled from Vietnam, Thailand, Portugal, and other countries in America (Brazil, Colombia, Mexico). Circles represent haplotypes, with bright red corresponding to samples collected from Southeast Vietnam, light turquoise the Central Highlands of Vietnam, purple the Northwest of Vietnam, while yellow indicates Thailand, pink Brazil, brick red Colombia, dark blue Mexico, and grey CIFC (Portugal). The size of the circles is proportional to the number of sequences in each haplotype, one transverse line represents simple mutations, and black nodes are non-sampled haplotypes.

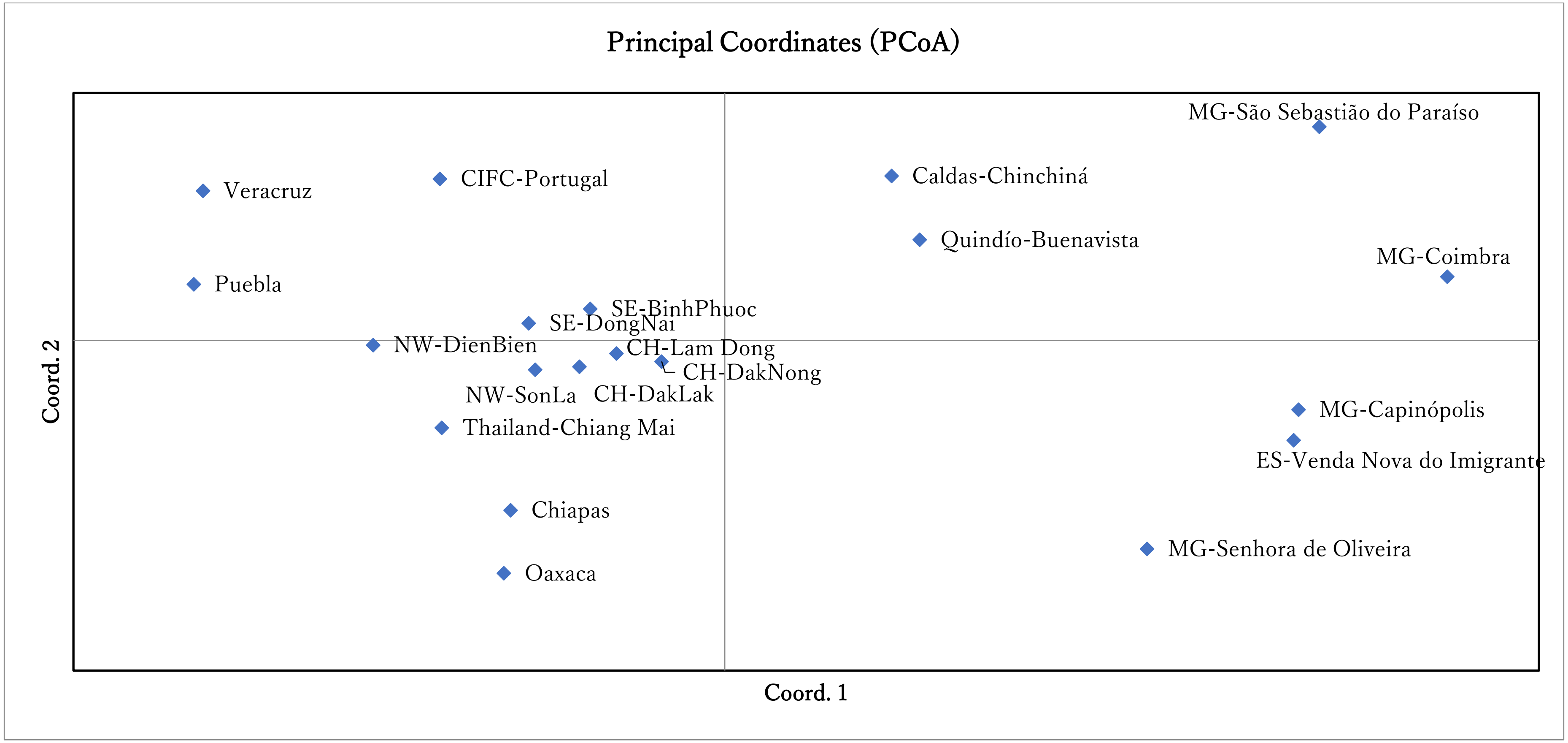


Figure 3.3. Principal Coordinate Analysis (PCoA) of the genetic (nucleotide) distance in the *Hemileia vastatrix* subpopulations in Southeast Asia and the Americas. Subpopulations that are closer together have smaller divergence. NW: Northwest, CH: Central Highands, SE: Southeast, CIFC: Centro de Investigação de Ferrugens do Cafeeiro, MG: Minas Gerais, ES: EspíritoSanto.

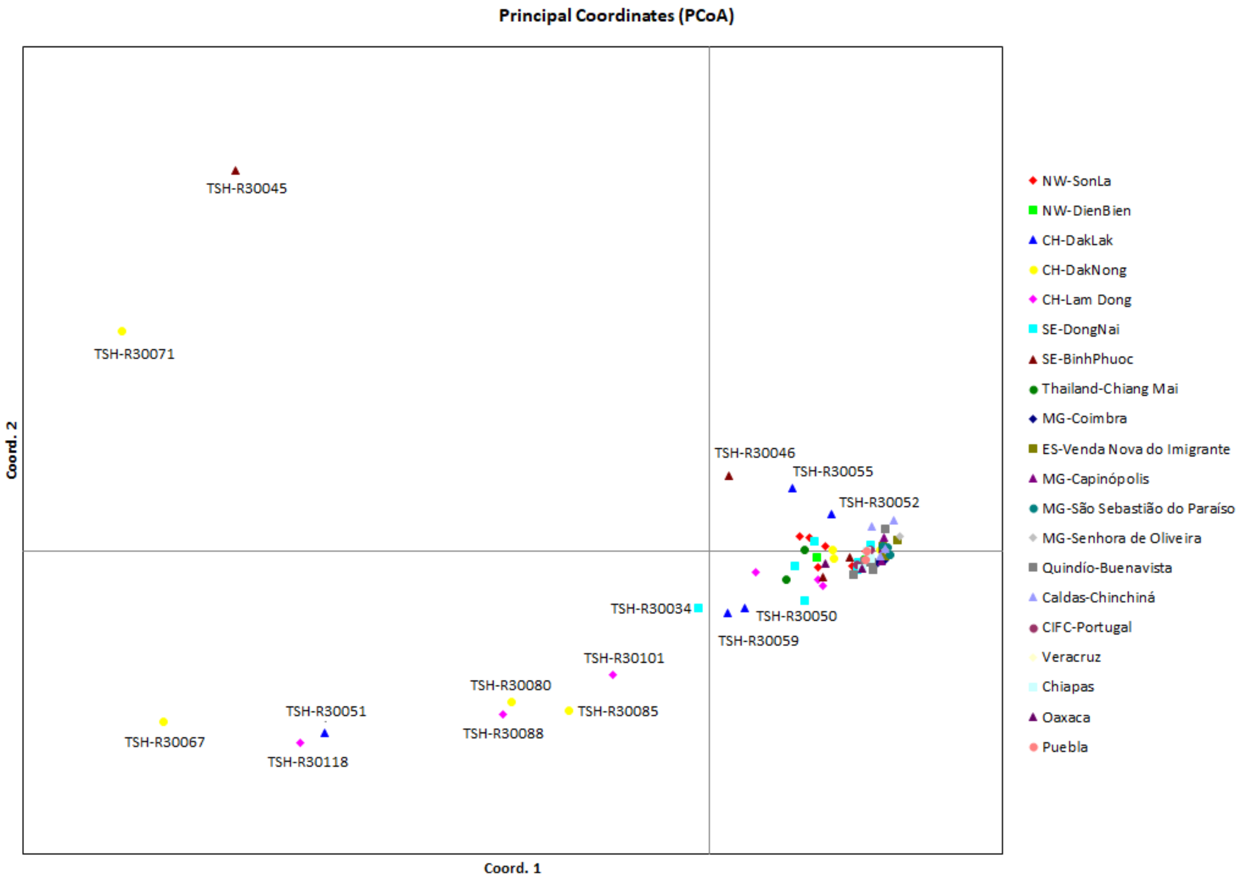


Figure 3.4. Principal coordinate analysis (PCoA) of the genetic (nucleotide) diversity of *Hemileia vastatrix* in Vietnam, Thailand, Portugal, and some countries in the Americas (Brazil, Colombia, Mexico). Individuals that are closer together have smaller genetic distances. NW: Northwest (Vietnam), CH: Central Highands (Vietnam), SE: Southeast (Vietnam), CIFC: Centro de Investigação de Ferrugens do Cafeeiro (Portugal), MG: Minas Gerais (Brazil), ES: Espírito Santo (Brazil).

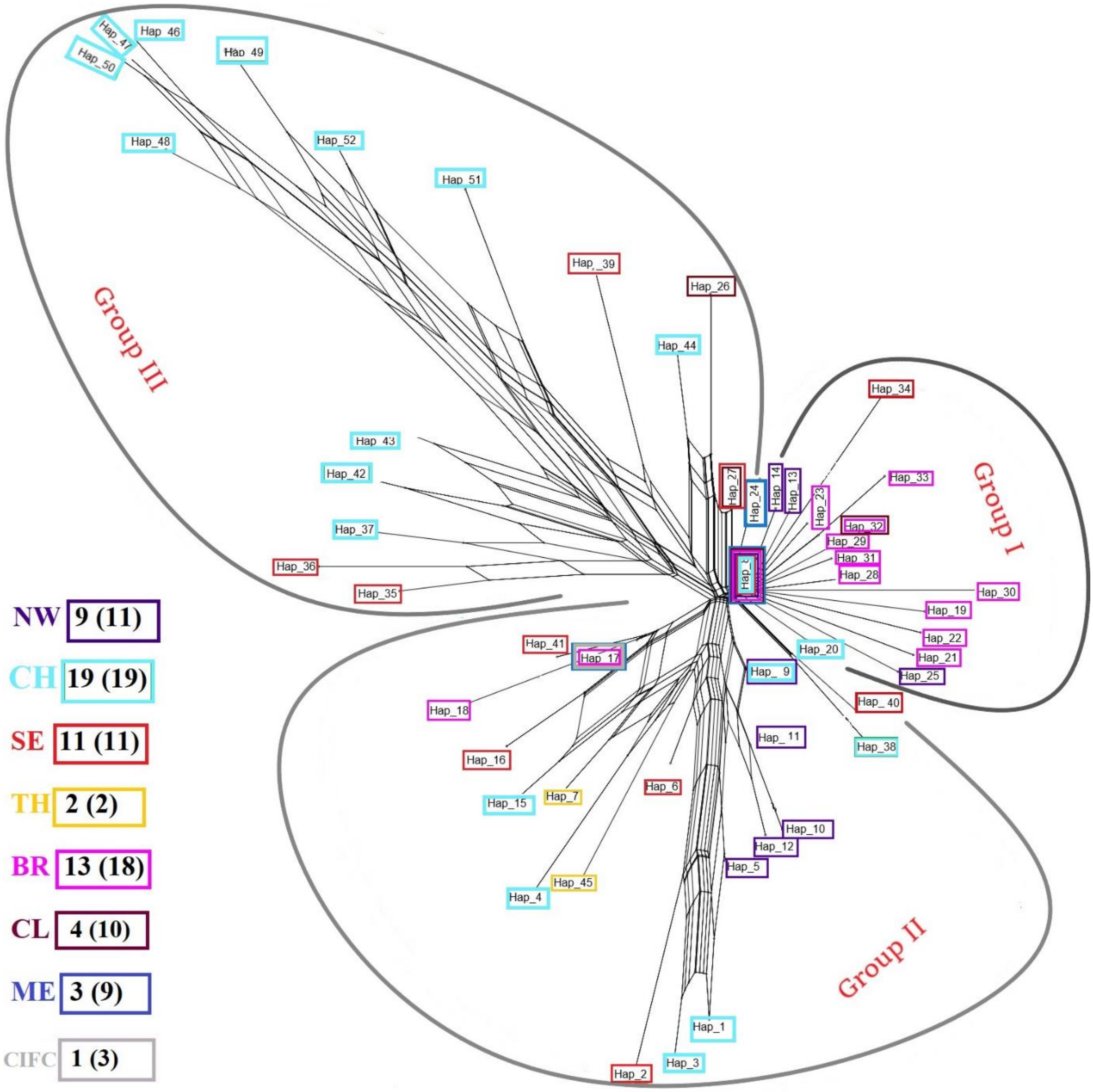


Figure 3.5. NeighborNet graph constructed from the pairwise F_{ST} values of *Hemileia vastatrix* samples from Vietnam, Thailand, Portugal, and other American countries (Brazil, Colombia, Mexico) based on their ITS regions. The haplotypes detected in seven areas are framed in boxes with different colors: purple indicates the Northwest (NW) of Vietnam, light turquoise the Central Highlands (CH) of Vietnam, bright red the Southeast (SE) of Vietnam, while yellow indicates Thailand (TH), pink Brazil (BR), brick red Colombia (CL), blue Mexico (ME), and grey the Centro de Investigação de Ferrugens do Cafeeiro (CIFC)-Portugal. After the abbreviation of each region the number of haplotypes detected in each area is stated, and the number in the parentheses indicates the number of sequences obtained from the region.

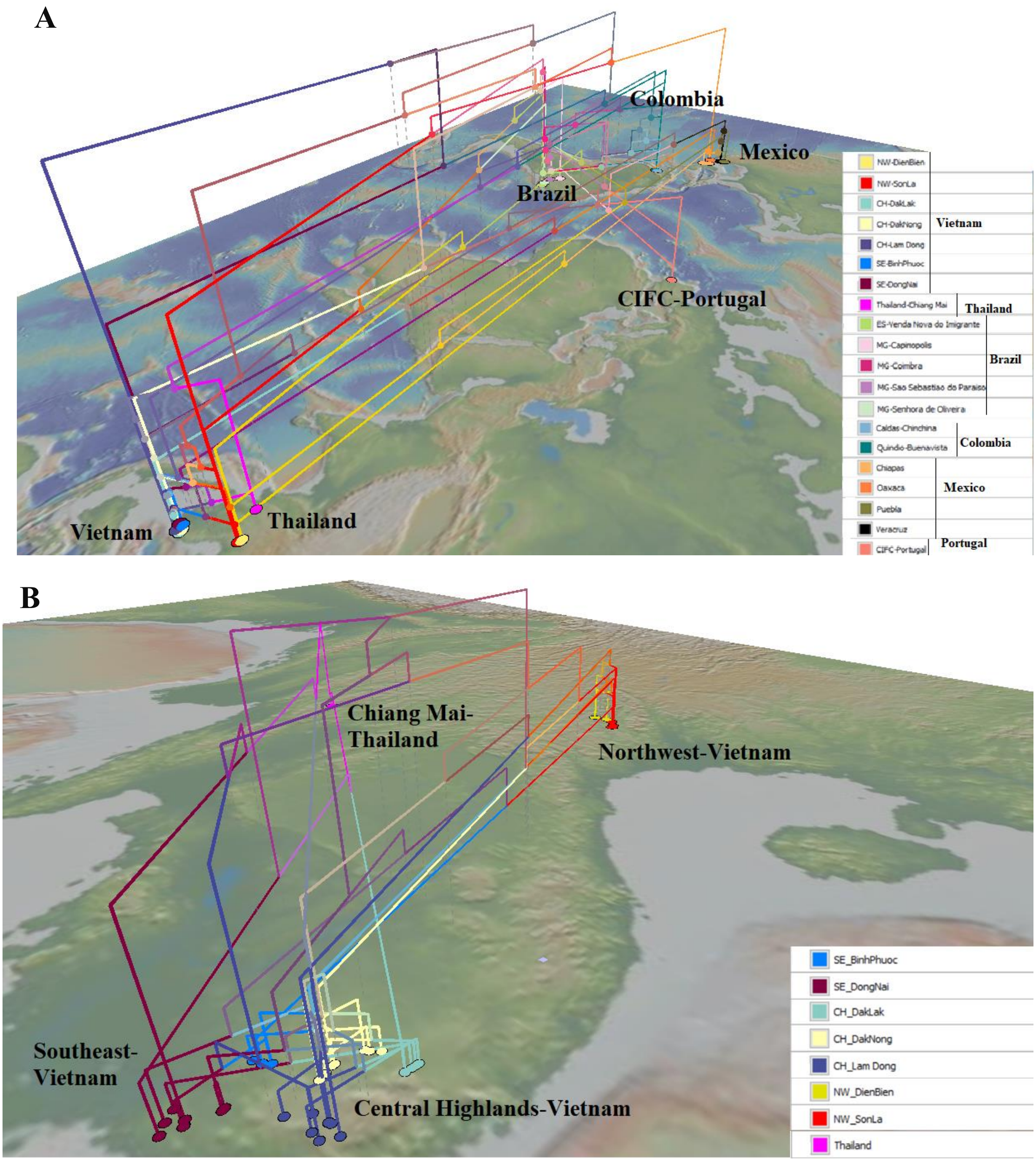


Figure 3.6. Geophylogeny of *Hemileia vastatrix* from Vietnam and, Thailand and other American countries (Brazil, Colombia, Mexico) using ITS sequences. Colored nodes on the map correspond to sampling locations. (A) Geophylogeny of *Hemileia vastatrix* from Vietnam, Thailand, and Central and South American countries, (B) Geophylogeny of *Hemileia vastatrix* from areas in Vietnam and Chiang Mai-Thailand. NW: Northwest (Vietnam), CH: Central Highands (Vietnam), SE: Southeast (Vietnam), CIFC: Centro de Investigação de Ferrugens do Cafeeiro (Portugal), MG: Minas Gerais (Brazil), ES: Espírito Santo (Brazil).

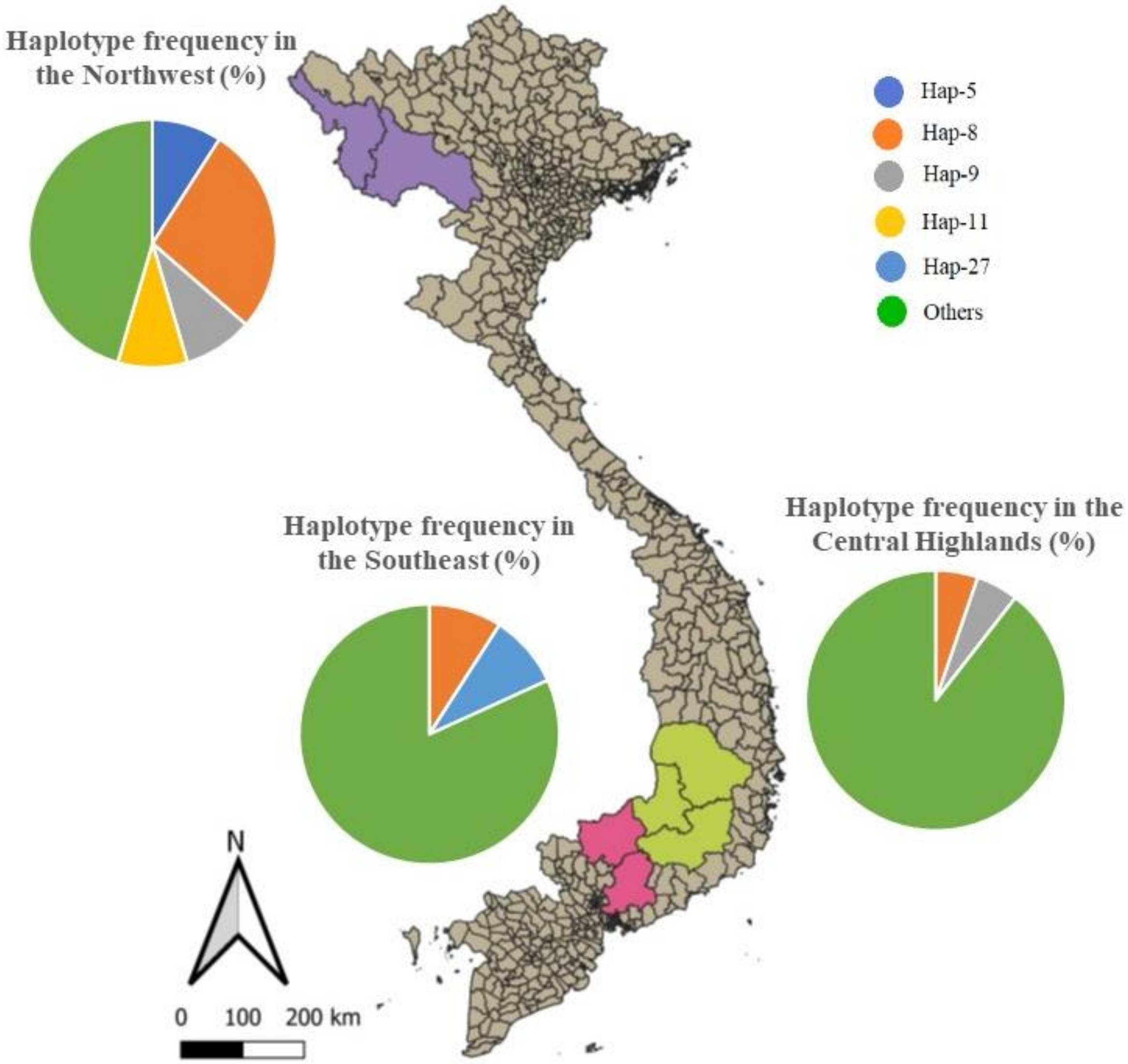


Figure 3.7. The frequency of ancestral *Hemileia vastatrix* haplotypes in different geographical sample regions of Vietnam. Colored areas on the map correspond to sampling regions: purple color indicates the Northwest, greenish-yellow signifies the Central Highlands, and reddish pink shows the Southeast.

Table 3.1. ITS sequences of *Hemileia vastatrix* dowloaded from Genbank NCBI

NCBI name	Country, year of collection	State	Latitude	Longitude	Accession #	Identifiers in this study	Host
H. vastatrix voucher CMPH-234	Mexico, 2017	Veracruz	19.43000	-96.90000	KX260253.1	Mexico_1	*Coffea arabica cv. Caturra Rojo*
H. vastatrix voucher CMPH-211	Mexico, 2017	Veracruz	19.17000	-96.96000	KX260252.1	Mexico_2	*C. arabica cv. Pluma Hidalgo*
H. vastatrix voucher CMPH-192	Mexico, 2017	Chiapas	16.70000	-92.83000	KX260251.1	Mexico_3	*C. arabica cv. Borbon*
H. vastatrix voucher CMPH-165	Mexico, 2017	Chiapas	15.00000	-92.15000	KX260250.1	Mexico_4	*C. canephora cv. Robusta*
H. vastatrix voucher CMPH-142	Mexico, 2017	Chiapas	15.05000	-92.07000	KX260249.1	Mexico_5	*C. arabica cv. Caturra Rojo*
H. vastatrix voucher CMPH-95	Mexico, 2017	Oaxaca	15.94000	-96.46000	KX260248.1	Mexico_6	*C. arabica cv. Caturra Rojo*
H. vastatrix voucher CMPH-80	Mexico, 2017	Oaxaca	15.95000	-96.45000	KX260247.1	Mexico_7	*C. arabica cv. Paca*
H. vastatrix voucher CMPH-14	Mexico, 2017	Puebla	20.60000	-97.69000	KX260246.1	Mexico_8	*C. arabica* X *C. canephora cv. Oro Azteca*
H. vastatrix isolate 37-37-9	Mexico, 2017	Puebla	20.37000	-97.86000	MF417748.1	Mexico_9	*C. arabica* X *C. canephora* (*Coffea* 832/1 X *C. arabica*) *cv. Costa Rica 95*
H. vastatrix isolate CT-9-8	Colombia, 2007	Quindío-Buenavista	4.35857	-75.74045	EF394132.1	Colombia_1	*C. arabica cv. Caturra*
H. vastatrix isolate CT-9-9	Colombia, 2008	Quindío-Buenavista	4.35857	-75.74045	EF394132.2	Colombia_2	*C. arabica cv. Caturra*
H. vastatrix isolate CT_13_v	Colombia, 2007	Quindío-Buenavista	4.35857	-75.74045	EF394130.1	Colombia_3	*C. arabica cv. Caturra*
H. vastatrix isolate CTInNa2	Colombia, 2007	Caldas-Chinchiná	4.98294	-75.61024	EF394129.1	Colombia_4	*C. arabica cv. Caturra*

CIFC: Centro de Investigação das Ferrugens do Cafeeiro

(Continued)

NCBI name	Country, year of collection	State	Latitude	Longitude	Accession #	Identifiers in this study	Host
H. vastatrix isolate CTSol_1	Colombia, 2007	Quindío-Buenavista	4.35857	-75.74045	EF394128.1	Colombia_5	*C. arabica cv. Caturra*
H. vastatrix isolate vColPa1_2	Colombia, 2007	Quindío-Buenavista	4.35857	-75.74045	EF394127.1	Colombia_6	*C. arabica cv. Caturra*
H. vastatrix isolate CT_18_N	Colombia, 2007	Caldas-Chinchiná	4.98294	-75.61024	EF394126.1	Colombia_7	*C. arabica cv. Caturra*
H. vastatrix isolate CTInNa1	Colombia, 2007	Caldas-Chinchiná	4.98294	-75.61024	EF394125.1	Colombia_8	*C. arabica cv. Caturra*
H. vastatrix isolate CT_12_v	Colombia, 2007	Quindío-Buenavista	4.35857	-75.74045	EF394124.1	Colombia_9	*C. arabica cv. Caturra*
H. vastatrix isolate CT_7_N	Colombia, 2007	Caldas-Chinchiná	4.98294	-75.61024	EF394123.1	Colombia_10	*C. arabica cv. Caturra*
H. vastatrix isolate raceXXII-535	NA	CIFC-Portugal	38.70152	-9.31236	EF394122.1	CIFC_1	*C. arabica cv. Caturra*
H. vastatrix isolate RaceII-1126	NA	CIFC-Portugal	38.70152	-9.31236	EF394120.1	CIFC_2	*C. arabica cv. Caturra*
H. vastatrix isolate RaceXXII-637	NA	CIFC-Portugal	38.70152	-9.31236	EF394119.1	CIFC_3	*C. arabica cv. Caturra*
H. vastatrix isolate Coimbra_7	Brazil, 2018	Minas Gerais-Coimbra	-20.84679	-42.84445	MF627828.1	Brazil_1	NA
H. vastatrix isolate Coimbra_6	Brazil, 2018	Minas Gerais-Coimbra	-20.84679	-42.84445	MF627827.1	Brazil_2	NA
H. vastatrix isolate Coimbra_4	Brazil, 2018	Minas Gerais-Coimbra	-20.84679	-42.84445	MF627825.1	Brazil_3	NA
H. vastatrix isolate Coimbra_3	Brazil, 2018	Minas Gerais-Coimbra	-20.84679	-42.84445	MF627824.1	Brazil_4	NA
H. vastatrix isolate Coimbra_2	Brazil, 2018	Minas Gerais-Coimbra	-20.84679	-42.84445	MF627823.1	Brazil_5	NA

CIFC: Centro de Investigação das Ferrugens do Cafeeiro

(Continued)

NCBI name	**Country, year of collection**	**State**	**Latitude**	**Longitude**	**Accession #**	**Identifiers in this study**	**Host**
H. vastatrix isolate Coimbra_1	Brazil, 2018	Minas Gerais-Coimbra	-20.84679	-42.84445	MF627822.1	Brazil_6	NA
H. vastatrix isolate Venda_Nova_do_Imigrante_7	Brazil, 2018	Espírito Santo-Venda Nova do Imigrante	-20.33291	-41.12964	MF627821.1	Brazil_7	NA
H. vastatrix isolate Venda_Nova_do_Imigrante_6	Brazil, 2018	Espírito Santo-Venda Nova do Imigrante	-20.33291	-41.12964	MF627820.1	Brazil_8	NA
H. vastatrix isolate Venda_Nova_do_Imigrante_5	Brazil, 2018	Espírito Santo-Venda Nova do Imigrante	-20.33291	-41.12964	MF627819.1	Brazil_9	NA
H. vastatrix isolate Venda_Nova_do_Imigrante_4	Brazil, 2018	Espírito Santo-Venda Nova do Imigrante	-20.33291	-41.12964	MF627818.1	Brazil_10	NA
H. vastatrix isolate Capinópolis_6	Brazil, 2018	Minas Gerais-Capinópolis	-18.68314	-49.57076	MF627814.1	Brazil_11	NA
H. vastatrix isolate Capinópolis_5	Brazil, 2018	Minas Gerais-Capinópolis	-18.68314	-49.57076	MF627813.1	Brazil_12	NA
H. vastatrix isolate Capinópolis_4	Brazil, 2018	Minas Gerais-Capinópolis	-18.68314	-49.57076	MF627812.1	Brazil_13	NA
H. vastatrix isolate São_Sebastião_do_Paraíso_6	Brazil, 2018	Minas Gerais-São Sebastião do Paraíso	-20.91775	-46.99968	MF627808.1	Brazil_14	NA
H. vastatrix isolate São_Sebastião_do_Paraíso_5	Brazil, 2018	Minas Gerais-São Sebastião do Paraíso	-20.91775	-46.99968	MF627807.1	Brazil_15	NA

CIFC: Centro de Investigação das Ferrugens do Cafeeiro

(Continued)

NCBI name	**Country, year of collection**	**State**	**Latitude**	**Longitude**	**Accession #**	**Identifiers in this study**	**Host**
H. vastatrix isolate São_Sebastião_do_Paraíso_4	Brazil, 2018	Minas Gerais-São Sebastião do Paraíso	-20.91775	-46.99968	MF627806.1	Brazil_16	NA
H. vastatrix isolate Senhora_de_Oliveira_5	Brazil, 2018	Minas Gerais-Senhora de Oliveira	-20.79387	-43.34312	MF627802.1	Brazil_17	NA
H. vastatrix isolate Senhora_de_Oliveira_4	Brazil, 2018	Minas Gerais-Senhora de Oliveira	-20.79387	-43.34312	MF627801.1	Brazil_18	NA
H. vastatrix (TSH-R59309)	Thailand, 2018	Chiang Mai	18.767	99.02616	TL_18TH13	TSH-R59309	NA
H. vastatrix (TSH-R59420)	Thailand, 2019	Chiang Mai	13.75398	100.50144	TL_19TH46	TSH-R59420	NA

CIFC: Centro de Investigação das Ferrugens do Cafeeiro

Table 3.2. List of haplotypes obtained from sequences of coffee leaf rust fungus in Vietnam, Thailand, Portugal, and some American countries (Brazil, Colombia, Mexico)

Hap	Frequency	Distribution																				Sequence ID.
		SE-DN	SE-BP	CH-DL	CH-DN	CH-LD	NW-SL	NW-DB	TL-CM	MG-CB	ES-VNI	MG-CP	MG-SSP	MG-SO	Q-B	C-C	V	C	O	P	C-P	
Hap_1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30055
Hap_2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30045
Hap_3	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30071
Hap_4	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30109
Hap_5	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30148
Hap_6	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30036
Hap_7	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	59309
Hap_8	15	1	0	0	0	1	1	2	0	3	1	0	1	0	1	1	0	1	1	1	0	Mexico_6, 30160, 30172, Brazil_8, Brazil_1, Mexico_4, Mexico_8, Colombia_3, Brazil_5, Brazil_6, Colombia_4, Brazil_14, 30022, 30107, 30137
Hap_9	2	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30120, 30073
Hap_10	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	30155
Hap_11	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30133
Hap_12	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30145
Hap_13	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30129
Hap_14	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	30167
Hap_15	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30052
Hap_16	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30028

SE-DN: Southeast-Dong Nai (Vietnam), SE-BP: Southeast-Binh Phuoc (Vietnam), CH-DL: Central Highlands-Dak Lak (Vietnam), CH-DN: Central Highlands-Dak Nong (Vietnam), CH-LD: Central Highlands-Lam Dong (Vietnam), NW-SL: Northwest-Son La (Vietnam), NW-DB: Northwest-Dien Bien (Vietnam), TL-CM: Thailand-Chiang Mai, MG-CB: Minas Gerais-Coimbra (Brazil), ES-VNM: Espírito Santo-Venda-Nova-do-Imigrante (Brazil), MG-CP: Minas Gerais- Capinopolis (Brazil), MG-SSP: Minas Gerais-Sao-Sebastiao-do-Paraiso (Brazil), MG-SO: Minas Gerais-Senhora-de-Oliveira (Brazil), Q-B: Quindío-Buenavista (Colombia), C-C: Caldas-Chinchiná (Colombia), V: Veracruz (Mexico), C: Chiapas (Mexico), O: Oaxaca (Mexico), P: Pueble (Mexico), C-P: Centro de Investigação de Ferrugens do Cafeeiro (Portugal).

(Continued)

Hap	Frequency	Distribution																				Sequence ID.
		SE-DN	SE-BP	CH-DL	CH-DN	CH-LD	NW-SL	NW-DB	TL-CM	MG-CB	ES-VNI	MG-CP	MG-SSP	MG-SO	Q-B	C-C	V	C	O	P	C-P	
Hap_17	9	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2	1	0	1	3	Mexico_3, Mexico_2, Mexico_1, Mexico_9, CIFC_1, CIFC_2, CIFCI_3, Brazil_15, Brazil_16
Hap_18	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	Brazil_12
Hap_19	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	Brazil_18
Hap_20	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30069
Hap_21	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	Brazil_7
Hap_22	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	Brazil_11
Hap_23	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	Brazil_17
Hap_24	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	Mexico_5, Mexico_7
Hap_25	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30125
Hap_26	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Colombia_1
Hap_27	2	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Colombia_2, 30018
Hap_28	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	Brazil_4
Hap_29	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	Brazil_2
Hap_30	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	Brazil_10
Hap_31	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	Brazil_13

SE-DN: Southeast-Dong Nai (Vietnam), SE-BP: Southeast-Binh Phuoc (Vietnam), CH-DL: Central Highlands-Dak Lak (Vietnam), CH-DN: Central Highlands-Dak Nong (Vietnam), CH-LD: Central Highlands-Lam Dong (Vietnam), NW-SL: Northwest-Son La (Vietnam), NW-DB: Northwest-Dien Bien (Vietnam), TL-CM: Thailand-Chiang Mai, MG-CB: Minas Gerais-Coimbra (Brazil), ES-VNM: Espírito Santo-Venda-Nova-do-Imigrante (Brazil), MG-CP: Minas Gerais- Capinopolis (Brazil), MG-SSP: Minas Gerais-Sao-Sebastiao-do-Paraiso (Brazil), MG-SO: Minas Gerais-Senhora-de-Oliveira (Brazil), Q-B: Quindío-Buenavista (Colombia), C-C: Caldas-Chinchiná (Colombia), V: Veracruz (Mexico), C: Chiapas (Mexico), O: Oaxaca (Mexico), P: Pueble (Mexico), C-P: Centro de Investigação de Ferrugens do Cafeeiro (Portugal).

(Continued)

Hap	Frequency	Distribution																				Sequence ID.
		SE-DN	SE-BP	CH-DL	CH-DN	CH-LD	NW-SL	NW-DB	TL-CM	MG-CB	ES-VNI	MG-CP	MG-SSP	MG-SO	Q-B	C-C	V	C	O	P	C-P	
Hap_32	7	0	0	0	0	0	0	0	0	1	0	0	0	0	3	3	0	0	0	0	0	Colombia_5, Brazil_3, Colombia_7, Colombia_8, Colombia_10, Colombia_9, Colombia_6
Hap_33	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	Brazil_9
Hap_34	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30046
Hap_35	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30031
Hap_36	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30034
Hap_37	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30098
Hap_38	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30065
Hap_39	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30024
Hap_40	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30042
Hap_41	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30043
Hap_42	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30050
Hap_43	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30059
Hap_44	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30104
Hap_45	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	59420
Hap_46	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30051
Hap_47	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30088
Hap_48	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30080
Hap_49	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30067
Hap_50	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30118
Hap_51	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30101
Hap_52	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30085

SE-DN: Southeast-Dong Nai (Vietnam), SE-BP: Southeast-Binh Phuoc (Vietnam), CH-DL: Central Highlands-Dak Lak (Vietnam), CH-DN: Central Highlands-Dak Nong (Vietnam), CH-LD: Central Highlands-Lam Dong (Vietnam), NW-SL: Northwest-Son La (Vietnam), NW-DB: Northwest-Dien Bien (Vietnam), TL-CM: Thailand-Chiang Mai, MG-CB: Minas Gerais-Coimbra (Brazil), ES-VNM: Espírito Santo-Venda-Nova-do-Imigrante (Brazil), MG-CP: Minas Gerais- Capinopolis (Brazil), MG-SSP: Minas Gerais-Sao-Sebastiao-do-Paraiso (Brazil), MG-SO: Minas Gerais-Senhora-de-Oliveira (Brazil), Q-B: Quindío-Buenavista (Colombia), C-C: Caldas-Chinchiná (Colombia), V: Veracruz (Mexico), C: Chiapas (Mexico), O: Oaxaca (Mexico), P: Pueble (Mexico), C-P: Centro de Investigação de Ferrugens do Cafeeiro (Portugal).

Table 3.3. The number of haplotype, haplotype diversity and nucleotide diversity of *Hemileia vastatrix* populations

No.	**Population ID.**	**No of haplotype**	**Haplotype diversity (Hd)**	**Nucleotide diversity (Pi)**
1	SE-Dong Nai	7	1	0.01175
2	SE-Binh Phuoc	4	1	0.02651
3	CH-Daklak	5	1	0.0194
4	CH-DakNong	7	1	0.02527
5	CH-Lam Dong	7	1	0.01452
6	NW-Son La	7	1	0.00484
7	NW-Dien Bien	3	0.833	0.00288
8	TL-Chiang Mai	2	1	0.01153
9	Veracruz	1	0	0
10	Chiapas	3	1	0.00231
11	Oaxaca	2	1	0.00346
12	Puebla	2	1	0.00115
13	Quindío-Buenavista	5	0.933	0.00412
14	Caldas-Chinchiná	2	0.5	0.00175
15	MG-Coimbra	5	0.933	0.00268
16	ES-Venda Nova do Imigrante	4	1	0.0069
17	MG- Capinópolis	3	1	0.00613
18	MG-São Sebastião do Paraíso	3	1	0.0023
19	MG-Senhora de Oliveira	2	1	0.0046
20	CIFC	3	1	0.00231
	Overal	52	0.952	0.00779

Table 3.4. Tajima’s D value of *Hemileia vastatrix* populations

Country	**Population ID.**	**Statistics**		
		Sample size	**Tajima's D**	**Tajima's D p-value**
Vietnam	NW-SonLa	7	-0.07375	0.49800
	NW-DienBien	4	-0.38921	0.51100
	CH-DakLak	5	-0.24640	0.52400
	CH-DakNong	7	-0.58352	0.32800
	CH-Lam Dong	7	0.06951	0.53200
	SE-DongNai	7	-0.83847	0.22600
	SE-BinhPhuoc	4	-0.56337	0.40200
Thailand	Chiang Mai	2	0.00000	1.00000
Brazil	MG-Coimbra	6	0.63465	0.73200
	ES-Venda Nova do Imigrante	4	-0.84307	0.08700
	MG-Capinópolis	3	0.00000	1.00000
	MG-São Sebastião do Paraíso	3	0.00000	1.00000
	MG-Senhora de Oliveira	2	0.00000	1.00000
Colombia	Quindío-Buenavista	6	-0.27307	0.45000
	Caldas-Chinchiná	4	-0.79684	0.16400
Mexico	Veracruz	2	0.00000	1.00000
	Chiapas	3	0.00000	1.00000
	Oaxaca	2	0.00000	1.00000
	Puebla	2	0.00000	1.00000
Portugal	CIFC	3	0.00000	1.00000
	Mean	4.15	-0.19518	0.67270
	S.d.	1.899	0.37308	0.33319

NW: Northwest, CH: Central Highands, ST: Southeast, CIFC: Centro de Investigação de Ferrugens do Cafeeiro, MG: Minas Gerais, ES: Espírito Santo.

Table 3.5. Analysis of molecular variance (AMOVA) based on ITS sequences of the *Hemileia vastatrix* populations

Source of Variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	19	210.921	1.257	17
Within populations	63	374.055	5.937	83
Total	82	584.976	7.195	100
Fixation Index (F_{ST})			0.175	
P value			0.001	

d.f. Degrees of freedom

Table 3.6. Genetic differentiation (measured by Φ_{PT}) of *Hemileia vastatrix* between populations in Vietnam and other countries

Pairwise population Linearized PhiPT Values																				
NW-SL	**NW-DB**	**CH-DL**	**CH-DN**	**CH-LD**	**SE-DN**	**SE-BP**	**TL-CM**	**MG-C**	**ES-VNI**	**MG-CP**	**MG-SSP**	**MG-SO**	**Q-B**	**C-C**	**C-P**	**V**	**C**	**O**	**P**	
0.000																				**NW-SL**
-0.053	0.000																			**NW-DB**
0.089	0.054	0.000																		**CH-DL**
0.083	0.048	-0.081	0.000																	**CH-DN**
0.176	0.133	-0.050	-0.084	0.000																**CH-LD**
0.031	-0.037	0.052	0.058	0.074	0.000															**SE-DN**
0.133	0.046	0.048	0.089	0.155	0.079	0.000														**SE-BP**
0.151	0.217	-0.133	-0.111	-0.008	0.005	-0.123	0.000													**TL-CM**
0.450	0.557	0.358	0.280	0.393	0.362	0.336	0.647	0.000												**MG-C**
0.315	0.344	0.236	0.181	0.299	0.258	0.212	0.434	0.122	0.000											**ES-VNI**
0.326	0.384	0.186	0.131	0.267	0.235	0.142	0.410	0.090	-0.062	0.000										**MG-CP**
0.512	0.637	0.301	0.237	0.373	0.379	0.221	0.602	0.441	0.371	0.369	0.000									**MG-SSP**
0.272	0.319	0.099	0.036	0.202	0.168	0.008	0.339	0.407	-0.089	0.039	0.582	0.000								**MG-SO**
0.237	0.215	0.203	0.191	0.244	0.147	0.167	0.315	0.229	0.223	0.209	0.307	0.253	0.000							**Q-B**
0.422	0.444	0.313	0.263	0.372	0.345	0.262	0.495	0.593	0.429	0.488	0.613	0.482	0.352	0.000						**C-C**
0.239	0.224	0.089	0.083	0.184	0.078	-0.005	0.310	0.613	0.447	0.411	0.663	0.518	0.204	0.515	0.000					**C-P**
0.130	0.056	-0.040	-0.040	0.088	-0.055	-0.185	0.231	0.730	0.462	0.511	0.820	0.624	0.256	0.521	0.078	0.000				**V**
0.171	0.121	0.066	0.057	0.154	0.069	0.031	0.286	0.570	0.295	0.337	0.663	0.335	0.194	0.508	0.271	0.254	0.000			**C**
-0.023	0.055	-0.133	-0.131	-0.014	-0.073	-0.137	-0.053	0.567	0.248	0.288	0.638	0.258	0.157	0.455	0.311	0.371	-0.257	0.000		**O**
0.018	-0.189	-0.081	-0.079	0.038	-0.133	-0.206	0.174	0.687	0.391	0.453	0.789	0.508	0.182	0.465	0.102	0.000	0.102	0.204	0.000	**P**

SE-DN: Southeast-Dong Nai (Vietnam), SE-BP: Southeast-Binh Phuoc (Vietnam), CH-DL: Central Highlands-Dak Lak (Vietnam), CH-DN: Central Highlands-Dak Nong (Vietnam), CH-LD: Central Highlands-Lam Dong (Vietnam), NW-SL: Northwest-Son La (Vietnam), NW-DB: Northwest-Dien Bien (Vietnam), TL-CM: Thailand-Chiang Mai, MG-CB: Minas Gerais-Coimbra (Brazil), ES-VNM: Espírito Santo-Venda-Nova-do-Imigrante (Brazil), MG-CP: Minas Gerais- Capinopolis (Brazil), MG-SSP: Minas Gerais-Sao-Sebastiao-do-Paraiso (Brazil), MG-SO: Minas Gerais-Senhora-de-Oliveira (Brazil), Q-B: Quindío-Buenavista (Colombia), C-C: Caldas-Chinchiná (Colombia), V: Veracruz (Mexico), C: Chiapas (Mexico), O: Oaxaca (Mexico), P: Pueble (Mexico), C-P: Centro de Investigação de Ferrugens do Cafeeiro (Portugal).

CHAPTER 4

SEQUENCING ANALYSIS OF CYTOCHROME *b* GENE IN *HEMILEIA VASTATRIX* IN RELATION TO POTENTIAL QoI FUNGICIDE RESISTANCE

4.1 Introduction

Coffee leaf rust epidemics have occurred and cause severe yield loss affecting the livelihoods of more than 100 million people worldwide (Talhinhas et al., 2017). An effective strategy for controlling CLR is applying chemical fungicides. Three fungicide classes are among the most effective for controlling rust fungi: quinone outside inhibitors (QoIs), sterol demethylation inhibitors (DMIs), and succinate dehydrogenase inhibitors (SDHIs) (Oliver, 2014). Of these, QoIs (strobilurins) are especially important because of their effectiveness at controlling a broad range of fungal and oomycete pathogens (Bartlett et al., 2002). QoI fungicides inhibit mitochondrial respiration by binding the Qo site of the cytochrome bc1 enzyme complex, which leads to energy deficiency (Bartlett et al., 2002). To control CLR and increase yield, QoI fungicides are applied alone or in mixtures with other systemic fungicides, like triazoles (DMI group) (Souza et al., 2011; Honorato et al., 2015) or copper-based fungicides (copper oxychloride and hydroxide, cuprous oxide, and Bordeaux mixture) (Zambolim, 2016; Costa et al., 2019). In Vietnam, coffee farmers have been using the QoI fungicides azoxystrobin, pyraclostrobin, picoxystrobin, and trifloxystrobin to control CLR and other fungal diseases (Circular no. 19/2021/TT-BNNPTNT dated 28/12/2021).

Field resistance to QoIs has been reported in many pathogenic agents (FRAC 2020). Resistance to QoI fungicides is usually related to a single-point mutation in the *CYTB* gene that changes the encoded amino acid (Sierotzki et al., 2000). Three amino acid substitutions are associated with QoI resistance in fungi: from phenylalanine to leucine at position 129 (F129L), from glycine to arginine at position 137 (G137R), and from glycine to alanine at position 143 (G143A) (Fernández-Ortuño et al., 2007). In rust fungi, the first report on a change in QoI fungicide sensitivity in *Puccinia horiana*, the causative agent of chrysanthemum white rust, was from England (Cook, 2001), although the *CYTB* gene of *P. horiana* was characterized, no mutations were found (Grasso et al., 2006). Several years later, in Japan, curative spraying of azoxystrobin against *P. horiana* isolates which were subjected to QoI selection pressure in the field was ineffective. Sequencing the *CYTB* gene of these isolates revealed that they did not contain the F129L mutation (Matsuura, 2019). Nonetheless, the F129L mutation in the first exon of the *CYTB* gene of the Asian soybean rust fungus, *P. pachyrhizi*, can lead to QoI resistance (Klosowski, 2016; Muller et al., 2021). Likewise, QoI resistance associated with G143A has never been reported in rust fungi previously because the presence of intron type I directly after this codon leads to the substitution at this position being lethal (Grasso et al., 2006). QoI fungicides are being applied in coffee cultivation (Souza et al., 2011; Honorato et al., 2015; Zambolim, 2016; Costa et al., 2019), thus it is crucial to monitor QoI resistance of CLR to devise disease management strategies.

The *CYTB* gene structure is different among rust fungal genera. *Phakopsora pachyrhizi CYTB* gene includes two exons and an intron, while *P. horiana* harbors three exons and two introns in its *CYTB* gene (Grasso et al., 2006). However, the *H. vastatrix CYTB* gene contains four exons and three introns (Grasso et al., 2006). Until now, although there have been no reports on QoI resistance in *H. vastatrix*, it is epidemiologically important to detect such QoI resistance relating gene mutation rapidly. Therefore, in this study, we designed specific PCR primers to amplify fragments of exons 2 to 4 in the *H. vastatrix CYTB* gene and tested them using CLR specimens collected in Vietnam. The purpose of this work was to examine mutations in the *CYTB* gene to propose an appropriate control strategy for CLR disease.

4.2 Materials and methods

In this study, I examined 40 CLR specimens (Table 4.1) identified based on rDNA-ITS sequences (Chapter 3). These specimens were collected in three main coffee planting regions in Vietnam including Northwest, Southeast and Central Highlands. Genomic DNA was isolated from the CLR specimens using the modified thermal-shock method (Chapter 3). Briefly, *H. vastatrix* urediniospores were collected from rust-infected leaves and immersed in a tube containing 30 μL Buffer 1. After a 15 min incubation at 95°C, the tube was placed in a deep freezer at –80°C for 10 min. Finally, it was centrifuged and the supernatant was transferred to a new tube. A NanoDrop DS-11 spectrophotometer (DeNovix, Delaware, USA) was used to evaluate the quality and quantity of the extracted DNA.

The *H. vastatrix CYTB* gene is large, with four exons and three introns (Grasso et al., 2006). Therefore, to survey potential nucleotide substitutions in the *CYTB* gene of *H. vastatrix* isolates, the three PCR primer pairs CLRE2-FW and CLRE2-RV, CLRE3-FW and CLRE3-RV, CLRE4-FW and CLRE4-RV, were designed (Figure 4.1 and Table 4.2) to amplify fragments containing exons 2 to 4 based on a whole reference sequence of the *CYTB* gene using the Primer-BLAST tool at NCBI (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi). The reference sequence of *H. vastatrix CYTB* gene fragment, including full sequence of exons 2, 3 and fragments of exons 1, 4 (Grasso et al., 2006), was obtained from NCBI (https://www.ncbi.nlm.nih.gov/nuccore) with GenBank accession number DQ209282. The primer melting temperature (Tm) ranged from 57°C to 63°C; the maximum Tm difference between two primers in a pair was 3°C and the GC content was 40% to 60%. The primer pairs with the smallest Tm difference were chosen.

Gradient PCR was conducted to determine the optimal annealing temperature for each primer pair. A 25 μL reaction mixture was prepared that contained 1–2 μL DNA template (30–50 ng/μL), 0.2 μM each primer (2.5 μL), 12.5 μL Gene RED PCR Mix Plus (Nippon Gene, Tokyo, Japan), and 5.5–6.5 μL autoclaved distilled water. The PCR process consisted of an initial 5 min denaturation at 95°C, 30 cycles of 95°C for 30 s, annealing from 48°C to 58°C for 45 s, and 72°C for 1 min, followed by a final 8 min extension at 72°C. The amplicon was confirmed by electrophoresis in 1.2% agarose gels. After being purified using the FastGene™ Gel/PCR extraction kit (Nippon Genetics, Tokyo, Japan), the PCR products were consigned to Eurofins Genomics (Tokyo, Japan) for sequencing. ATGC software (GENETYX, Tokyo, Japan) was used to assemble the obtained sequences. Finally, ClustalX ver. 2.0 (Larkin et al., 2007) was used to identify the nucleotide substitutions in exons 2 to 4.

4.3 Results

The primers CLRE2-FW and CLRE2-RV amplified a 280 bp fragment that contains site F129 (exon 2). The primers CLRE3-FW and CLRE3-RV amplified a 475 bp fragment containing sites G137 and G143 (exon 3). The primers CLRE4-FW and CLRE4-RV amplified a 691 bp fragment including a part of exon 4 (Table 4.2, Figure 4.1). After gradient PCR, the optimal annealing temperature for the primer pairs CLRE2-FW and CLRE2-RV, and CLRE4-FW and CLRE4-RV was determined to be 57°C, while that for the primer pair CLRE3-FW and CLRE3-RV was 53°C (Table 4.2). After aligning the nucleotide sequences, amino acids were deduced and compared to those of *H. vastatrix CYTB* in GenBank (ABB54711). No mutations at positions F129, G137, and G143 were detected in any isolates analyzed (Figures 4.2A, 4.2B). Nonetheless, all isolates had third-base nucleotide substitutions in codons T145 and T148 compared with reference sequence DQ209282 (Figure 4.2C). The thymine (T) at both codons in the DQ209282 sequence was replaced with adenine (A) in all Vietnamese isolates, although the deduced amino acid remained unchanged (threonine).

4.4 Discussion and conclusion

The *CYTB* gene structure of *H. vastatrix* was revealed with three sandwich introns (Grasso et al., 2006). In this study, the primer pairs CLRE3-FW, CLRE3-RV, CLRE4-FW and CLRE4-RV were designed based on sequences of introns 2 and 3 and exon 4 (Figure 4.1). An intron between exons 1 and 2 (intron 1) of the *CYTB* gene of *H. vastatrix* (Figure 4.1) was considered for the designed forward primer CLRE2-FW. However, I could not design a forward primer in the target intron 1 because no sequence data were available. No forward primer located in exon 1 amplified a product. Finally, the primer CLRE2-FW was designed which included 10 nucleotides in exon 1 and 15 nucleotides in exon 2, while CLRE2-RV covered part of intron 2 (Figure 4.1). The success of amplifying the exon 2 fragment using these primers questions the existence of intron 1 in the *H. vastatrix CYTB* gene. The presence of this intron needs to be confirmed because the frequent loss of introns was reported in the fungal *CYTB* gene (Yin et al., 2012).

The main mechanism of QoI resistance in fungal pathogens is the nucleotide substitutions in some positions including 129, 137, and 143 in their *CYTB* gene (Fernández-Ortuño et al., 2008). Howbeit, in this study, no mutations were detected at these positions in 40 *H. vastatrix* isolates in Vietnam although some host plants were exposed to QoI fungicides several times (data not shown). The G143A mutation has not been detected in the rust fungus and the presence of type I intron immediately after this codon might be one of the reasons (Grasso et al., 2006; Fernández-Ortuño et al., 2008). Similar to *H. vastatrix*, F129L mutation was not observed in *P. horiana* isolates in Japan (Matsuura, 2019). Nonetheless, the F129L mutation associated with QoI resistance was detected in 51% of the *P. pachyrhizi* populations in Brazil in 2013–2014 (Klosowski, 2016). Subsequently, these F129L mutant isolates increased rapidly in Brazil (Müller et al., 2021). In soybean cultivation, rust-resistant cultivar is yet planted, thus, chemical sprays including triazoles and strobilurins have been applied two or three times per season as an effective method to control rust disease in Brazil (Zambolim et al., 2022). The periodic QoIs applications could be a reason for developing resistance in soybean rust fungal populations in Brazil. As QoIs are being applied to control CLR disease, it is important to monitor fungicide resistance in the rust fungal populations subjected to fungicide selection pressure in the field to propose rational and timely management strategy for disease control. Meanwhile, to avoid QoI resistance, farmers need to be guided to save QoI fungicide applications.

Even though F129L, G137R, and G143A mutations were not detected in 40 *H. vastatrix* isolates in Vietnam, it cannot be excluded that other mechanisms are involved in resistance in this fungus. Indeed, two *Colletotrichum nymphaeae* isolates causing strawberry anthracnose fruit rot showed resistance to azoxystrobin without mutations in the *CTYB* gene (Chechi et al., 2020). In the case of *P. horiana*, F129L mutation was not detected in some isolates collected in 2017 (Matsuura, 2019). However, several years later, the substitutions L299F, L275F + L299F, and N256S + L299F in the third exon were identified in some isolates possessing azoxystrobin resistance (Matsuzaki et al., 2020; 2021). In this study, the amplicon using primers CLRE4-FW and CLRE4-RV did not cover positions L275F, L299F, and N256S. We need to analyze these positions and others in exon 4 of *H. vastatrix* in the future. In the meantime, field survey and bioassay will be necessary to be conducted in order to know whether performance of QoI fungicides is still maintained or not.

- ◆ The content of this chapter was accepted to be published in Tropical Plant Pathology Journal-Special Issue: Fungicide Resistance and Management Strategies with the title “Sequence analysis of cytochrome b gene in Vietnamese isolates of *Hemileia vastatrix*, the causal agent of coffee leaf rust, in relation to potential QoI fungicide resistance”.

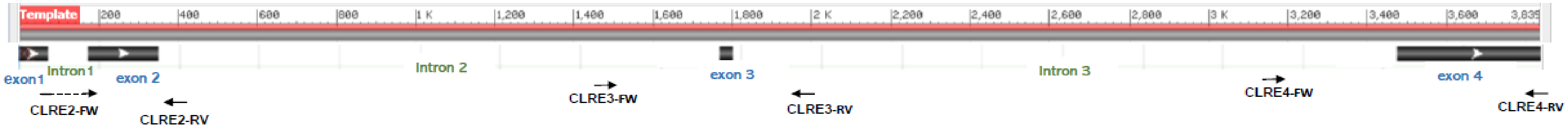


Figure 4.1. The positions of the exons of the *Hemileia vastatrix CYTB* gene and designed primers.

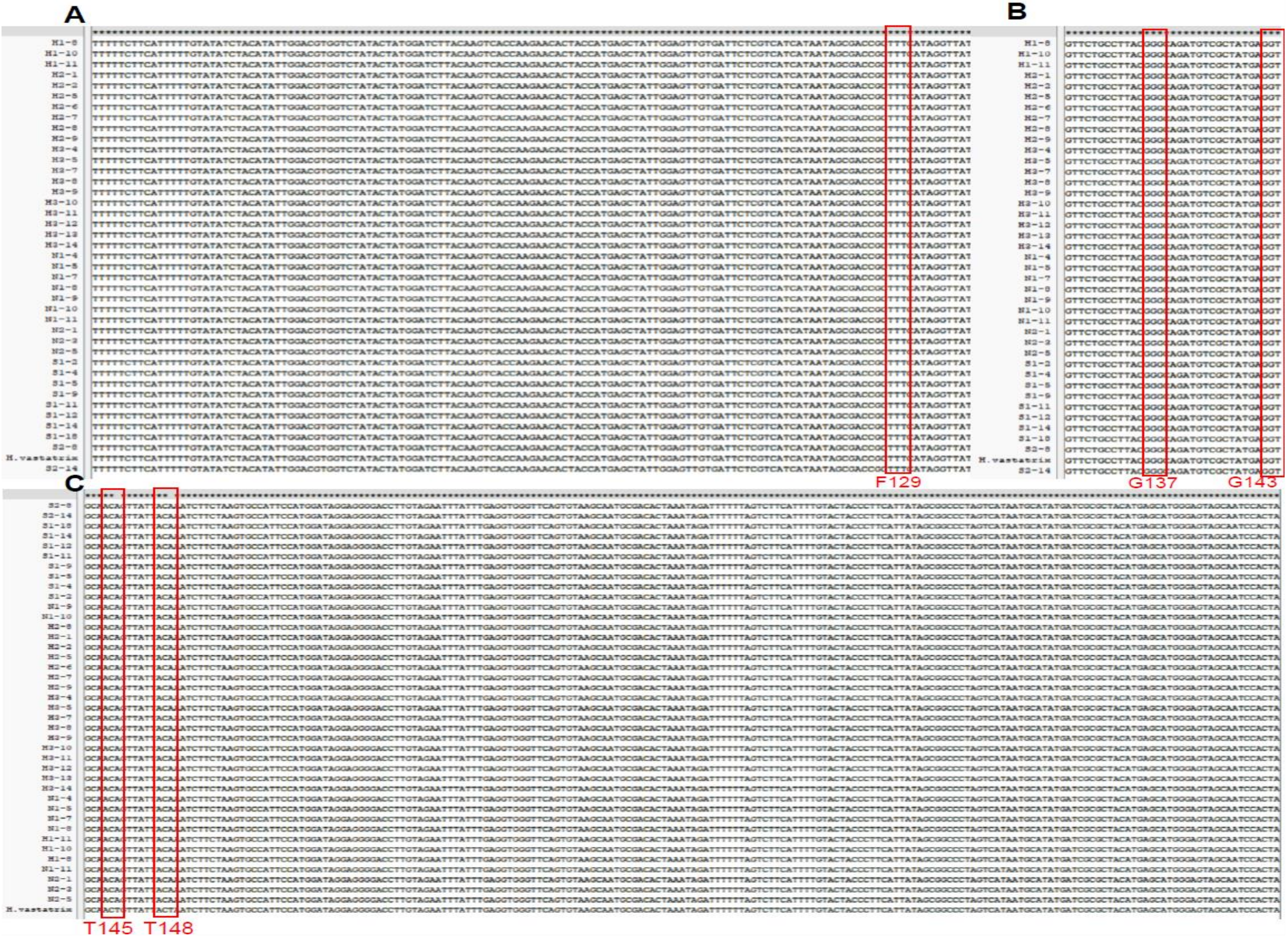


Figure 4.2. Sequence alignment of *CYTB* gene fragments of *Hemileia vastatrix* collected in Vietnam and GenBank isolate DQ209282. (A) Sequence alignment of exon 2 fragment, (B) Sequence alignment of exon 3 fragment, and (C) Sequence alignment of exon 4 fragment.

Table 4.1. Geographical location of specimens using for checking QoI fungicide resistance

Area	**Province**	**Latitude**	**Longitude**	**Specimen ID.**	**TSH ID.**	**Coffee host species**
The Southeast	Dong Nai	10.9837	107.2536	S1_2	TSH-R30018	*Coffea canephora*
		10.9775	107.2700	S1_4	TSH-R30021	
		10.9467	107.3131	S1_5	TSH-R30022	
		10.8579	107.1626	S1_9	TSH-R30024	
		10.9729	107.0765	S1_11	TSH-R30028	
		11.4090	107.3903	S1_12	TSH-R30030	
		11.2424	107.3906	S1_14	TSH-R30031	
		10.7703	107.2639	S1_18	TSH-R30036	
	Binh Phuoc	11.7416	107.1452	S2_8	TSH-R30044	
		11.8008	107.2175	S2_14	TSH-R30045	
The Central Highlands	Dak Lak	12.5533	108.1810	H1_8	TSH-R30050	
		12.6214	108.1920	H1_10	TSH-R30052	
		12.6674	108.1960	H1-11	TSH-R30055	
	Dak Nong	12.5737	107.8304	H2_1	TSH-R30062	
		12.5677	107.8235	H2_2	TSH-R30063	
		12.4719	107.6916	H2_5	TSH-R30069	
		12.1819	107.6379	H2_6	TSH-R30071	
		12.1555	107.6402	H2_7	TSH-R30073	
		12.0618	107.6771	H2_8	TSH-R30080	
		11.9840	107.7004	H2_9	TSH-R30085	
	Lam Dong	11.7108	107.9789	H3_4	TSH-R30089	
		11.6855	107.9926	H3_5	TSH-R30090	
		11.7637	108.1193	H3_7	TSH-R30094	
		11.7764	108.1537	H3_8	TSH-R30098	
		11.7848	108.1842	H3_9	TSH-R30101	
		11.7626	108.3454	H3_10	TSH-R30104	
		11.6236	108.2564	H3_11	TSH-R30107	
		11.6283	108.1409	H3_12	TSH-R30109	
		11.6095	108.1197	H3_13	TSH-R30116	
		11.5590	107.8309	H3_14	TSH-R30118	
The Northwest	Son La	21.2389	103.8599	N1_4	TSH-R30120	*Coffea arabica var. Catimor*
		21.2593	103.8687	N1_5	TSH-R30125	
		21.3538	103.8124	N1_7	TSH-R30129	

(Continued)

Area	**Province**	**Latitude**	**Longitude**	**Specimen ID.**	**TSH ID.**	**Coffee host species**
The Northwest	Son La	21.5177	103.6437	N1_8	TSH-R30133	*Coffea arabica var. Catimor*
		21.5244	103.6375	N1_9	TSH-R30137	
		21.5063	103.6534	N1_10	TSH-R30145	
		21.5387	103.6202	N1_11	TSH-R30148	
	Dien Bien	21.5653	103.4651	N2_1	TSH-R30155	
		21.555	103.2813	N2_3	TSH-R30159	
		21.5039	103.1969	N2_5	TSH-R30167	

Table 4.2. Information on the primer pairs designed to amplify exon fragments of the *Hemileia vastatrix* cytochrome *b* gene

Primer name	**Sequence (5'->3')**	**Tm (°C)**	**Percentage of GC (%)**	**Optimal annealing temperature (°C)**	**Product length (bp)**
CLRE2-FW	TGTTGAGTATGGATGACTTATCAGG	58.18	40	57	280
CLRE2-RV	GTAGCTTTCGGCTGGCTCTA	59.54	55		
CLRE3-FW	TCGGTTCTATGCAGGATCGC	59.97	55	53	475
CLRE3-RV	TGATCATCGTCGTTCAGCGT	59.83	50		
CLRE4-FW	ACAGTCGATGTTCCCGTAGC	59.83	55	57	691
CLRE4-RV	TTCAGGAACGATCGACGGTG	60.11	55		

FW: Forward primer, RV: Reverse primer.

CHAPTER 5

GENERAL DISCUSSION AND RECOMMENDATIONS

5.1 General discussion

5.1.1 The current situation of coffee leaf rust disease in Vietnam

The size of urediniospores (length, width, and cell wall thickness) from all specimens fell in a small contiguous range, i.e., similar. Of forty-two species of *Hemileia* known, only two species including *H. vastatrix* and *H. coffeicola* could infect coffee so far (Waller, 1982). The urediniospores sizes were the same as *H. vastatrix* and *H. coffeicola* as described by Ritschel et al. (2005). Nonetheless, the typical morphology on which we can safely relay to distinguish between *H. vastatrix* and *H. coffeicola* is the appearance of the rigid spine ridge dividing echinulate and smooth areas on the surface of urediniospores. The former species has this ridge (Ritschel et al., 2005). Therefore, based on the dimension as well as the morphology of urediniospores of 41 specimens collected in three main coffee areas in Vietnam, the pathogenic agent causing coffee leaf rust in coffee plantations in Vietnam was identified as *H. vastatrix*. Indeed, previous studies pointed out that *H. coffeicola* was found to infect robusta coffee in wet environments in West and Central Africa while *H. vastatrix* is the major agent leading to CLR disease in the world (Waller, 1982). Until now, there was no information about the appearance of *H. coffeicola* in other areas of the world (Talhinhas et al., 2017). *Hemileia vastatrix* was also first detected in northern Vietnam in 1890 (Hennings, 1895). Therefore, according to these information, the occurrence of only *H. vastatrix* in Vietnam is rationable. The result in this study is in agreement with previous studies when *H. vastatrix* related to CLR disease outbreak in Peru (Cinthia et al., 2017), Brazil (Maia et al., 2013), and other countries in recent years (Talhinhas et al., 2017).

The survey of 85 plantations/small-holder farms across three main coffee planting regions (Northwest, Central Highlands, and Southeast) revealed CLR disease existed in 41 visited plots from South to North, and from low altitudes to high mountainous areas. In which, the disease incidence and severity were lowest in the Southeast. The Southeast used to be one of the main coffee-planting regions in Vietnam. However, the majority of coffee farms were converted to black pepper, rubber tree, and fruit trees in recent years. The remaining coffee farms are small acreage and scattered distance from each other. Moreover, coffee is usually intercropped with black pepper in the Southeast. These could be the reasons for low disease incidence in the Southeast because host presence and density are important factors affecting plant disease incidence (Burdon and Chilvers, 1982). Black pepper is a cash crop in Vietnam, therefore, when intercropped with coffee, farmers apply an appropriate nutritional regime for coffee. This might be a rational excuse for the lowest disease severity in this region.

This study showed that the disease incidence in Central Highlands was higher than that in the Southeast. In the coffee cultivation map of Vietnam, Central Highlands, which is the main coffee area in Vietnam, includes both large plantations and small-holder farms. Historically, robusta coffee has been cultivated in this region as a monocrop with high yields (Le et al., 2021). Moreover, our survey showed that these plantations/farms are gathered to form large coffee-growing areas (data not shown). The appearance of the coffee host in large continuous areas might be the reason for the higher disease incidence in this region (Hamelin et al., 2000; Keiper et al., 2006; Cabral et al., 2016) . With the experience of coffee cultivation for years, farmers apply appropriate water and fertilizers regism for coffee growth. In addition, *C. canephora*, which is a natural resistance to *H. vastatrix* (McCook, 2006), covers almost the whole region. Thus, the disease severity in Central Highlands is moderate.

The Northwest had the highest disease incidence and severity among the three survey regions. In the Northwest, most plantations are located in the high mountainous areas while some farms with large acreage are scatted in lower areas. This topographic feature leads to natural cultural practice in coffee cultivation with minimum involvement of human beings. Besides, *C. arabica cv. Catimor*, which is a descendant of *C. arabica*, is the primary host cultivar in this region. Being different from *C. canephora*, *C. arabica* is highly susceptible to *H. vastatrix* (McCook, 2006; Talhinhas et al., 2017). Historically, *H. vastatrix* was first detected on *C. arabica* in northern Vietnam. Although *C. arabica* was cut down for Catimor coffee in northern Vietnam (Anonymous, 2019), urediniospores of the rust fungus could still exist by parasitizing coffee plants in other plantations (the replacement of coffee varieties was not synchronized well and did not occur at the same time in all plantations of a region) or wild coffee plants and infected Catimor coffee, making this coffee variety susceptible to coffee rust. The presence of the susceptible coffee host with poor cultural practices leads to this result.

The field survey in this study showed that CLR fungus infected and caused disease in both *C. canephora* and *C. arabica cv. Catimor*. These species show some resistance to rust fungus although the levels of resistance vary between them. The first species is naturally resistant to rust fungus while the second is the hybrid. The rust disease incidence and severity were recorded as high levels in both *C. canephora* lines and *C. arabica cv. Catimor* derivatives in Vietnam remind us to conduct more research in disease-resistance breeding. Finding mechanism and process in losing resistance in coffee varieties are also necessary. The resistance breakdown could be the reasons for new race emergence or the development of rust populations (Talhinhas et al., 2017). Random mutations in the genome followed by the natural or anthropogennic selection processes or new combinations of genes generated through sexual or somatic recombinations could make the population more diverse or develop new races (Ramírez-Camejo et al., 2022). On the other hand, the reproduction mode of *H. vastatrix* was revealed by population genetics analyses by eight simple sequence repeat markers (SSRs) using isolates from the Old World (Congo, Cameroon, and India) and New World (Latin American and Caribbean) between 2014 – 2019. The results revealed the asexual reproduction of *H. vastatrix* across all areas (Ramírez-Camejo et al., 2022). Moreover, urediniospores of *H. vastatrix* contain two different haploid nuclei. It is clear that sexual reproduction generates new gene combinations, however, in rust fungi, cryptosexuality (hidden sexual reproduction within asexual spore) in *H. vastatrix* (Carvalho et al., 2011), somatic hybridization in *Puccinia recondita f. sp. tritici* (Prt) (Park et al., 1999) and *Puccinia graminis f. sp. tritici* (Pgt) (Ug99 strain) (Li et al., 2019) also lead to novel virulence types emergence. Somatic hybridization occurred between many rust species. This phenomenon involves the mixture of dikaryotic hyphae, nuclear exchange, or whole chromosome exchange (Park and Wellings, 2012).

In the coffee season 2018-2019 in Vietnam, abnormal rains for several months in flowering time created a low yield in coffee production. In addition, coffee farmers experienced low prices during this season. These factors resulted in poor investment in coffee production in the next year. Moreover, the price of fertilizers increased at the end of the year 2018 also contributed to the limited investment in coffee plantations. In Central Highlands, some farmers even converted from coffee to other cash fruit trees such as avocado and durian or abandoned coffee plantations. The poor cultural practice on coffee plantations leading to the high disease severity was recorded in this study. Indeed, the survey in this study revealed high levels of rust severity found in some abandoned coffee plantations of the Central Highlands. Unfavorable climatic conditions and cultural practices also promote disease outbreaks in Central America and Colombia from 1987 to 2003 (Avelino et al., 2015) and in Colombia during the period from 2008-2011 (Cristancho et al., 2012). Therefore, it is important to have more research on factors including climate and cultural practices on the coffee leaf rust disease.

5.1.2 The current population of *Hemileia vastatrix* in Vietnam

The real picture of the *H. vastatrix* populations in Vietnam is now disclosed through genetic diversity and population structure analyses in this study. Firstly, the genetic differentiation of *H. vastatrix* in Vietnam was low but the haplotype diversity was high. Secondly, *H. vastatrix* population in Vietnam was unstructured based on sampling sites or coffee host species. Most haplotypes in Vietnam are unique. Mutations in the *H. vastatrix* populations of Vietnam are likely to be spontaneous without being affected by geographical factors or the host's resistance genes. This result is in accordance with a study on *H. vastatrix* population in Peru. Analysis of the ITS sequence of *H. vastatrix* isolates in Peru also indicated a large undifferentiated population with high haplotype diversity and low nucleotide diversity. Furthermore, the majority of haplotypes in Peru were singleton (Cinthia et al., 2017; Cinthia et al., 2021). Likewise, analysis of 112 *H. vastatrix* isolates from Brazil by using amplified fragment length polymorphisms (AFLP) showed a low level of genotypic diversity, in which 90% of the genetic variance occurred among isolates within the population. *Hemileia vastatrix* population in Brazil was also unstructured with respect to either geographic region or coffee host (Cabral et al., 2016). Contrary to the result of the genetically unstructured populations at the domestic level isolates (Vietnam, Peru, and Brazil), other studies that applied restriction site-associated DNA sequencing (RADseq) on isolates across continents revealed a structured population of *H. vastatrix* marked with host specialization. In which, one lineage infected tetraploid coffee including *C. arabica* and interspecific hybrids, whereas the other two lineages could infect diploid coffee species (Silva et al., 2018; Rodrigues et al., 2022). There are several differences that lead to these contrary results. The first variation between these studies relates to the methodology which was conducted in population genetic analyses. This study and the other in Peru focused only on the rDNA-ITS region, these studies analyzed nucleotide differentiation in the whole genome, while Brazil's study applied the AFLP method. The second difference concerns the origin of the samples. These studies used samples collected in many countries across the continent, while my study, that of Peru and Brazil, only collected samples in the same country. Lastly, while Brazil's, Peru's, and my studies used isolates collected during the same period, the samples used in these studies were collected at different time periods from 1953 to 2013. Whether sample collections in the studies at the national or intercontinental level, the reality is that *H. vastatrix* populations have low genetic variation. It is of paramount importance to find the reason for this phenomenon.

It is incontrovertible that sexual reproduction leads to genetic diversity in the population (Mboup et al., 2009; Saleh et al., 2014). However, most studies on CLR fungus indicated low genetic differentiation in *H. vastatrix* population. Recently, a study gathered global scale *H. vastatrix* isolates from Central Africa, Southeast Asia, the Caribbean, and South and Central America using eight simple sequence repeat markers (SSRs) emphasized the clonal reproduction in *H. vastatrix* population. Moreover, the loss of host resistance could be the result of the process of directional selection rather than recombination in *H. vastatrix* population (Ramírez-Camejo et al., 2022). Another study respected that migration might contribute to poor genetic variation in clonal populations (Cabral et al., 2016). Indeed, shared haplotypes between Northwest, Southeast, and Central Highlands in Vietnam and the South and Central American countries in my study is a solid proof for this hypothesis. These haplotypes might be brought to Northwest Vietnam and then spread to other regions. Moreover, adjacent genetic distance in PCoA analysis and close genetic relationships in geophylogenetic tree of *H. vastatrix* isolates in Vietnam also support the migration events of this fungus in the past. The migration of this fungus was evidenced in Brazil's with the high number of migrants found in the states of Bahia and Espırito Santo (Cabral et al., 2016). Brazil is the first largest coffee producer worldwide and the presence of vast extended areas of coffee supported the expansion of this fungus in the whole country (Cabral et al., 2016). Similarly, Vietnam has large continuous coffee cultivation areas as well in the Central Highlands and the Northwest, it is, therefore, the migration of this fungus might occur leading to the disease incidence and severity in these regions higher. Urediniospores of *H. vastatrix* can survive for several weeks in the wild (McCook, 2006) and are easily spread by the wind, rain, insects, animals, and humans (Avelino et al., 2015). Long-distance dispersion of this fungus with the aid of wind happened in Brazil (Bowden et al., 1971) and other areas (McCook, 2006). However, the unstructured population of *H. vastatrix* in Vietnam might be affected by human activities rather than wind. Vietnam's wind is typical by monsoon seasons including northeast monsoon and southwest monsoon. Therefore, the expansion of this fungus from northern to southern Vietnam might derive from anthropogenic actions. For example, the gathering of coffee pickers in the fields during harvesting season can promote urediniospores being detached and spreading. When they move to other regions, attached urediniospores might be brought along with their belongings. Historically, human activities contributed to disease dispersion through the continental and created severe epidemics affecting global coffee production (McCook, 2006). The last explanation for the current situation of *H. vastatrix* population in Vietnam is the infected seedlings. The fact that in Vietnam, some growers propagated seedlings for their coffee cultivation expansion or commercial to local farmers without being fully educated about CLR disease leading to the accidentally spread of the disease (data not shown).

Cytochrome b gene sequencing revealed undifferentiation between *H. vastatrix* isolates in Vietnam. This means that there were no nucleotide substitutions at some potential positions associated with QoI resistance. There are several possibilities for undetected mutations in the *CYTB* gene of *H. vastatrix* in this study. First, Vietnamese farmers might not apply QoIs in coffee farms. This can not occur because QoI fungicides azoxystrobin, pyraclostrobin, picoxystrobin, and trifloxystrobin were commercial in Vietnam to control CLR and other fungal diseases (Circular no. 19/2021/TT-BNNPTNT dated 28/12/2021). In southern Vietnam, farmers usually intercrop coffee with black pepper, while coffee is cultivated as a mono-crop in large areas of northern Vietnam. Southern farmers sometimes apply fungicides including biological and chemical compounds to control root rot, charcoal rot, anthracnose, and so forth on black pepper, while coffee is grown naturally in high mountains in the northwest and some areas of the Central Highlands (data not shown). Second, other mechanisms rather than F129L, G137R, and G143A mutations might involve in resistance in this fungus. For example, *Colletotrichum nymphaeae*, which causes strawberry anthracnose fruit rot, showed resistance to azoxystrobin without mutations in the *CTYB* gene (Chechi et al., 2020). Another example in *P. horiana*, chrysanthemum rust, substitutions L299F, L275F + L299F, and N256S + L299F in the third exon were identified in some isolates possessing azoxystrobin resistance (Matsuzaki et al., 2020; 2021). Lastly, these *H. vasatrtix* isolates might still be sensitive to QoI fungicides. Historically, QoI resistance arose in chrysanthemum rust in England (Cook, 2001), Japan (Matsuura, 2019; Matsuzaki et al., 2020; 2021) and soybean rust (Klosowski, 2016; Muller et al., 2021). In soybean cultivation, the rust-resistant cultivar is yet planted, thus, chemical sprays including triazoles and strobilurins have been applied two or three times per season as an effective method to control rust disease in Brazil (Zambolim et al., 2022). In Japan, some farmers applied QoIs several times per season to control chrysanthemum rust (Matsuura, 2019). These periodic QoIs applications could be a reason for developing resistance in rust populations in Brazil and Japan. Meanwhile, coffee is a long-term crop and is usually cultivated in large and high mountainous areas, therefore, the treatment by QoI fungicides is infrequently. The infrequent application of fungicides in coffee farms might be a reason for the remaining sensitivity of QoIs in *H. vastatrix* population in Vietnam. Anyhow, the evolution of QoI fungicide resistance in *H. vastatrix* subjected to fungicide selection pressure in the field needs to be monitored carefully to best enable farmers to control CLR disease. In the meantime, field survey and bioassay will be necessary to be conducted in order to know whether the performance of QoI fungicides is still maintained or not. The fact that QoI fungicides effectively controlled CLR alone or combined with other agents (Souza et al., 2011; Honorato et al., 2015; Zambolim, 2016; Costa et al., 2019). From the result of this study, Vietnamese farmers can continue to combine QoIs with DMIs or copper-based fungicides to control CLR. However, to avoid QoI resistance, farmers need to be guided to save QoI fungicide applications.

5.2 Recommendations

Coffee leaf rust fungus infected not only *C. canephora* but also *C. arabica cv. Catimor* in most cutivated regions Vietnam. The former species is naturally resistant to CLR while the latter is a hybrid between HdT and Caturra coffee. The successful infection of rust on these varieties reminds us to conduct more research on factors including climatic conditions and cultural practices that can promote disease outbreaks. Likewise, generating new rust-resistant varieties is necessary as many reports indicated the re-outbreaks of CLR even in areas cultivating current hybrid varieties. Because of this primary importance, the breeding of resistant cultivars should be carried out through collaborations between national research institutions and international research and development institutions with support from governmental and international funding organizations.

In this research, Northwest was determined as the first establishment whereas other sources of *H. vasatrix* emerged in southern Vietnam. In addition, *H. vastatrix* from Mexico and Brazil was highlighted as having a close genetic relationship with those from these regions. However, I cannot ascertain that the *H. vastatrix* population in Vietnam spread from these countries owing to the first rust detection in Vietnam in 1890, while the first confirmation of rust establishment in Brazil and Mexico was in the 1970s and 1980s, respectively. The sources of *H. vastatrix* and migration routes into Vietnam are the big questions. In order to answer these inquiries, I propose to collect CLR samples from Old World, New World, French colonies, and Southeast Asian coffee growing regions as well as identify coffee cultivars in each collection site in Vietnam. Moreover, surveying, and monitoring races and disease incidences in Southeast Asia need to be conducted in international collaborations (including CIFC).

The accelerated developmental progress of QoIs resistance was determined in other rust fungi. Consequently, research on QoIs resistance of CLR needs to be implemented periodically. Specimens of CLR fungus should be collected for this research and for other purposes. This study proposes PCR approach using three newly designed primer pairs for examining the progress of nucleotide mutations on the *CYTB* gene of *H. vastatrix* for these specimens to timely propose a proper approach to control CLR. Moreover, the other mechanisms of QoI resistance of *H. vastatrix* need to be considered other than F129L, G137R, and G143A mutations. In this study, the amplified product using primers CLRE4-FW and CLRE4-RV did not cover positions L275F, L299F, and N256S. I suggest to analyze these positions and others in exon 4 of *H. vastatrix* in the future. In the meantime, it is necessary to conduct field survey and bioassay in order to know whether performance of QoI fungicides is still maintained or not.

The survey in this study revealed the occurrence as well as the current population of *H. vastatrix* in Vietnam. Based on these results and the disease triangle between susceptible host, conducive environment, and pathogenic agents, I propose some strategies for sustainable coffee production in Vietnam. To manage CLR disease effectively, the Vietnamese Government and local farmers need to take responsibility. The Government should have a strict plant quarantine program that considers clean seedlings and quarantine belts as the center of the action. Especially, in this study, the development of CLR disease was raised as a result of joining neighbor farms to create large extended coffee areas in Vietnam. Thus, to prevent disease outbreaks in the future, Vietnam needs to plan intercropped regions to interrupt the presence of the host. Moreover, agricultural extension officers should combine with local farmers to visit coffee planting regions frequently to monitor the occurrence of CLR and other diseases. The purpose of this work is to propose a timely disease-managing program when it has just occurred. In this study, moderate and high disease incidence and severity were recorded on *C. canephora* lines and *C. arabica cv. Catimor* derivatives which are currently rust-resistant varieties, respectively. Therefore, it is time for Vietnam to conduct more research on creating quality and rust-resistance breedings. Besides investing in the breeding program, other research related to pathogenic agents and control disease also needs to be supported. Basic researches on coffee cultivation management with respect to planting methods, e.g., mixed planting (selection of crops to be mixed), shade culture (kinds and density of shade trees), fertilization, application of fungicides and pesticides. Depending on topography and climatology, different patterns of multiple cropping systems with coffee and/or agroforestry would be developed in the three coffee-growing regions. In addition, free distribution of clean seeds or seedlings of new resistant cultivars to smallholder farmers is strongly preferred as a practical approach to supporting them, and controlling Vietnamese coffee quality together with CLR disease in the future. Lastly, an equally important thing in this disease control strategy is the education of farmers and related workforces. They need to accumulate knowledge on CLR disease as well as cultural practice techniques on coffee production.

It is of paramount importance for farmers to acquire knowledge of multiple cropping systems, agroforestry, planting techniques together with coffee diseases. Smallholder farmers need to apply either multiple cropping systems or agroforestry because they are ecology-friendly and secure alternative ways to obtain food and invest in sustainable coffee cultivation on small lands. In this study, the high disease severity in some abandoned plantations/farms in the Central Highlands indicates the importance of cultural practices in preventing CLR. Therefore, farmers need to provide rational fertilizers and water for coffee to enhance the defense system of the host plant. On the other hand, the unstructured population with low genetic variation exposed in this study showed possible migration events with the aid of wind, human activities, or unknown factors in Vietnam. The fact that urediniospores of *H. vastatrix* were easily detached from the infected leaves and survival several days before attaching a new host. They can germinate and infect the host successfully or not depending on the proper environment. In order to prevent urediniospore detaching and germinating, farmers should apply drip irrigation rather than watering or sprinkler for coffee. My field survey also revealed the higher disease severity in plantations/farms that have high tree density and little ventilation (data not shown). Coffee growers need to prune and eliminate weeds frequently to create proper air circulation as well as prevent moisture accumulation in farms. Moreover, it is crucial to visit coffee plantations/farms frequently to conduct prompt treatments when diseases occur. Once detecting symptoms of CLR disease in the plantations/farms, producers have to separate infected leaves/twigs and burn them immediately. Meanwhile, infected clothes and tools need to be cleaned before reusing them to work in the field. The treatment by biological and chemical compounds requires to be conducted to cure coffee plants of CLR disease after that. Especially, farmers are required to combine QoIs with other fungicides to prevent resistance in *H. vastatrix* population.

Acknowledgments

I would like to give my first and foremost gratitude to my supervisor Associate Professor Izumi Okane, Laboratory of Plant Parasitic Mycology, Graduate School of Life and Environmental Sciences, the University of Tsukuba for supporting and guiding me on fungal taxonomy work and study of coffee leaf rust in Vietnam, and throughout my stay in the University of Tsukuba, Japan.

Second, I am greatly indebted to the Ministry of Education, Culture, Sports, Sciences, Technology of Japan and the Board of Directors of the Biotechnology Center of Ho Chi Minh City for financial support and for creating conditions as well as arranging other co-workers to help me collect coffee leaf rust samples in Vietnam.

Third, I wish to express my big appreciation to Emeritus Professor Yoshitaka Ono, College of Education, Ibaraki University; Associate Professor Yoshiaki Tsuda, Sugadaira Research Station, Mountain Research Center, University of Tsukuba; Doctor Hideo Ishii, National Institute for Agro-Environmental Sciences, Japan; and Professor Yuichi Yamaoka, Laboratory of Plant Parasitic Mycology, Graduate School of Life and Environmental Sciences, the University of Tsukuba for helpful pieces of knowledge on analyses as well as for giving comments on my papers. I am truly happy with their kind assistance.

Four, I am deeply indebted to Professor Watanabe Kazuo, Graduate School of Life and Environmental Sciences, the University of Tsukuba for his great guidance and for helping me joining in the University of Tsukuba.

Five, I would like to thank Dinh Anh Hoa, Nguyen Thi Thuy Duong, Tran Chi Hieu, and Nguyen Thi Ngoc Bich for their assistance in collecting samples in Vietnam.

Next, I am indebted to my friends in the Laboratory of Plant Parasitic Mycology, at the University of Tsukuba for their kind support during my stay in Japan.

I am sincerely thankful to all the people who were involved and give great advice throughout this study. I am also grateful to all staff in the Agricultural sciences for their kind support during my stay at the University of Tsukuba.

Lastly, I am grateful to all my family members for their precious support in my life.

References

1. Aime, M. C. and McTaggart A. R. (2020). A higher-rank classification for rust fungi, with notes on genera. *Fungal Systematics and Evolution* 7, 21–47. doi: 10.3114/fuse.2021.07.02.
2. Ali, S., Gladieux, P., Leconte, M., Gautier, A., Justesen, A. F., Hovmøller, M. S., Enjalbert, J., and Vallavieille-Pope, C. (2014). Origin, Migration Routes and Worldwide Population Genetic Structure of the Wheat Yellow Rust Pathogen *Puccinia striiformis f.sp. tritici. PLoS Pathogens* 10(1), e1003903. doi:10.1371/journal.ppat.1003903.
3. Anonymous (2019). Country coffee profile: Vietnam. International Coffee Organization. pp. 9–10.
4. Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., Läderach, P., Anzueto, F., Hruska, A. J., and Morales, C. (2015). The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. *Food Sececurity* 7: 303–321. DOI 10.1007/s12571-015-0446-9.
5. Backer, C. A. and Bakhuizen Van Den Brink, R. C. (1965). *Flora of Java (Spermatophytes only)* (Vol. 2). Angiospermae, families 111–160. (pp.322–323). N. V. P. Noordhoff – Groningen – The Netherlands.
6. Bartlett, D. W., Clough, J. M., Godwin, J. R., Hall, A. A., Hamer, M., and Parr-Dobrzanski, B. (2002). The strobilurin fungicides. *Pest Management Science* 58:649–662. DOI: 10.1002/ps.520.
7. Belan, L. L., Belan, L. L., Rafael, A. M., Gomes, C. A. G., Alves, F. R., Junior, W. C. J., and Moraes, W. B. (2020). Standard area diagram with color photographs to estimate the severity of coffee leaf rust in *Coffea canephora*. *Crop Protection* 130: 105077. https://doi.org/10.1016/j.cropro.2020.105077.
8. Berkeley, M. J. and Broome, C. E. (1869) *Hemileia vastatrix*. Gardeners' Chronicle, 6, 1157.
9. Bowden, J., Gregory, P. H., and Johnson, C. G. (1971). Possible wind transport of coffee leaf rust across the Atlantic Ocean. *Nature* 229, 500–1.
10. Bryant, D. and Moulton, V. (2013). Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. *Molecular Biology* 21(2), 255–265. DOI: 10.1093/molbev/msh018.
11. Bunn, C., Läderach, P., Ovalle Rivera, O., and Kirschke, D. (2015). A bitter cup: climate change profile of global production of Arabica and Robusta coffee. *Climatic Change* 129, 89–101. https://doi.org/10.1007/s10584-014-1306-x.
12. Burdon, J. J and Chilvers, G. A. (1982). Host Density as a Factor in Plant Disease Ecology. A*nnual Review of Phytopathology* 20 (1): 143-166. 10.1146/annurev.py.20.090182.001043. https://www.annualreviews.org/doi/abs/10.1146/annurev.py.20.090182.001043.
13. Cabral, P. G. C., Maciel-Zambolim, E., Oliveira, S. A. S., Caixeta, E. T., and Zambolim, L. (2016). Genetic diversity and structure of *Hemileia vastatrix* populations on *Coffea* spp. *Plant Pathology* 65, 196–204. Doi: 10.1111/ppa.12411.
14. Carvalho, C. R., Fernandes, R. C., Carvalho, G. M., Barreto, R. W., and Evans, H. C. (2011). Cryptosexuality and the genetic diversity paradox in coffee rust, *Hemileia vastatrix*. *PLoS One* 6(11):e26387. doi: 10.1371/journal.pone.0026387. Epub 2011 Nov 15. PMID: 22102860; PMCID: PMC3216932.
15. Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. and Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3 (Bethesda)* 1(3):171-82. doi: 10.1534/g3.111.000240. PMID: 22384329; PMCID: PMC3276136.

16. Chechi, A., Ivey, M. L. L., Kaufman, R. R., Bryson, K. P., and Schnabel, G. (2020). Quinone outside inhibitor-resistant *Colletotrichum nymphaeae* isolates from strawberry lack mutations in *cytb* gene. *Journal of Plant Pathology* 102, 681–683.
17. Cinthia, Q.-A., Roberto, M.-S., César, L.-B., Rosa, E.-J., Villanueva-Caceda, J., and Monzón, C. (2017). Genetic diversity of *Hemileia vastatrix* of two coffee producing areas in Peru. *Revista Mexicana de Fitopatología* 35(3), 418–436. DOI: 10.18781/R.MEX.FIT.1612-7.
18. Cinthia, Q.-A., Roberto, M.-S., Rosa, E.-J., Giovanni, B., Marisela, Y.-L., and César, L.-B. (2021). Spatial and Temporal Genetic Diversity and Population Structure of *Hemileia vastatrix* from Peruvian Coffee Plantations. *Plant Pathology Journal* 37(3): 280–290. https://doi.org/10.5423/PPJ.OA.10.2020.0192
19. Circular No. 19/2021/TT-BNNPTNT dated 28/12/2021 (2021). Danh mục thuốc bảo vệ thực vật được phép sử dụng tại Việt Nam. Retrieved from https://snnptnt.thuathienhue.gov.vn/UploadFiles/TinTuc/2022/3/18/phuluc1(2).pdf.
20. Cook, R. T. A. (2001). First report in England of changes in the susceptibility of *Puccinia horiana*, the cause of chrysanthemum white rust, to triazole and strobilurin fungicides. *Plant pathology* 50, 792. doi: 10.1046/j.1365-3059.2001.00615.x.
21. Costa, G. A., Lira, J. B., Lopes, R. L. F. and Lopes, U. P. (2019) Tank mix application of copper hydroxide either with cyproconazole or pyraclostrobin fungicides reduced the control of coffee leaf rust. *Crop Protection* 124, 104856. https://doi.org/10.1016/j.cropro.2019.104856.
22. Cristancho, M. A., Rozo, Y., Escobar, C., Rivillas, C. A., and Gaitan, A., L. (2012). Outbreak of coffee leaf rust (*Hemileia vastatrix*) in Colombia. *New Disease Reports* 25, 19. [doi:10.5197/j.2044-0588.2012.025.019].
23. Cummins, G. B. and Hiratsuka, Y. (2003). Illustrated genera of rust fungi. Third ed. *American Phytophathological Society*, St. Paul, MN. APS Press, St. Paul, MN.
24. Davis, A. P. (2011). *Psilanthus mannii*, the type species of *Psilanthus*, transferred to *Coffea*. *Nordic Journal of Botany* 29: 471 – 472.
25. Davis, A. P., Chadburn, H., Moat, J., O'Sullivan, R., Hargreaves, S. and Nic Lughadha, E. (2019). High extinction risk for wild coffee species and implications for coffee sector sustainability. *Science Advance* 5:eaav3473. https://doi.org/10.1126/sciadv.aav3473
26. D'haeze, D., Deckers, J., Raes, D., Tran, A. P., and Nguyen, D. M. C. (2003). Over-irrigation of *Coffea canephora* in the Central Highlands of Vietnam revisited Simulation of soil moisture dynamics in Rhodic Ferralsols. *Agricultural Water Management* 63: 185–202. doi:10.1016/S0378-3774(03)00181-1.
27. D'haeze, D., Deckers, J., Raes, D., Tran, A. P., and Loi, H. V. (2004). Environmental and socio-economic impacts of institutional reforms on the agricultural sector of Vietnam Land suitability assessment for Robusta coffee in the Dak Gan region. *Agriculture, Ecosystems and Environment* 105: 59–76. doi:10.1016/j.agee.2004.05.009.
28. D'Oliveira, B. (1971). *Report of the Coffee Rusts Research Centre, Oeiras*, Portugal.
29. Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14(8): 2611-20. doi: 10.1111/j.1365-294X.2005.02553.x. PMID: 15969739.
30. Excoffier, L. and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x.
31. Fernandes, L. H. M., Resende, M. L. V., Costa, B. G., Pereira, R. B., Costa, B. H. G., Monteiro, A. C. A., and Ribeiro Junior, P. M. (2013) Acibenzolar-S-Methyl no controle da ferrugem e da Cercosporiose do cafeeiro em condic¸~oes de campo. *Coffee Science* 8, 24–32.

32. Fernandes, R. de C., Evans, H. C., and Barreto, R. W. (2009). Confirmation of the occurrence of teliospores of Hemileia vastatrix in Brazil with observations on their mode of germination. *Tropical Plant Pathology* 34(2). https://doi.org/10.1590/s1982-56762009000200005.
33. Fernández-Ortuño, D., Torés, J. A., de Vicente, A., and Pérez-García, A. (2008). Mechanisms of resistance to QoI fungicides in phytopathogenic fungi. *International Microbiology* 11(1):1-9. PMID: 18683626. doi: 10.1016/s0958-1669(96)80043-7
34. Ferreira, T., Shuler, J., Guimarães, R., and Farah, A. (2019). CHAPTER 1. Introduction to Coffee Plant and Genetics. In *Coffee*, Cambridge: Royal Society of Chemistry, p. 1–25.
35. Fraczek, M. G., Zhao, C., Dineen, L., Lebedinec, R., Bowyer, P., Bromley, M., and Delneri, D. (2019). Fast and reliable PCR amplification from *Aspergillus fumigatus* spore suspension without traditional DNA extraction. *Current Protocols in Microbiology* 54, e89. doi: 10.1002/cpmc.89.
36. Fungicide Resistance Action Committee (2020). List of first confirmed cases of plant pathogenic organisms resistant to disease control agents. CropLife Int., Brussels, Belgium. Retrieved from https://www.frac.info/docs/default-source/publications/list-of-resistant-plant-pathogens/list-of-first-confirmed-cases-of-plant-pathogenic-organisms-resistant-to-disease-control-agents_05_2020.pdf?sfvrsn=7073499a_2.
37. Grasso, V., Palermo, S., Sierotzki, H., Garibaldi, A., and Gisi, U. (2006). *CYTB* gene structure and consequences for resistance to Qo inhibitor fungicides in plant pathogens. *Pest Management Science* 62(6): 465-72. doi: 10.1002/ps.1236. PMID: 16688790.
38. Haddad, F., Maffia, L. A., Mizubuti, E. S. G., and Teixeira, H. (2009). Biological control of coffee rust by antagonistic bacteria under field conditions in Brazil. *Biological Control* 49, 114–119. https://doi.org/10.1016/j.biocontrol.2009.02.004.
39. Haddad, F., Saraiva, R. M., Mizubuti, E. S. G., Romeiro, R. S., and Maffia, L. A. (2013). Antifungal compounds as a mechanism to control *Hemileia vastatrix* by antagonistic bacteria. *Tropical Plant Pathology* 38, 398–405. https://doi.org/10.1590/S1982567620130005000004 .
40. Haddad, F., Saraiva, R. M., Mizubuti, E. S. G., Romeiro, R. S., and Maffia, L. A. (2014). Isolation and selection of *Hemileia Vastatrix* antagonists. *European Journal of Plant Pathology* 139, 763–772. https://doi.org/10.1007/s10658-014-0430-9.
41. Hamelin, R. C, Hunt, R. S, Geils, B. W., Jensen, G. D, Jacobi, V., and Lecours, N. (2000). Barrier to gene flow between Eastern and Western populations of *Cronartium ribicola* in North America. *Phytopathology* 90, 1073–8.
42. Hammer, Ø., Harper, D. A. T., and Paul D. R. (2001). Past: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica* 4 (1). http://palaeo-electronica.org/2001_1/past/issue1_01.htm.
43. Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution* 59(8): 1633–1638. https://doi.org/10.1111/j.0014-3820.2005.tb01814.x.
44. Hennings, P. (1895). Neue und interessante Pilze aus dem Königl. Botanischen Museum in Berlin. III. *Hedwigia* 34: 10–13.
45. Honorato, J., Zambolim, L., do Nascimento Lopes, U., Pedro Lopes, U., and da Silva Silveira Duarte, H. (2015). DMI and QoI fungicides for the control of coffee leaf rust. *Australasian Plant Pathology* 44, 575–581. https://doi.org/10.1007/s13313-015-0373-4.
46. Huson, D. H. and Bryant, D. (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution* 23(2): 254-267.
47. Keiper, F. J., Haque, M. S., Hayden, M. J., and Park, R. F. (2006). Genetic diversity in Australian populations of *Puccinia graminis f. sp. avenae*. *Phytopathology* 96, 96–104.

48. Klosowski, A. C., May De Mio, L. L., Miessner, S., Rodrigues, R., and Stammler, G. (2016). Detection of the F129L mutation in the *CYTB* gene in *Phakopsora pachyrhizi. Pest Management Science* 72(6):1211-5. doi: 10.1002/ps.4099. Epub 2015 Sep 28. PMID: 26296393.
49. Kolmer, J. A., Ordonez, M. E., and Groth, J. V. (2009). The Rust Fungi. In: *Encyclopedia of Life Sciences (ELS)*. John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0021264.
50. Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547–1549. doi:10.1093/molbev/msy096.
51. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21): 2947–2948. DOI: 10.1093/bioinformatics/btm404.
52. Le, Q. V, Cowal, S., Jovanovic, G., and Le, D. T (2021). A Study of Regenerative Farming Practices and Sustainable Coffee of Ethnic Minorities Farmers in the Central Highlands of Vietnam. *Frontiers in Sustainable Food Systems* 5:712733. doi: 10.3389/fsufs.2021.712733.
53. Leigh, J. W. and Bryant, D. (2015). POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution 6*, 1110–1116. https://doi.org/10.1111/2041-210X.12410.
54. Li, F., Upadhyaya, N. M., Sperschneider, J., Matny, O., Nguyen, H. P., Mago, R., Raley, C., Miller, M. E., Silverstein, K. A. T., Henningsen, E., Hirsch, C. D., Visser, B., Pretorious, Z. A., Steffenson, B. J., Schwessinger, B., Dodds, P. N., and Figueroa, M. (2019). Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. *Nature Communications* 10, 5068. https://doi.org/10.1038/s41467-019-12927-7.
55. Librado, P and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25(11):1451-2. doi: 10.1093/bioinformatics/btp187. Epub 2009 Apr 3. PMID: 19346325.
56. Lopes, U. P., Zambolim, L., Neto, P. N. S., Souza, A. F., Capucho, A. S., and Rodrigues, F. A. (2013). Effect of foliar application of potassium silicate on the progress of coffee leaf rust. *Tropical Plant Pathology* 38, 547–551. https://doi.org/10.1590/S1982-56762013000600012.
57. Maia, T. A., Maciel-Zambolim, E., Caixeta, E. T., Mizubuti, E. S. G., and Zambolim, L. (2013). The population structure of *Hemileia vastatrix* in Brazil inferred from AFLP. *Australasian Plant Pathology* 42, 533–542. DOI 10.1007/s13313-013-0213-3.
58. Matsuura, S. (2019). Does QoI (strobilurin) resistance in isolates of *Puccinia horiana*, the causal agent of chrysanthemum white rust, occur in western Japan? *Journal of Plant Diseases and Protection* 126, 469–473. https://doi.org/10.1007/s41348-019-00224-w.
59. Matsuzaki, Y., Harada, T., and Iwahashi, F. (2020). Amino acid substitutions responsible for different QoI and SDHI sensitivity patterns in *Puccinia horiana*, the causal agent of chrysanthemum white rust. *Plant Pathology* 70(2):377-386. DOI: 10.1111/ppa.13298.
60. Matsuzaki, Y., Harada, T. and Iwahashi, F. (2021). New cytochrome *b* haplotypes, harboring L299F or N256S + L299F substitutions, were found in azoxystrobin-resistant *Puccinia horiana*, the causal agent of chrysanthemum white rust. *European Journal Plant Pathology* 160, 963–972. https://doi.org/10.1007/s10658-021-02299-4.
61. Mayne, W.W. (1932). Physiological specialisation of *Hemileia vastatrix* Berk. & Br. *Nature* 129, 510.
62. Mboup, M., Leconte, M., Gautier, A., Wan, A. M., Chen, W., de Vallavieille-Pope, C., and Enjalbert, J. (2009). Evidence of genetic recombination in wheat yellow rust populations

of a Chinese oversummering area. *Fungal Genetics and Biology* 46(4): 299-307. doi: 10.1016/j.fgb.2008.12.007. Epub 2009 Jan 14. PMID: 19570502.

63. McCook, S. (2006). Global rust belt: *Hemileia vastatrix* and the ecological integration of world coffee production since 1850. *Journal of Global History* 1: 177–195. doi:10.1017/S174002280600012X.
64. McCook S. (2017). Environmental history of coffee in Latin America. *Latin American History*. https://doi.org/10.1093/acrefore/9780199366439.013.440.
65. McCook, S. (2019). Coffee is not forever: a global history of coffee leaf rust. *Ohio University Press* pp. 67–68. ISBN 13: 9780821423875.
66. Müller, M. A., M. Müller, A., Stammler, G., Stammler, G., and Larissa May De Mio, L. (2021). Multiple resistance to DMI, QoI and SDHI fungicides in field isolates of *Phakopsora pachyrhizi*. *Crop protection* 145, 105618. doi: 10.1016/j.cropro.2021.105618
67. Nei, M. and Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences* 76 (10), 5269–5273. http://dx.doi.org/10.1073/ pnas.76.10.5269.
68. Noronha-Wagner, M. and Bettencourt, A. J. (1967). Genetic study of the resistance of *Coffea* spp. to leaf rust. I. The identification and behaviour of four factors conditioning disease reaction in *Coffea arabica* to twelve hysiologic races of *H. vastatrix*. *Canadian Journal of Botany* 45, 2021-2031.
69. Oliver, R. P. (2014). A reassessment of the risk of rust fungi developing resistance to fungicides. *Pest Management Science* 70(11):1641-5. doi: 10.1002/ps.3767. Epub 2014 Mar 26. PMID: 24616024.
70. Park, R. F., Burdon, J. J., and Jahoor, A. (1999). Evidence for somatic hybridization in nature in *Puccinia recondita f. sp. tritici*, the leaf rust pathogen of wheat. *Mycological Research* 103(6):715-723. DOI: 10.1017/s0953756298007631.
71. Park, R. F. and Wellings, C. R. (2012). Somatic hybridization in the Uredinales. *Annual Review of Phytopathology* 50:219-39. doi: 10.1146/annurev-phyto-072910-095405. PMID: 22920559.
72. Parks, D. H., Mankowski, T., Zangooei, S., Porter, M. S., Armanini, D. G., Baird, D. J., Langille, M. G. I., and Beiko, R. G. (2013). GenGIS 2: Geospatial analysis of traditional and genetic biodiversity, with new gradient algorithms and an extensible plugin framework. *PLoS One* 8:7, e69885. https://doi.org/10.1371/journal.pone.0069885.
73. Peakall, R. and Smouse P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28, 2537–2539. DOI: 10.1093/bioinformatics/bts460.
74. Pereira, R. B., Lucas, G. C., Perina, F. J., and Alves, E. (2012). Essential oils for rust control on coffee plants. *Ciência e Agrotecnologia* 36, 16–24. https://doi.org/10.1590/S1413-70542012000100002
75. Ramírez-Camejo, L. A, Eamvijarn, A., Díaz-Valderrama, J. R., Karlsen-Ayala, E., Johnson, E., Pruvot-Woehl, S., Meija, L. C., Montagnon, C., and Aime, M. C. (2022). Global analysis of *Hemileia vastatrix* populations shows clonal reproduction for the coffee leaf rust pathogen throughout most of its range. *Phytopathology* 112(3):643-652. doi: 10.1094/PHYTO-06-21-0255-R. Epub 2022 Mar 9. PMID: 34428920.
76. Ritschel, A. (2005). Monograph of the genus *Hemileia* (Uredinales). In: *Bibliotheca Mycologica*, vol.200 (Bresinsky, A., Butin, H., & Tudzinski, P., eds). Stuttgart: J. Cramer.
77. Rodrigues, A. S. B., Silva, D. N., Várzea, V., Paulo, O. S., and Batista, D. (2022). Worldwide Population Structure of the Coffee Rust Fungus *Hemileia vastatrix* Is Strongly Shaped by Local Adaptation and Breeding History. *Phytopathology* 112(9):1998-2011. doi: 10.1094/PHYTO-09-21-0376-R. Epub 2022 Aug 2. PMID: 35322716.

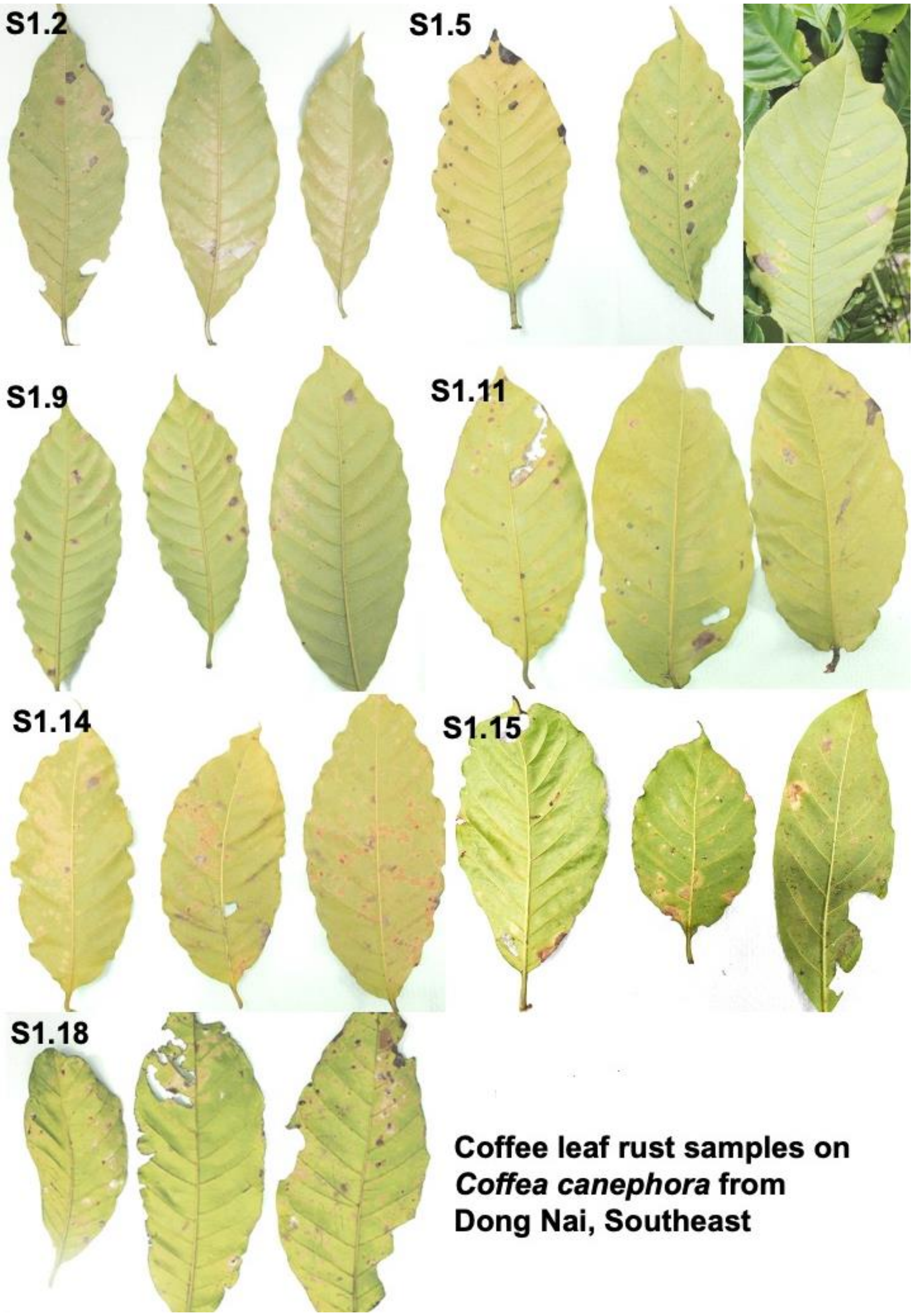
78. Rodrigues, C. J. Jr., Bettencourt, A. J., and Rijo, L. (1975) Races of the pathogen and resistance to coffee rust. *Annual Review of Phytopathology* 13, 49–70.
79. Rothfos B. (1980). Coffee diseases caused by microscopical fungi and how they are controlled" in Coffee production. Gordian Max Riech. P 137-139.
80. Ryan, W. B. F., Carbotte, S. M., Coplan, J. O., O'Hara, S., Melkonian, A., Arko, R., Weissel, R.A., Ferrini, V., Goodwillie, A., Nitsche, F., Bonczkowski, J., and Zemsky R. (2009). Global Multi-Resolution Topography (GMRT) synthesis data set. *Geochemistry, Geophysics, Geosystems* 10, Q03014, doi:10.1029/2008GC002332.
81. Saleh, D., Milazzo, J., Adreit, H., Fournier, E., and Tharreau, D. (2014). South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, *Magnaporthe oryzae*. *New Phytologist* 201(4):1440-1456. doi: 10.1111/nph.12627. Epub 2013 Dec 10. PMID: 24320224; PMCID: PMC4265293.
82. Santos, F. S., Souza, P. E., Resende, M. L. V., Pozza, E. A., Miranda, J. C., and Je Pests Acrossand Manerba, F. C. (2007). Effect of vegetal extracts on the progress of foliar diseases in organic coffee. *Fitopatologia Brasileira* 32, 59–63.
83. Sierotzki, H., Wullschleger, J., and Gisi, U. (2000). Point mutation in *CYTB* gene conferring resistance to strobilurin fungicides in *Erysiphe graminis* f sp *tritici* field isolates. *Pesticide Biochemistry and Physiology* 68:107–112.
84. Silva, D. N., Várzea, V., Paulo, O. S., and Batista, D. (2018). Population genomic footprints of host adaptation, introgression and recombination in coffee leaf rust. *Molecular Plant Pathology* 19(7):1742-1753. doi: 10.1111/mpp.12657. Epub 2018 Feb 22. PMID: 29328532; PMCID: PMC6638104.
85. Silva, H. S. A., Tozzi, J. P. L., Terrasan, C. R. F., and Bettiol, W. (2012). Endophytic microorganisms from coffee tissues as plant growth promoters and biocontrol agents of coffee leaf rust. *Biological Control* 63, 62–67. https://doi.org/10.1016/j.biocontrol.2012.06.005
86. Silva, M. C., Várzea, V., Guerra-Guimarães, L., Azinheira, H. G., Fernandez, D., Petitot, A. – S., Bertrand, B., Lashermes, P., and Nicole, M. (2006). Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Brazilian Journal of Plant Physiology* 18(1), 119–147. https://doi.org/10.1590/S1677-04202006000100010.
87. Souza, A. F., Zambolim, L., Jesus Jr., V. C., and Cecon, P. R. (2011). Chemical approaches to manage coffee leaf rust in drip irrigated trees. *Australasian Plant Pathology* 40 (3): 293.
88. Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* 123, 585–595.
89. Talhinhas, P., Batista, D., Diniz, I., Vieira, A., Silva, D. N., Loureiro, A., Tavare, I., Pereira, A. P., Ira, H. G. A., Guimaraes, L. G.-G., Varzea, V., and Silva, M. D. (2017). The coffee leaf rust pathogen *Hemileia vastatrix*: one and a half centuries around the tropics. *Molecular plant pathology* 18(8), 1039–1051. DOI: 10.1111/mpp.12512.
90. Tavaré S. (1986). Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. *Lectures on Mathematics in the Life Sciences* 17: 57–86.
91. Thirumalachar, M. J., and Narasimhan, M. J. (1947). Studies on the Morphology and Parasitism of *Hemileia* Species on Rubiaceae in Mysore. *Annals of Botany* 11(41): 77–89. http://www.jstor.org/stable/42907003.
92. Tuchscherer, M. (2003). Coffee in the Red Sea Area from the Sixteenth to the Nineteenth Century. In W. Clarence-Smith and S. Topik (Eds.), *The Global Coffee Economy in Africa, Asia, and Latin America*, 1500-1898 (pp. 50-60). Cambridge: Cambridge University Press. Doi:10.1017/CBO9780511512193.013.
93. Vandermeer, J., Perfecto, I., and Philpott, S. (2010). Ecological complexity and pest control in organic coffee production: uncovering an autonomous ecosystem service. *BioScience* 60,

527–537. https://doi.org/10.1525/bio.2010.60.7.8

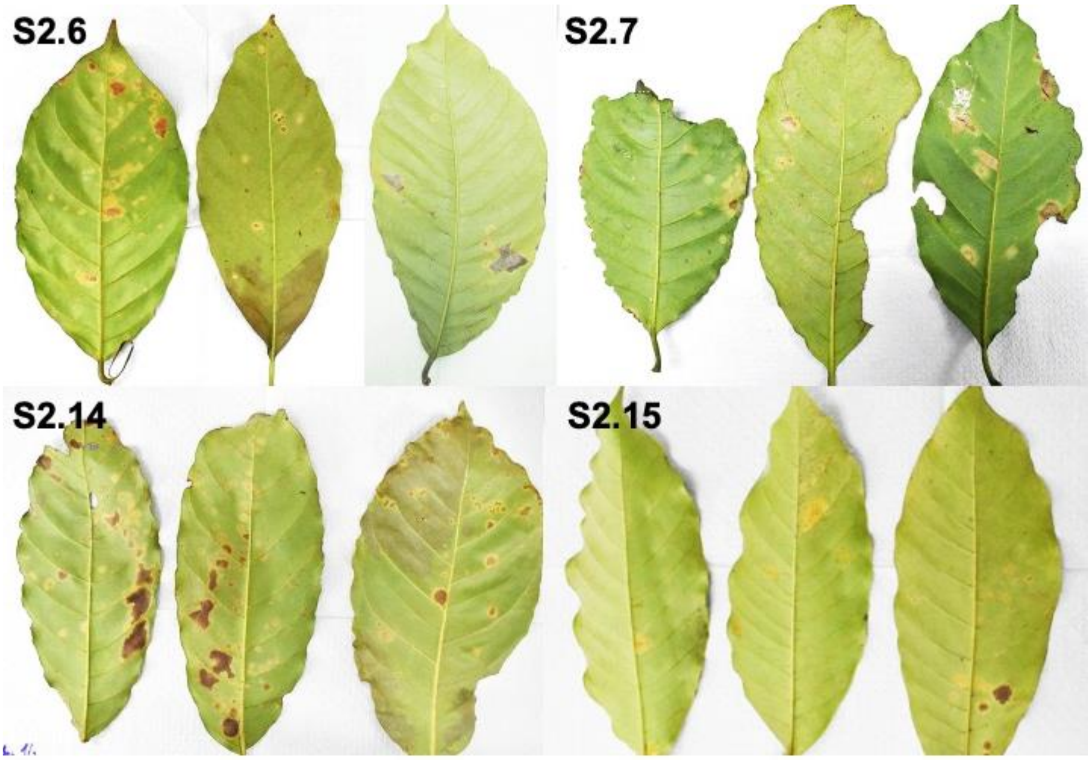
94. Waller, J. M. (1982). Coffee rust – epidemiology and control. *Crop Protection* 1 (4): 385-404. DOI 10.1016/0261-2194(82)90022-9. ISSN 0261-2194.
95. Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.2307/2408641.
96. White, T. J., Bruns, T., Lee, S., and Taylor, J. W. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M. A., Gelfand, D. H., Sninsky, J. J., & White, T. J. (eds) *PCR protocols: a guide to methods and applications*. *Academic Press*, New York: 315–322.
97. Yin, L. F., Hu, M. J., Wang, F., Kuang, H., Zhang, Y., Schnabel, G., Li, G. Q., and Luo, C. X. (2012). Frequent gain and loss of introns in fungal *CYTB* genes. *PLoS One* 7(11): e49096. doi: 10.1371/journal.pone.0049096. Epub 2012 Nov 7. PMID: 23145081; PMCID: PMC3492308.
98. Zambolim, L. (2016) Current status and management of coffee leaf rust in Brazil. *Tropical Plant Pathology* 41, 1–8. https://doi.org/10.1007/s40858-016-0065-9.
99. Zambolim, L., Reis, E. M., Guerra, W. D., Juliatti, F. C., and Menten, J. O. M. (2022). Integrated Management of Asian Soybean Rust. *European Journal of Applied Sciences* 10(2): 602-633. DOI:10.14738/aivp.102.12215.

Appendix:

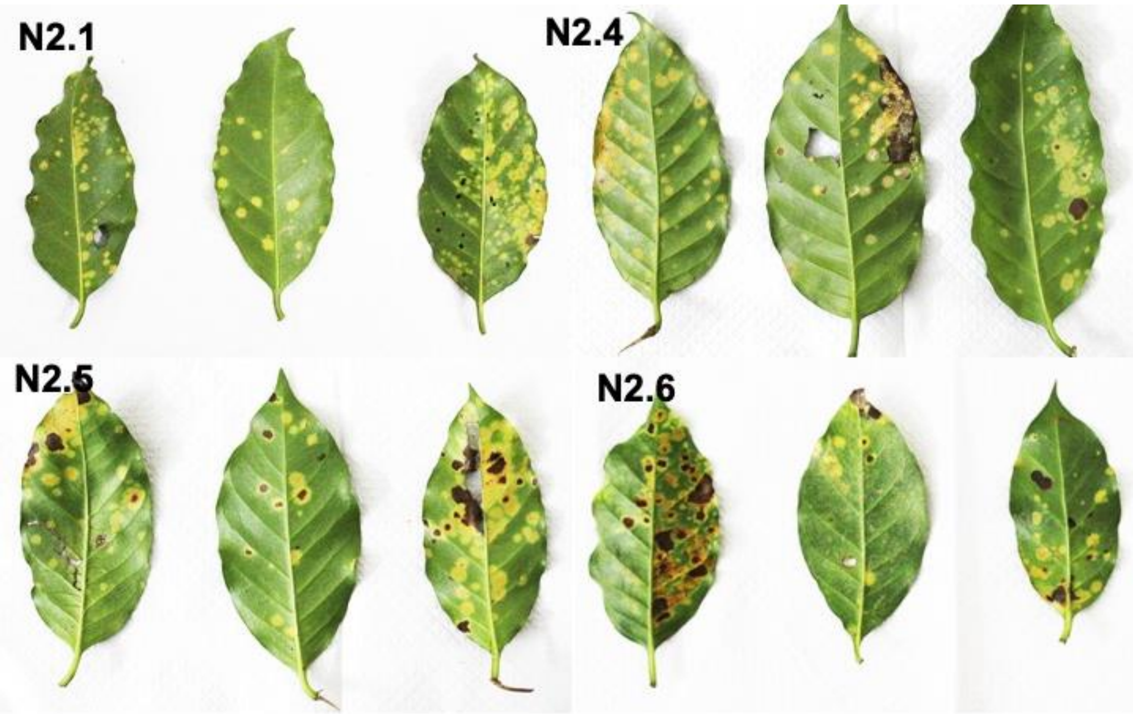
Appendix 1: Coffee leaf rust samples collected from *Coffea canephora* and *Coffea arabica cv. Catimor* in Vietnam.



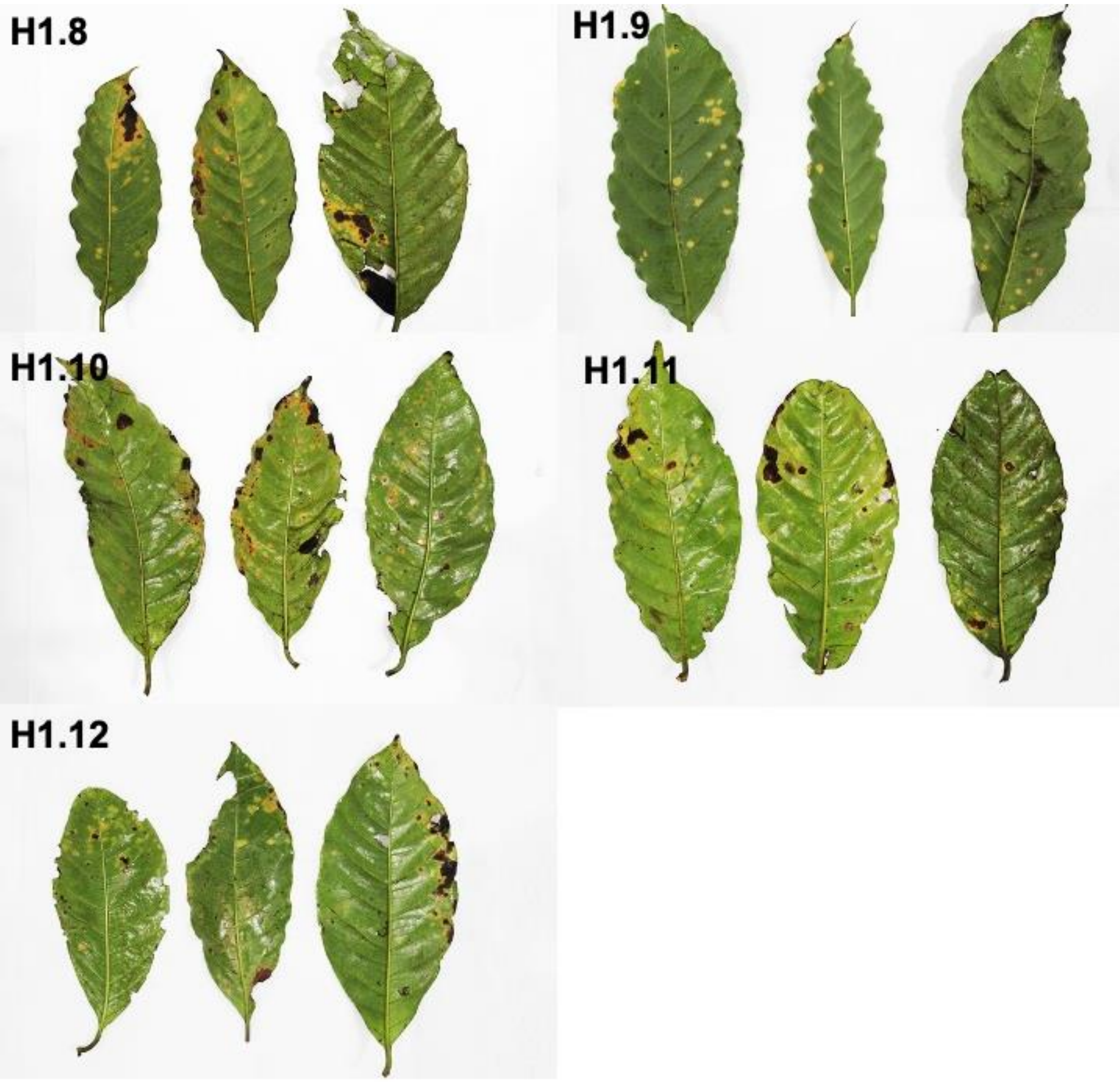
Coffee leaf rust samples on *Coffea canephora* from Dong Nai, Southeast



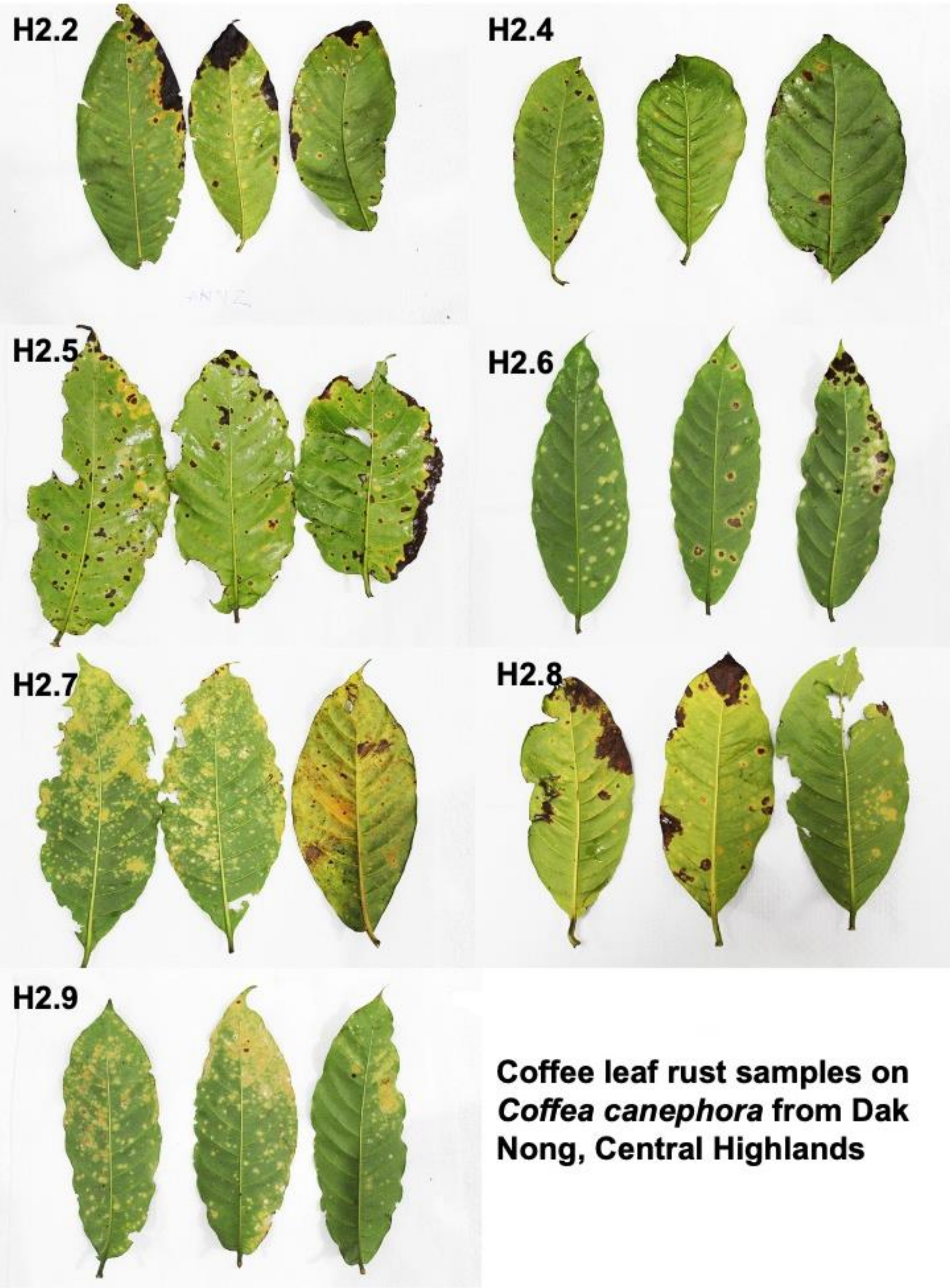
Coffee leaf rust samples on *Coffea canephora* from Binh Phuoc, Central Highlands



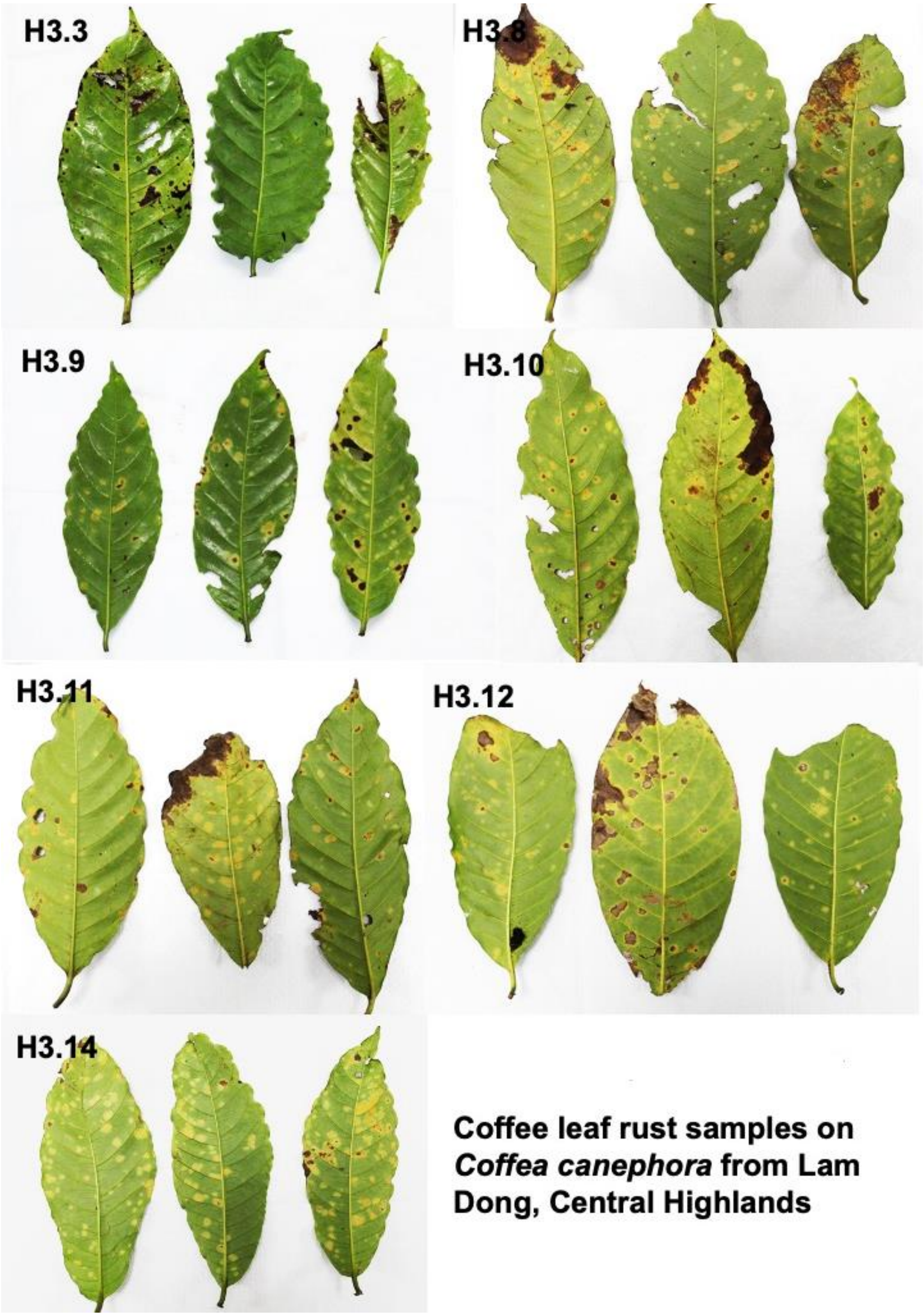
Coffee leaf rust samples on *Coffea arabica* from Dien Bien, Northwest



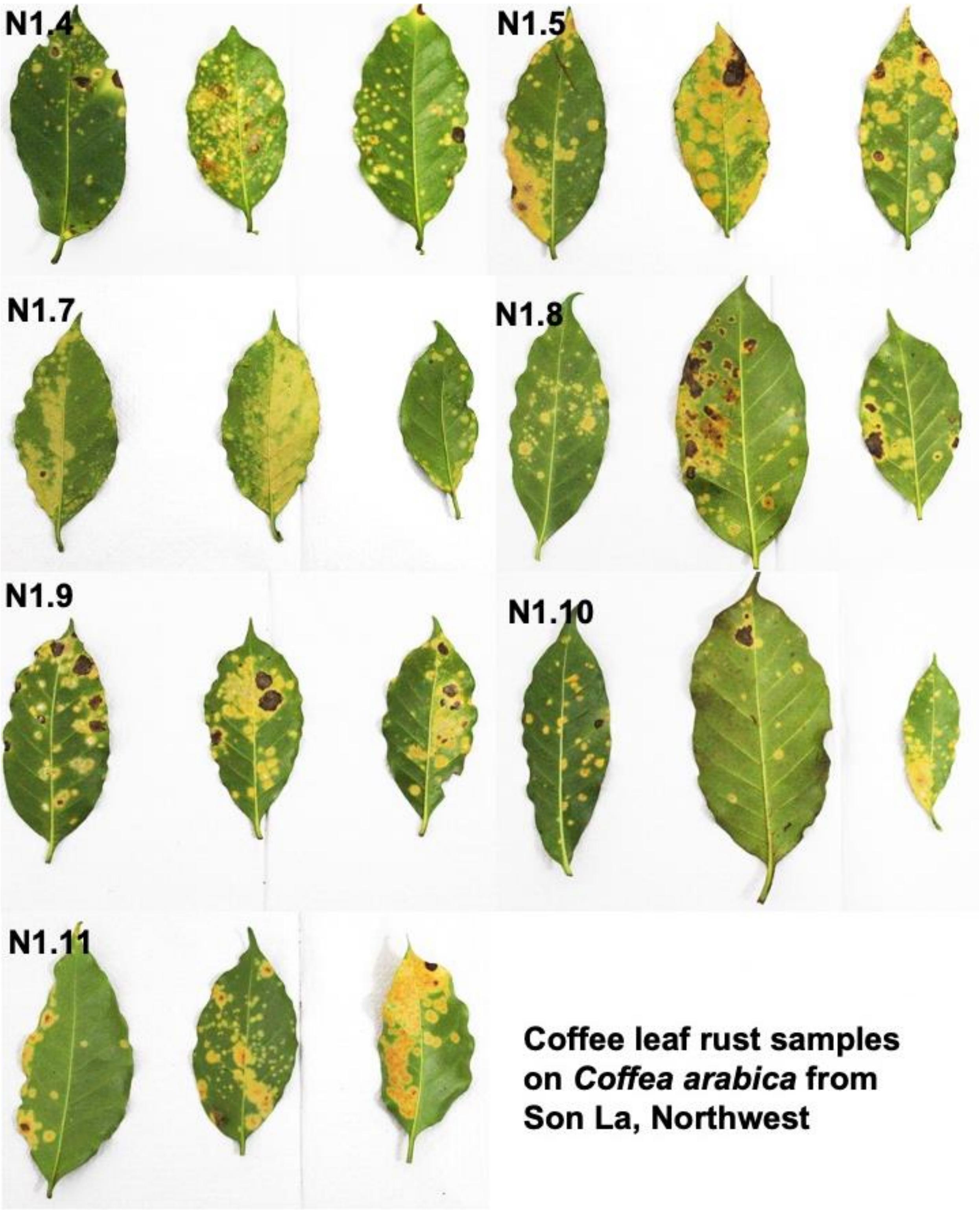
Coffee leaf rust samples on *Coffea canephora* from Dak Lak, Central Highlands



Coffee leaf rust samples on *Coffea canephora* from Dak Nong, Central Highlands

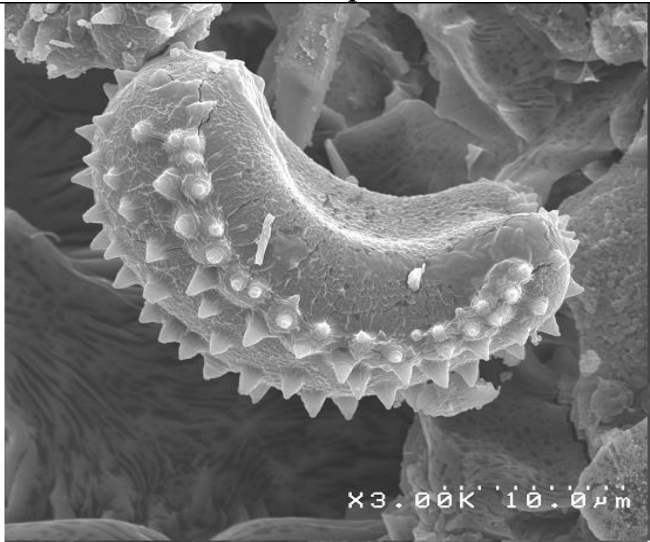
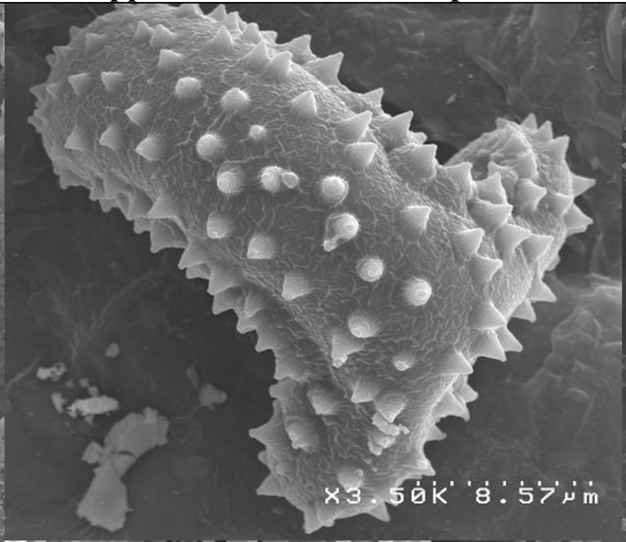
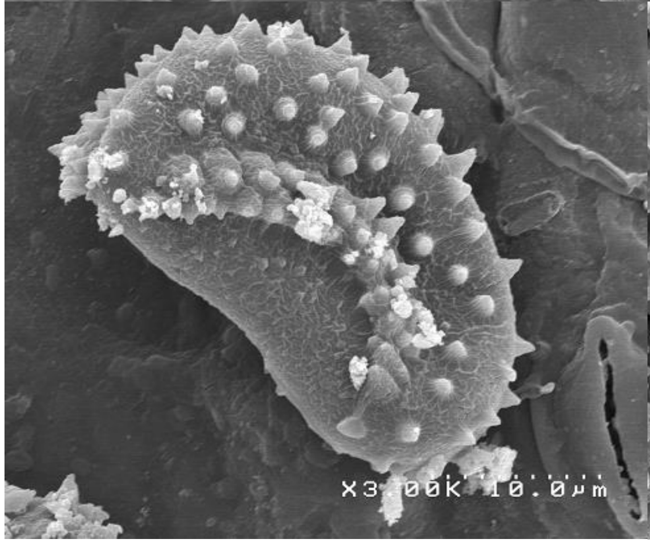
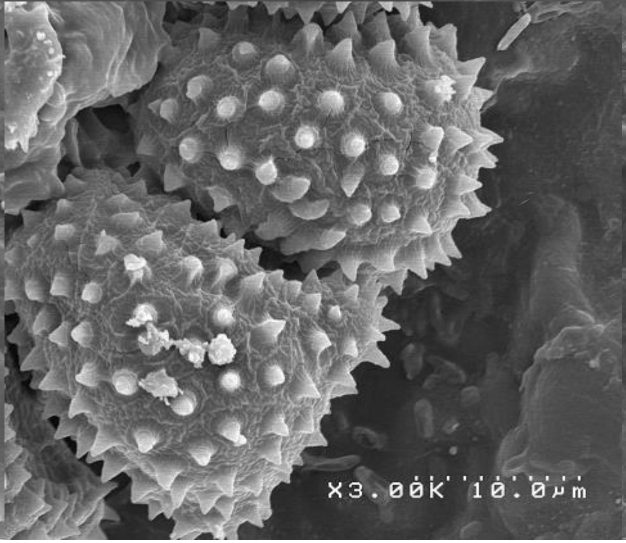


Coffee leaf rust samples on *Coffea canephora* from Lam Dong, Central Highlands

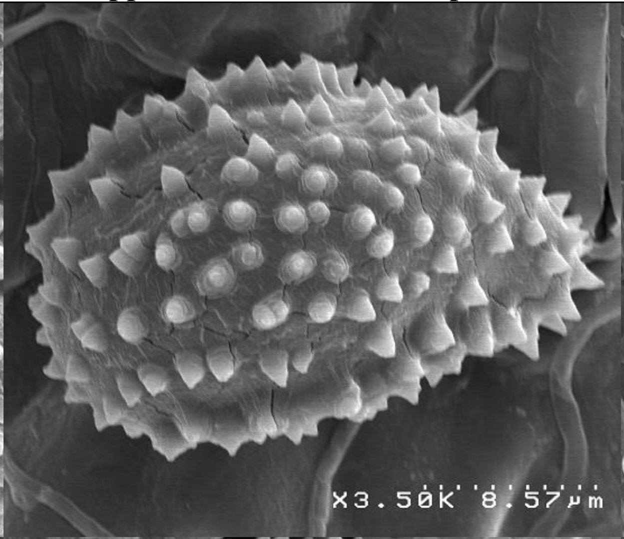
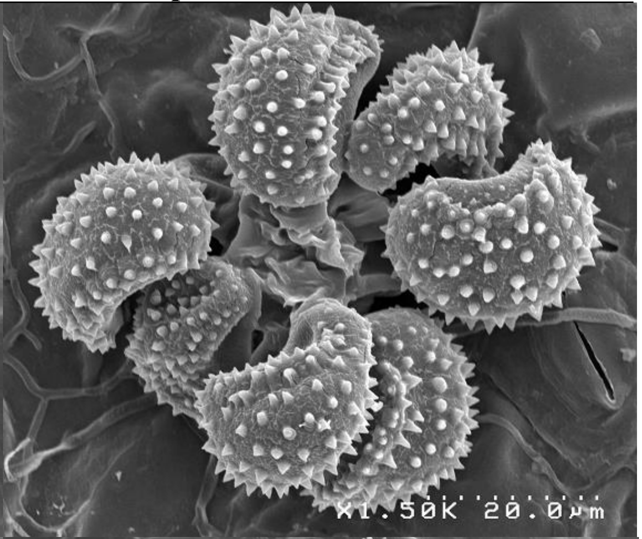
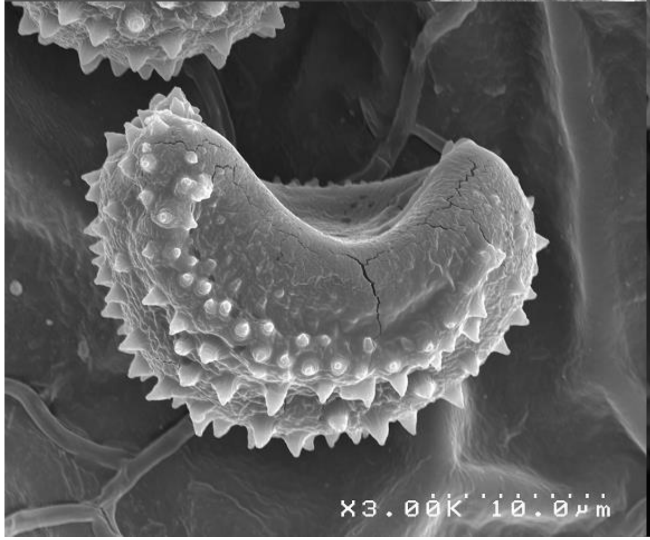
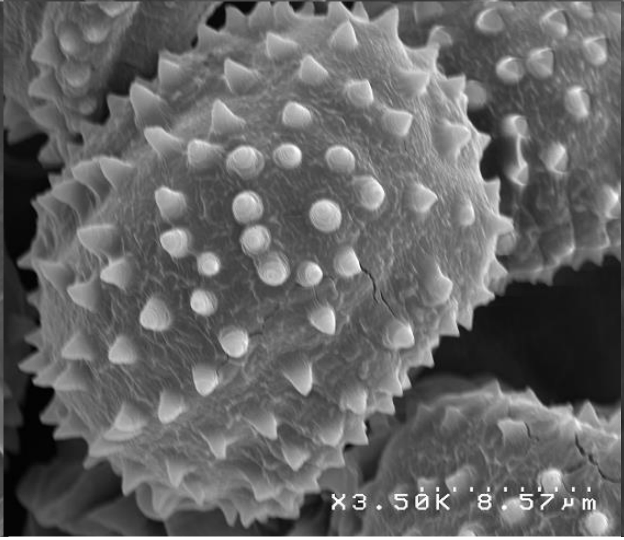
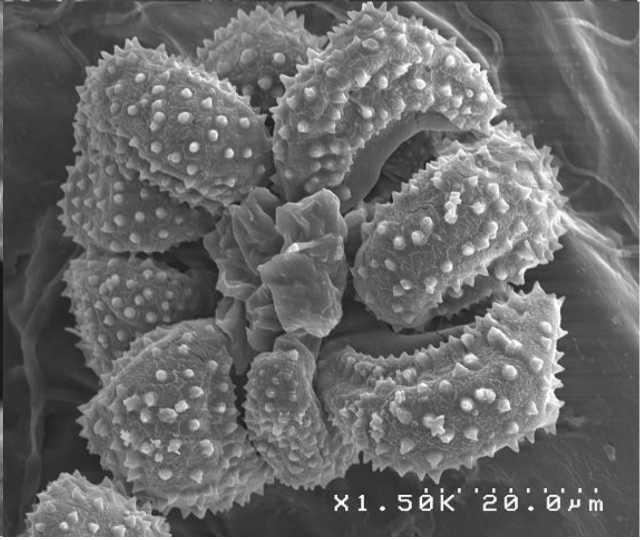


Coffee leaf rust samples on *Coffea arabica* from Son La, Northwest

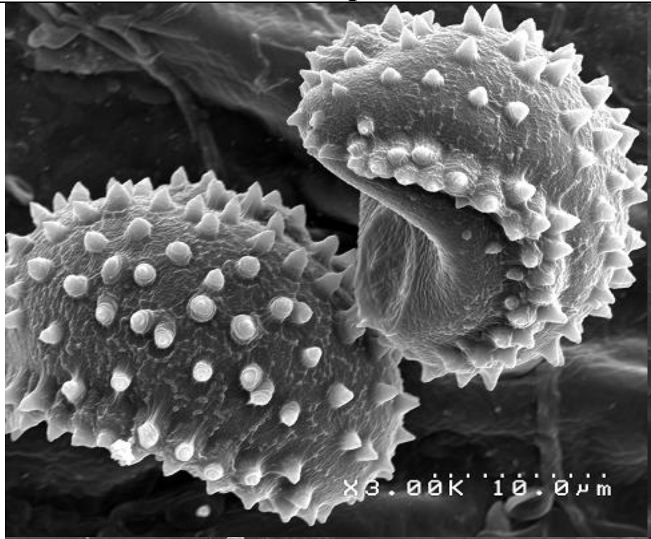
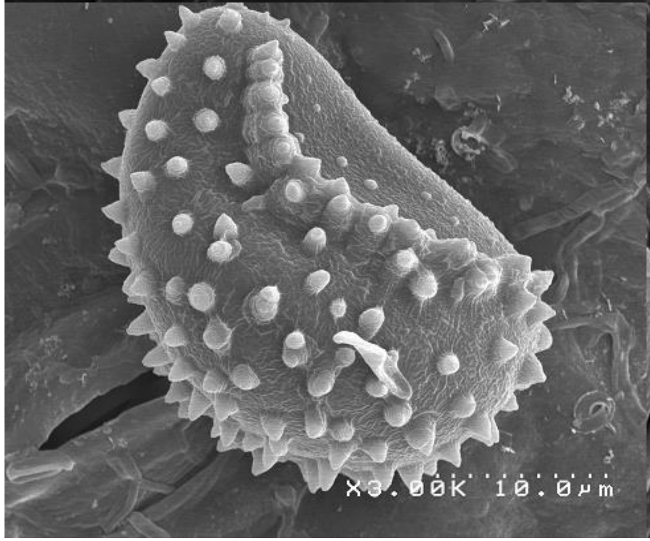
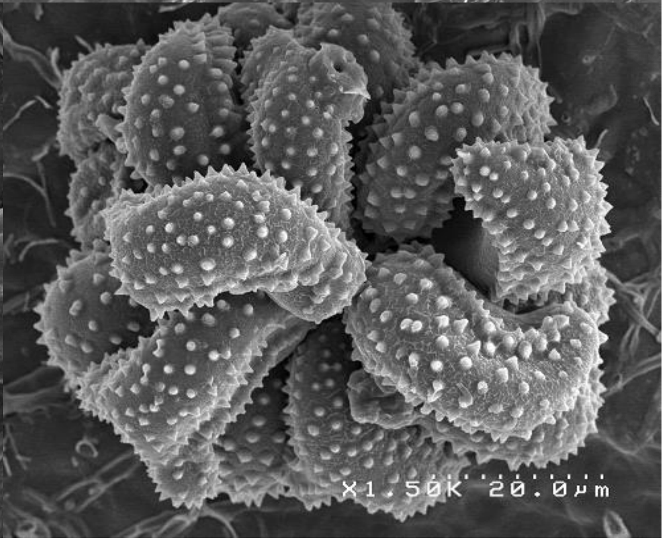
Appendix 2: Morphology of urediniospore: upper surface of urediniospore and suprastomatal uredinium of coffee rust fungus collected in Vietnam

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30018			
TSH-R30022			

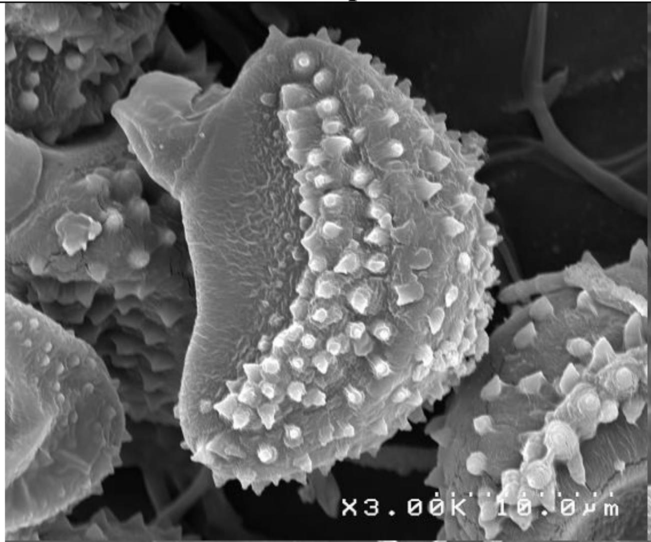
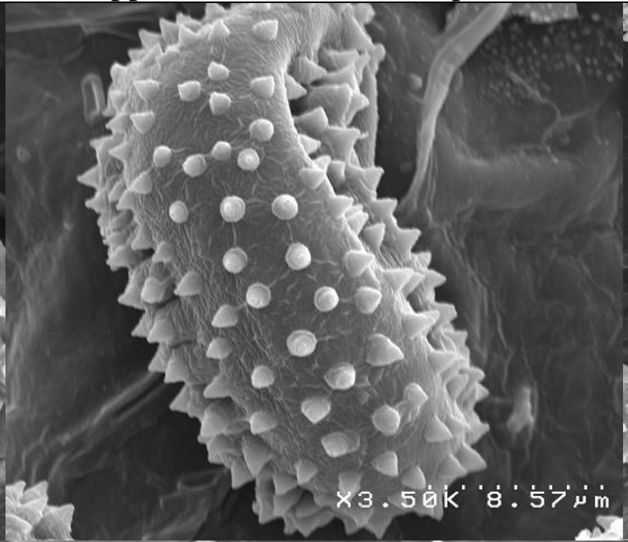
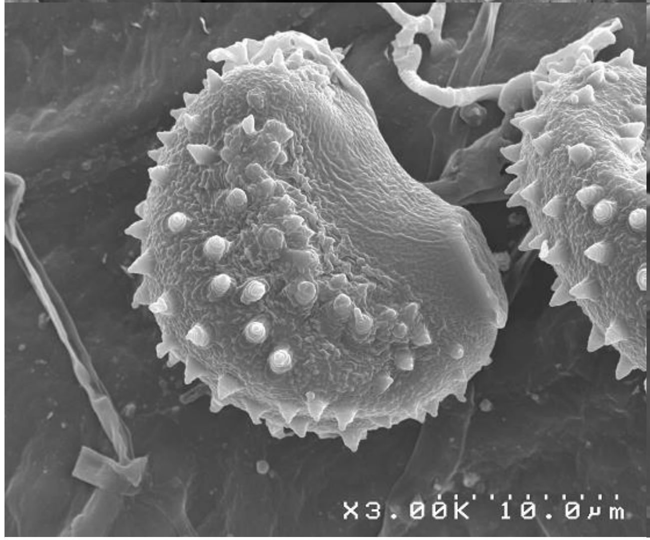
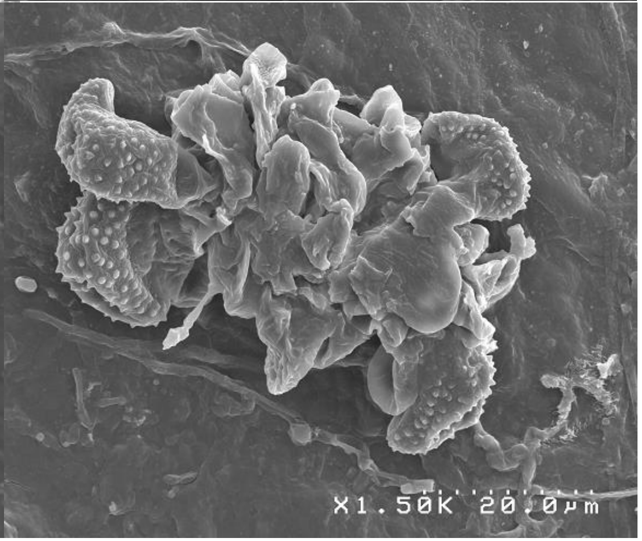
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30024			
TSH-R30028			

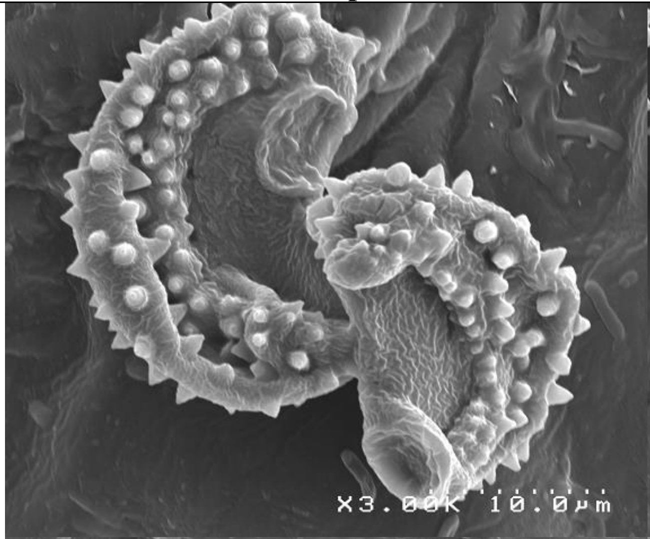
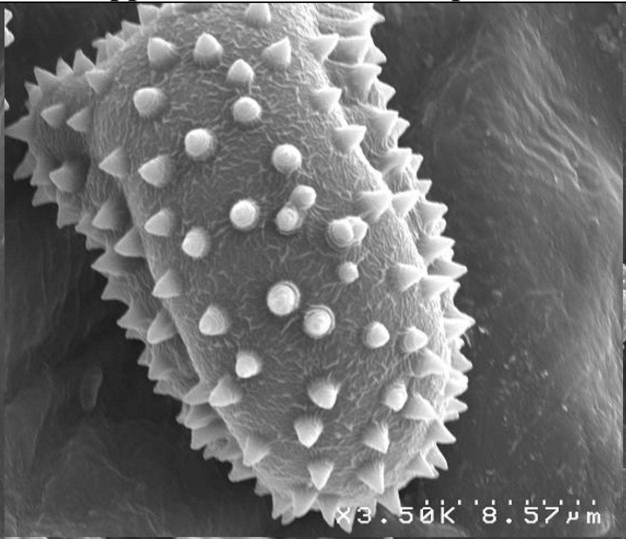
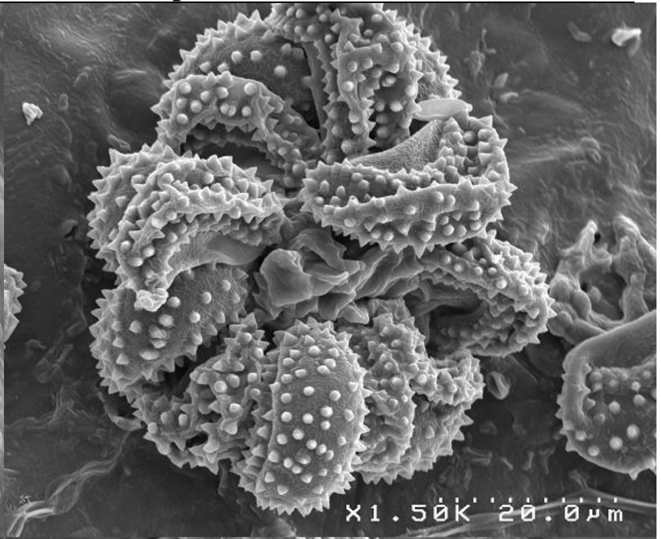
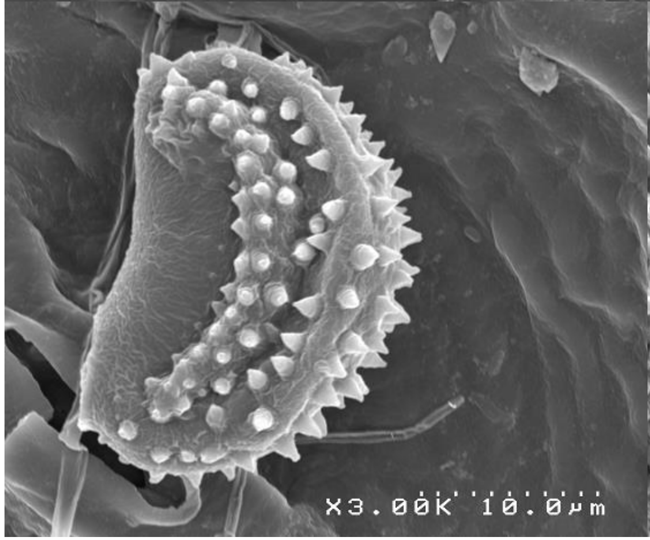
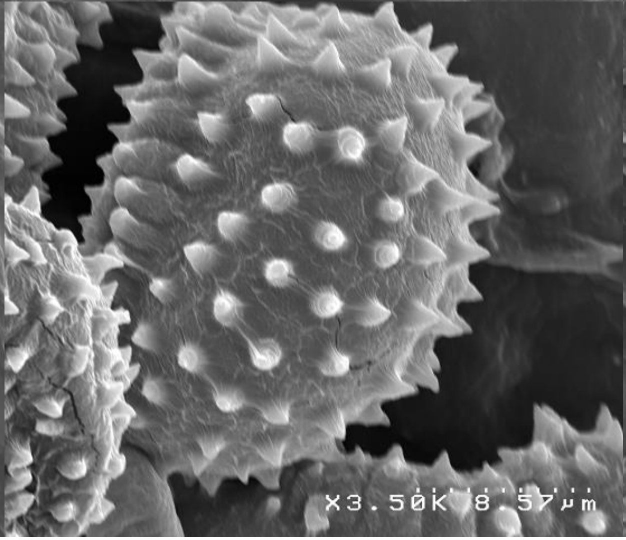
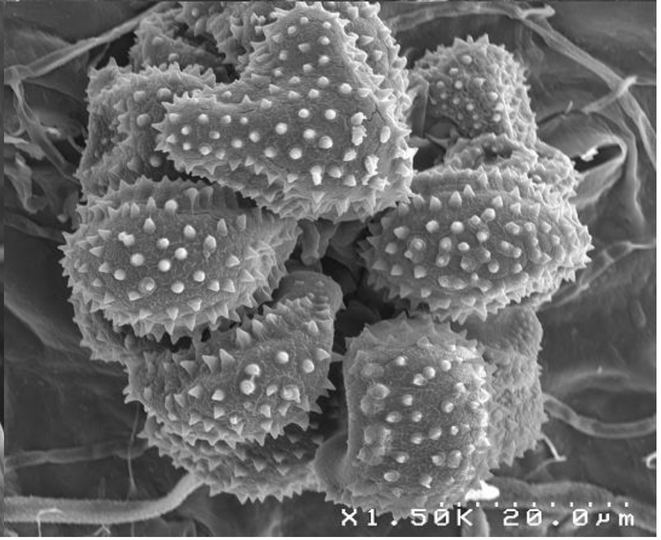
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30031			
TSH-R30034			

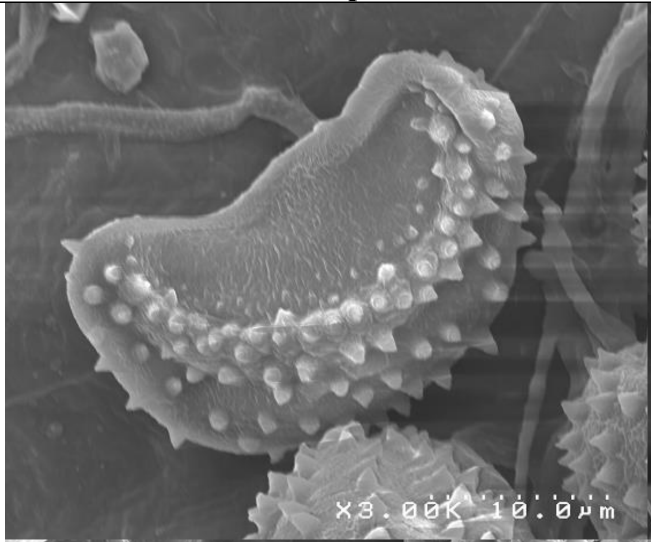
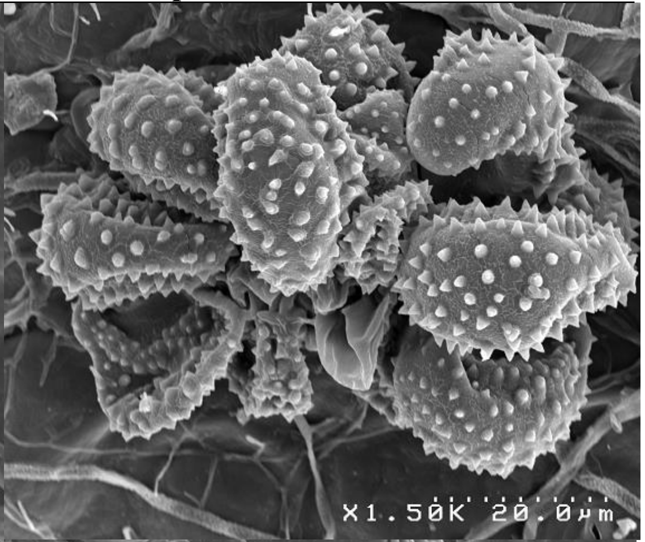
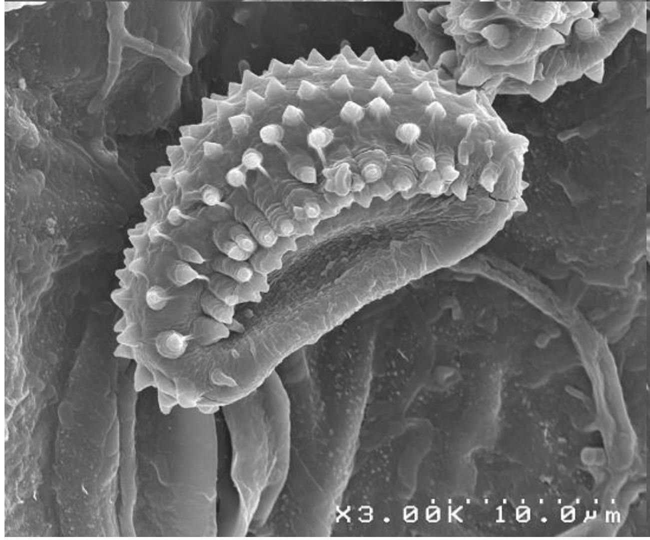
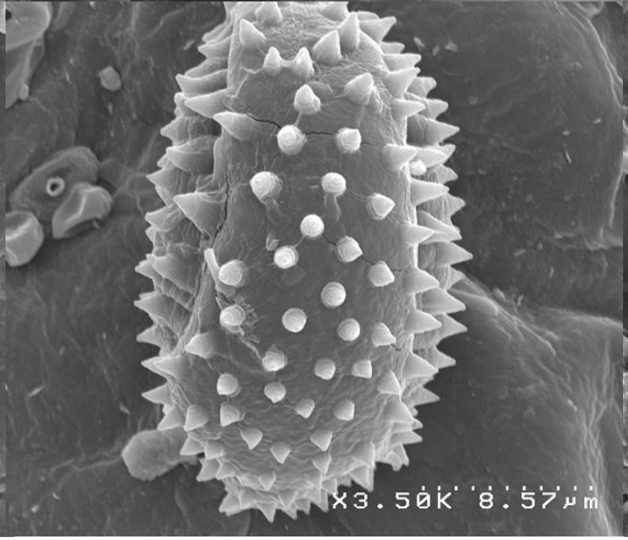
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30036			
TSH-R30042			

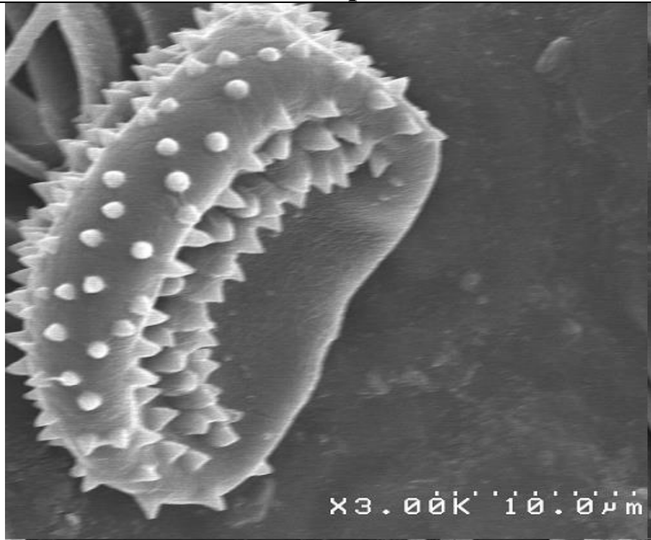
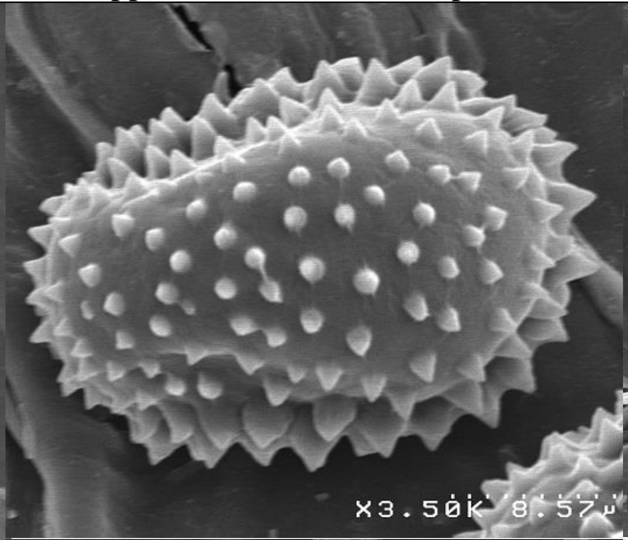
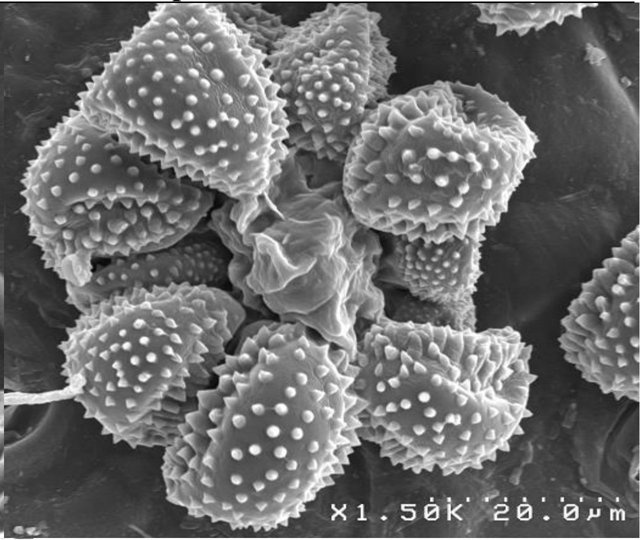
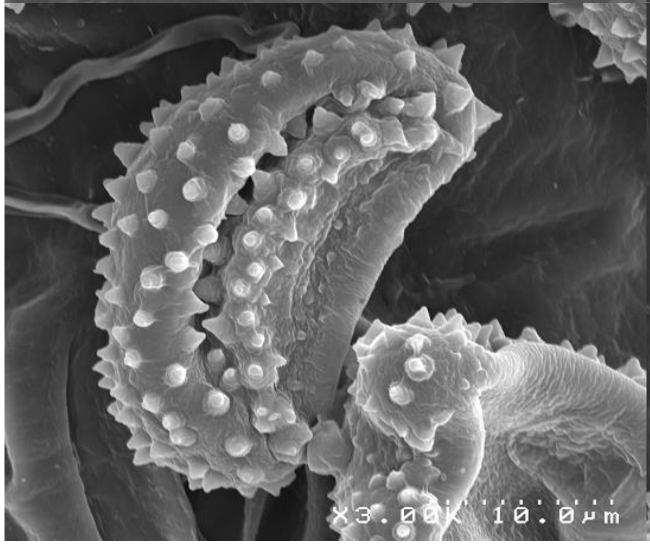
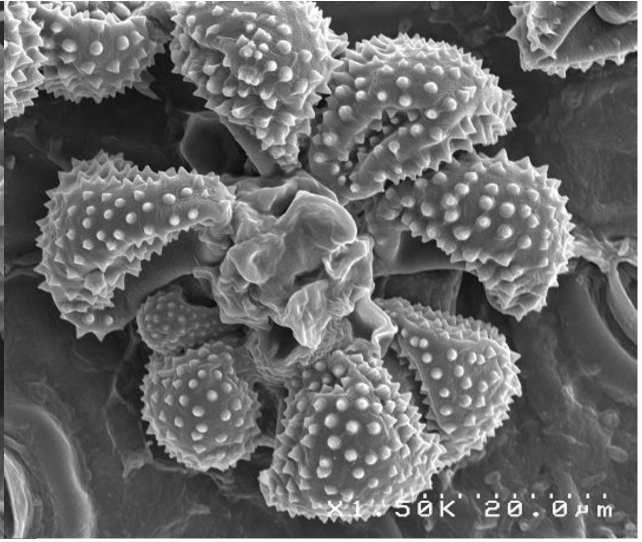
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30043			
TSH-R30045			

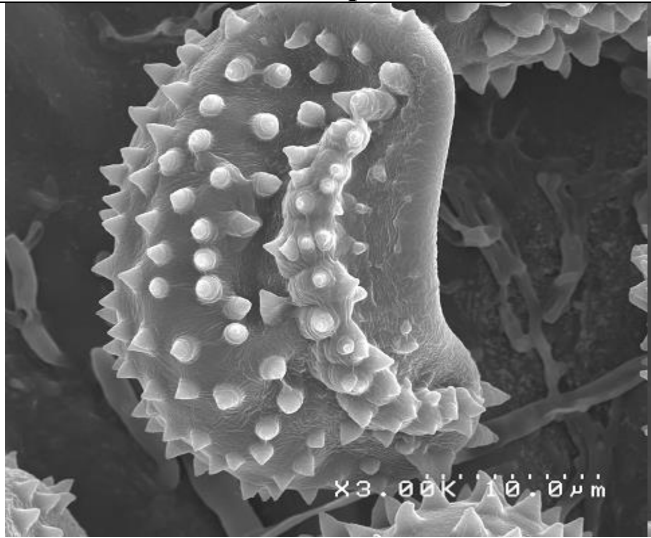
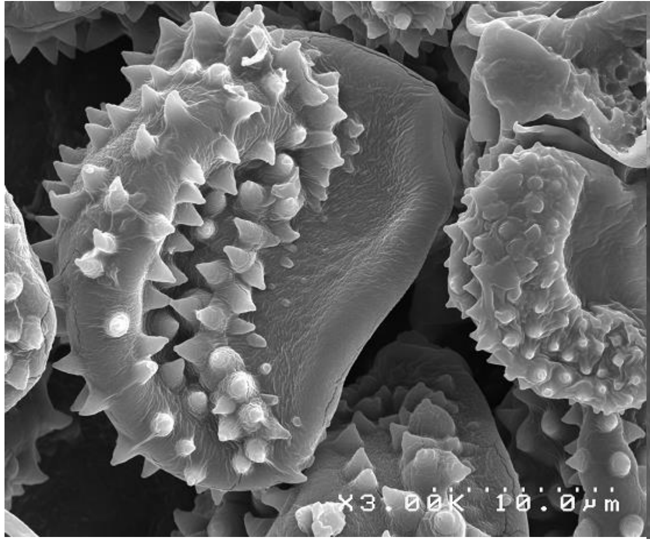
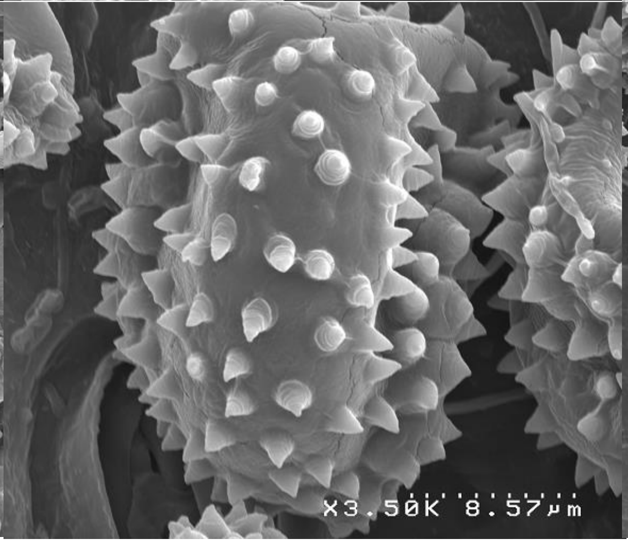
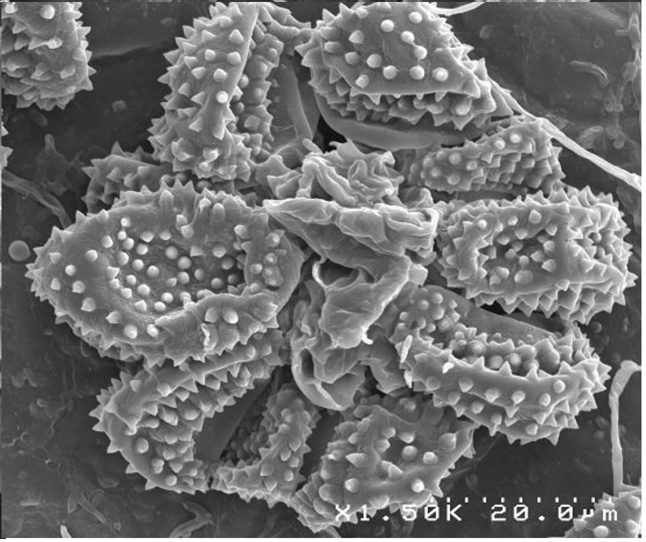
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30046			
TSH-R30050			

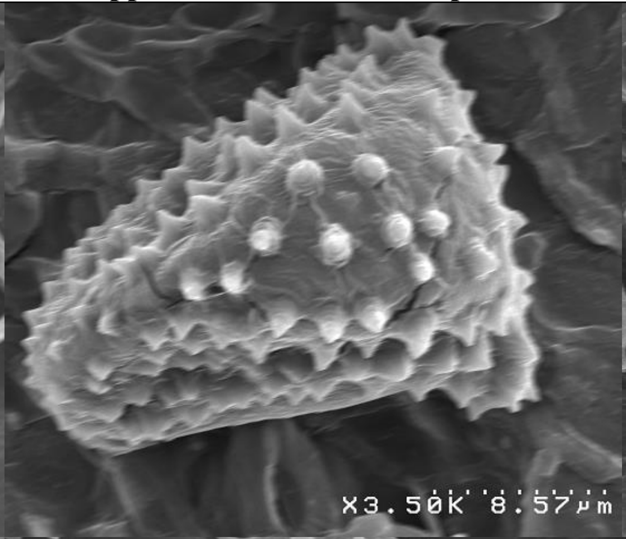
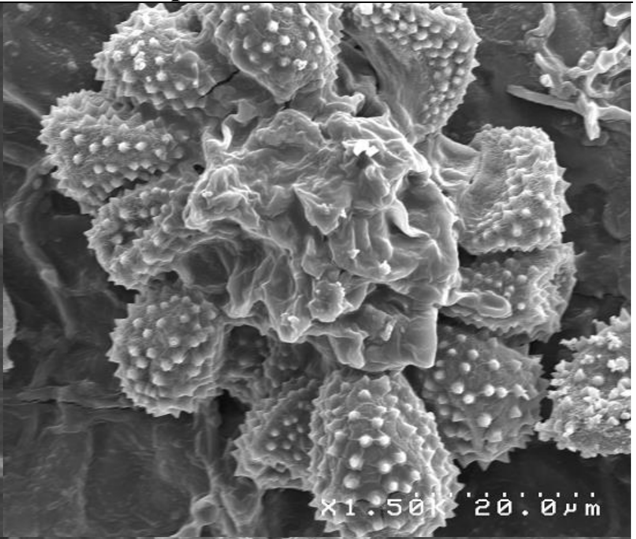
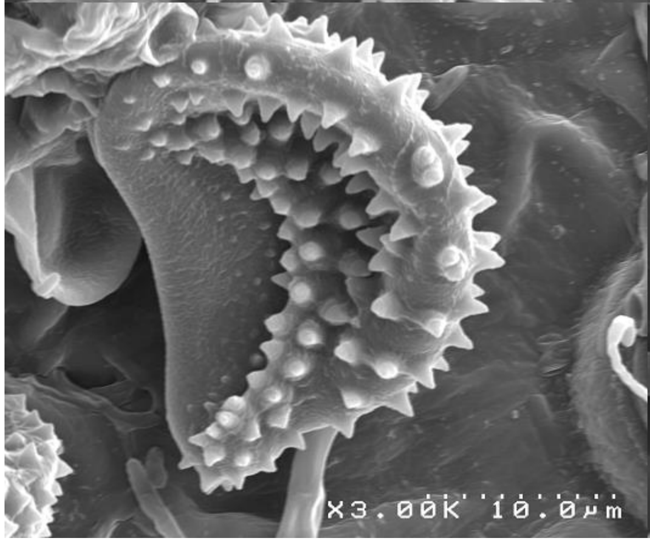
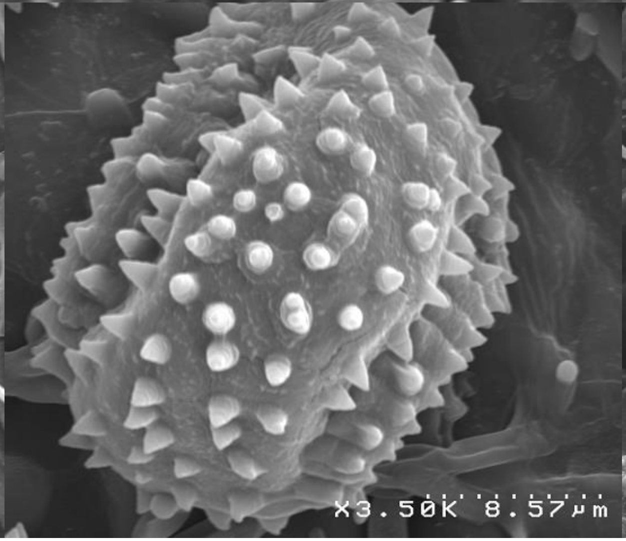
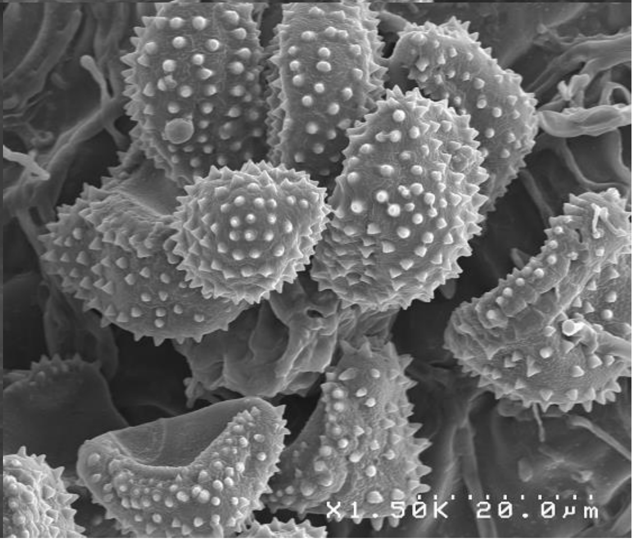
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30051			
TSH-R30052			

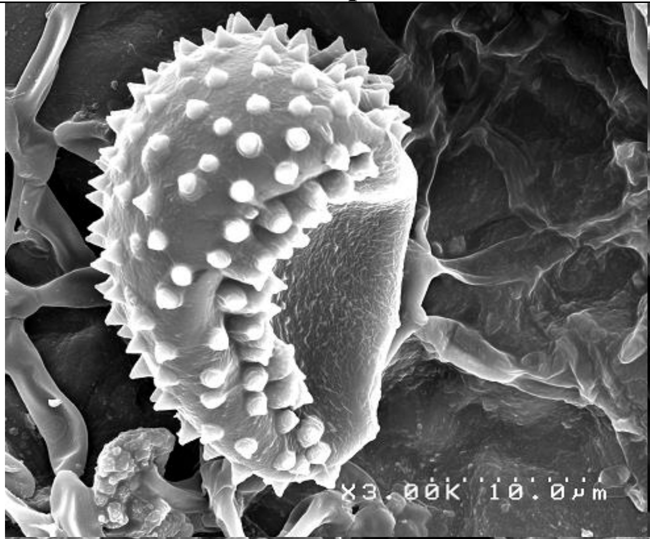
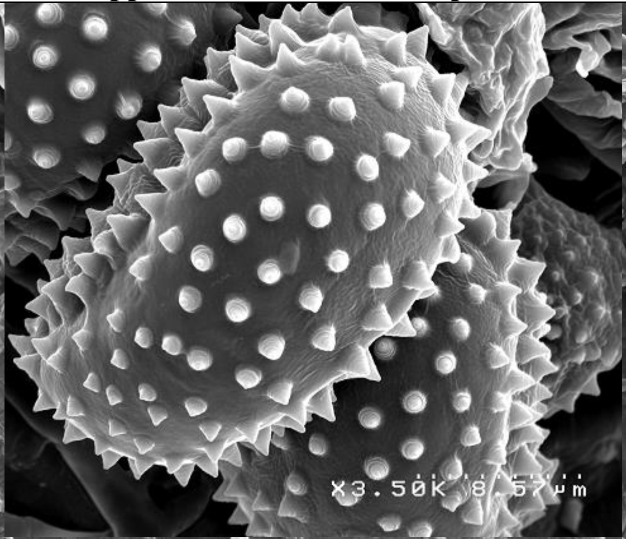
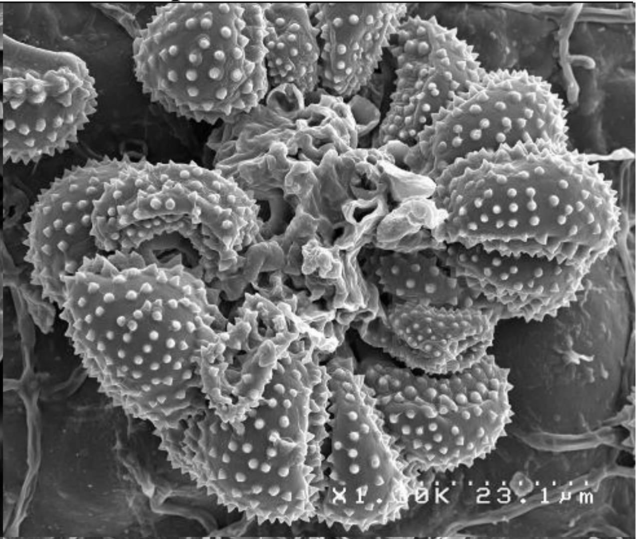
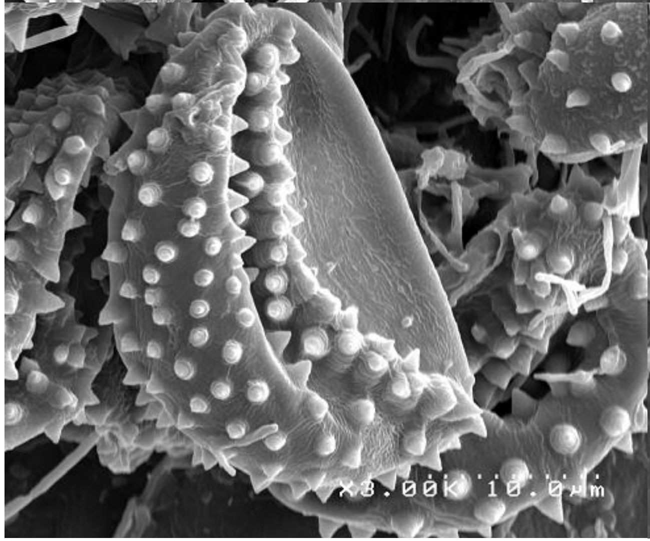
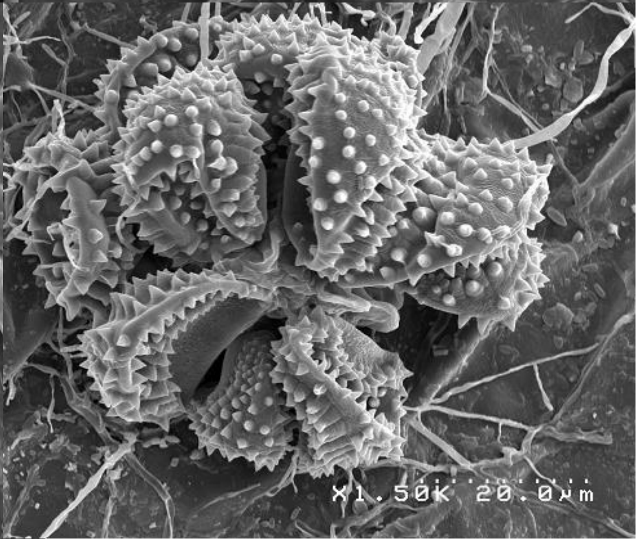
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30055			
TSH-R30059			

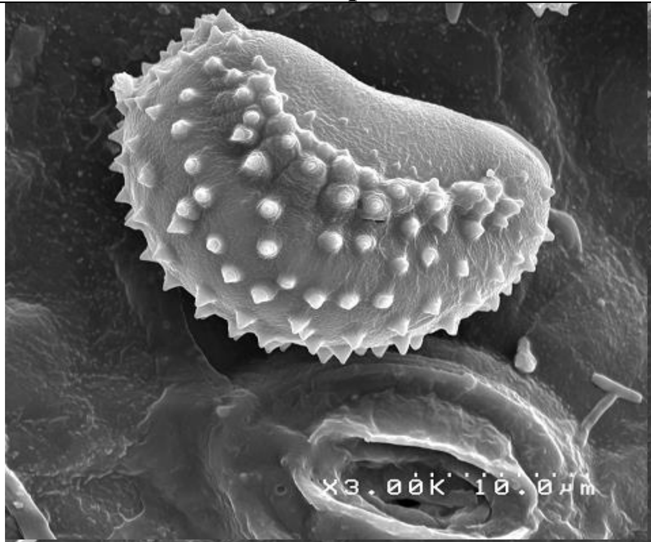
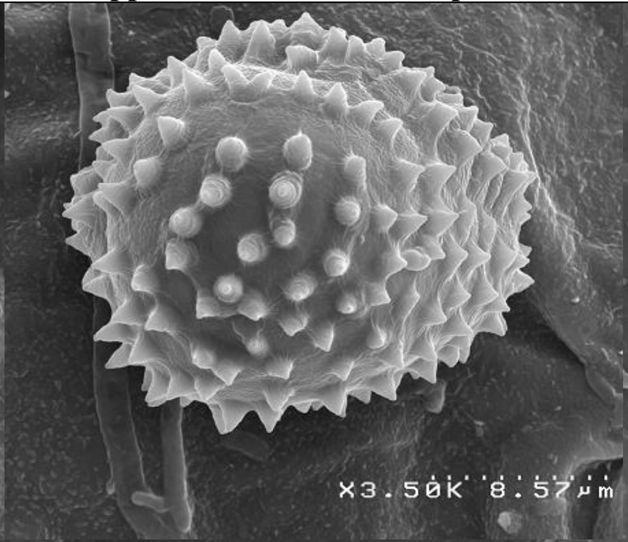
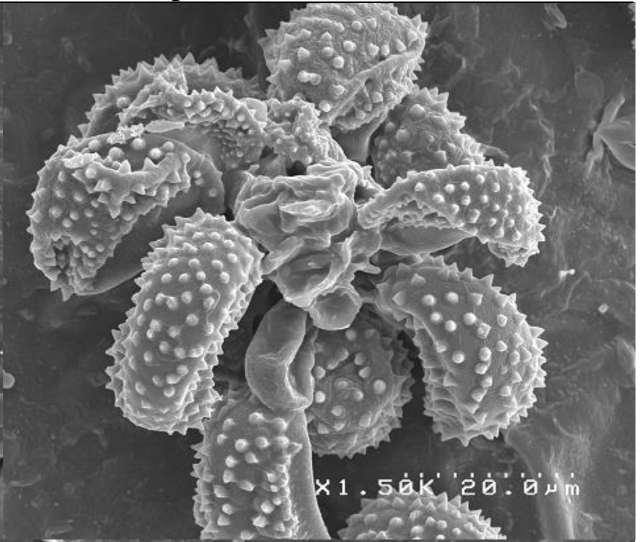
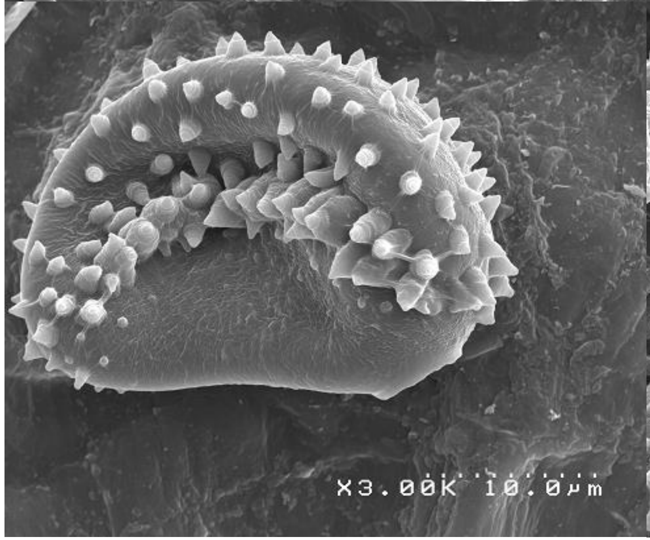
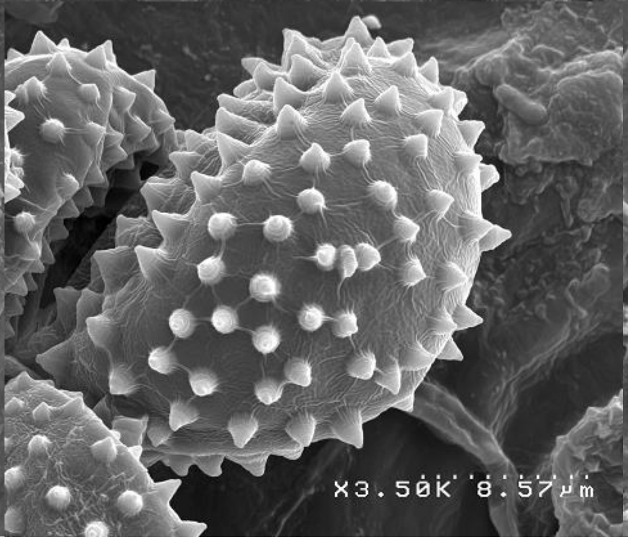
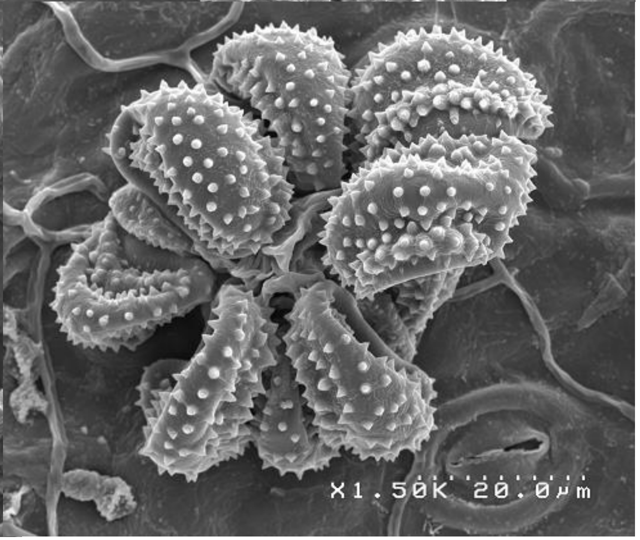
(Continued)

Specimens	**Urediniospore**	**Upper surface of urediniospore**	**Suprastomatal uredinium**
TSH-R30063			
TSH-R30067			

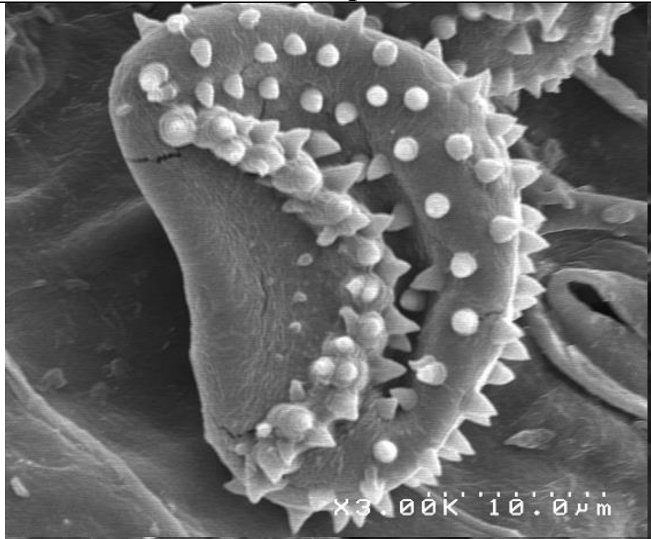
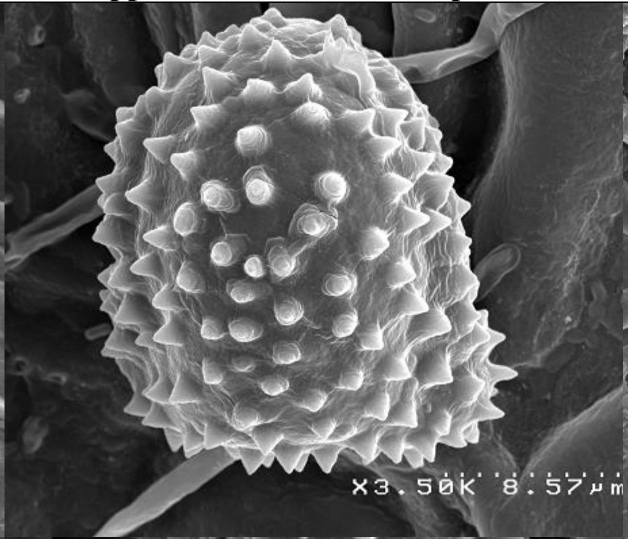
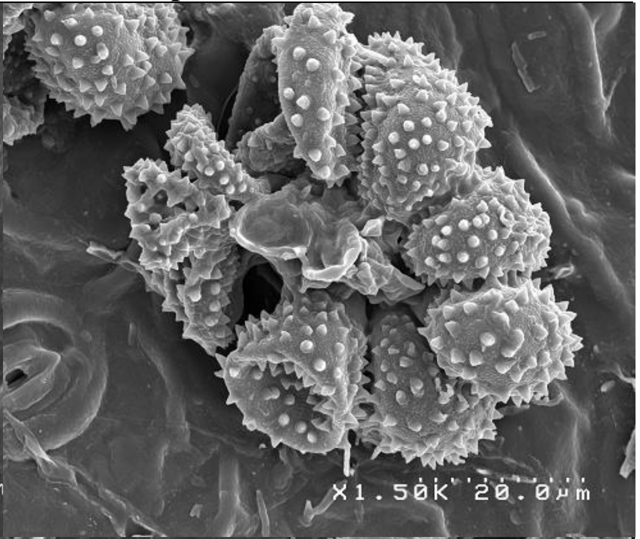
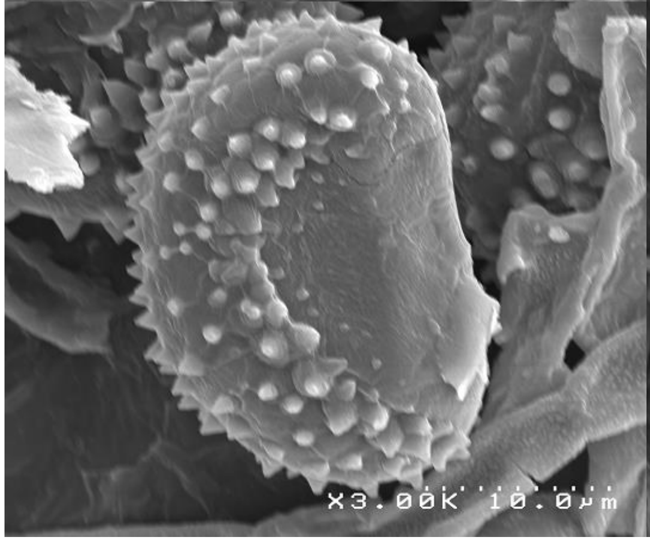
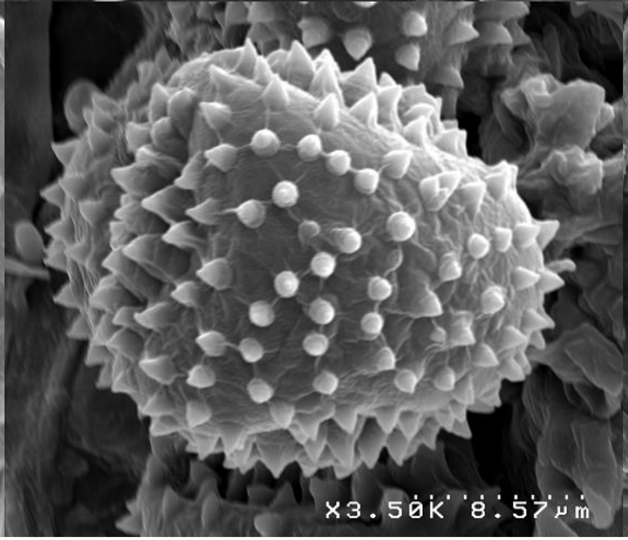
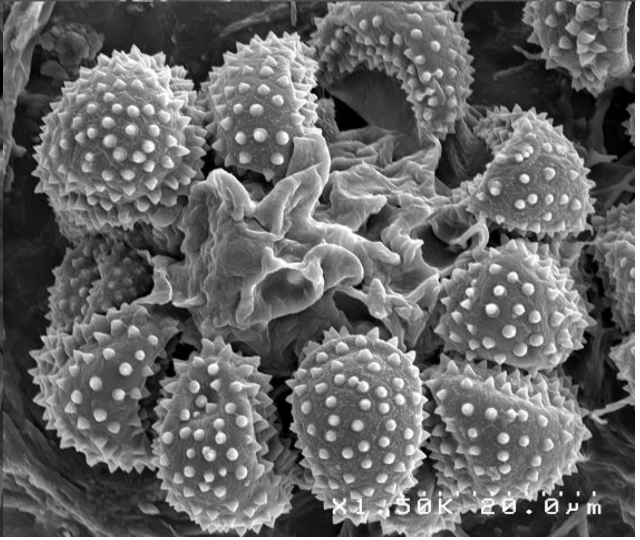
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30069			
TSH-R30071			

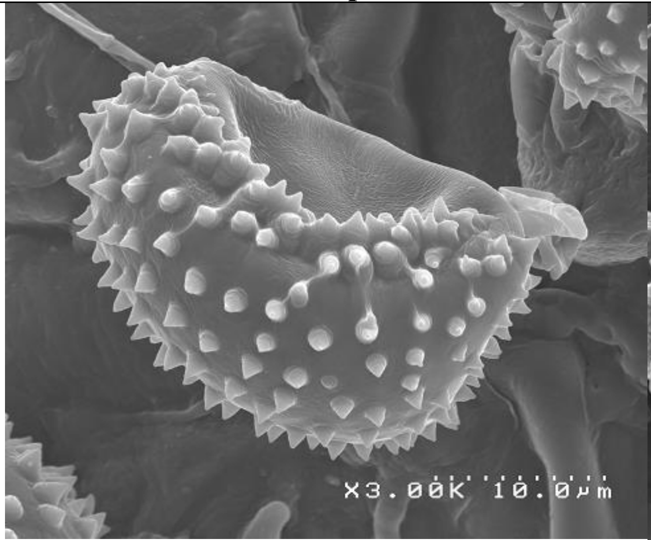
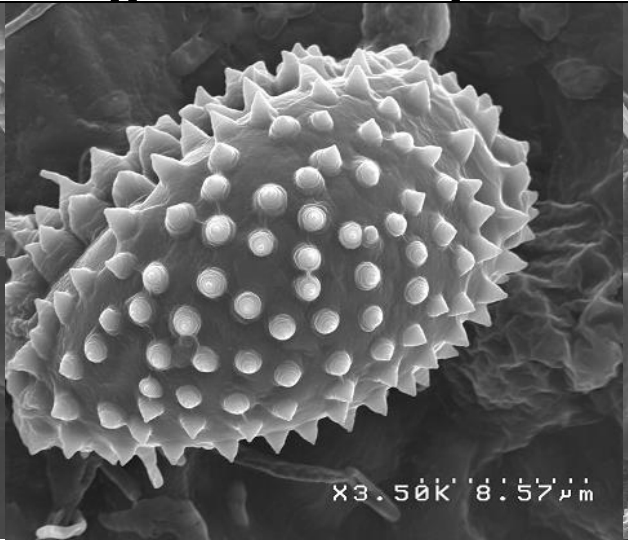
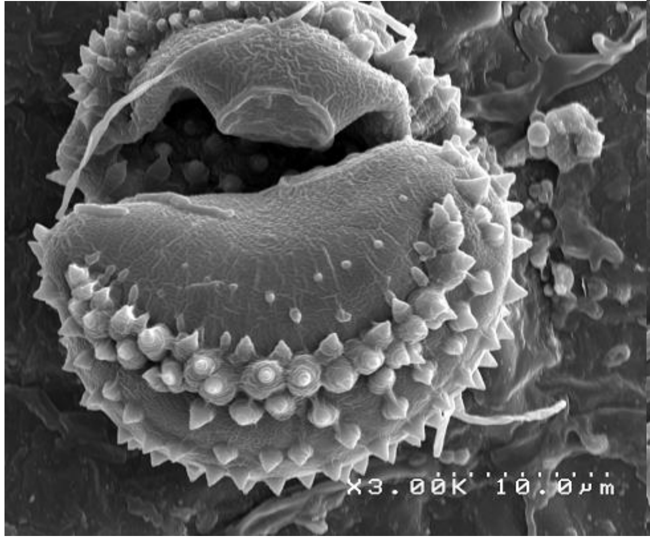
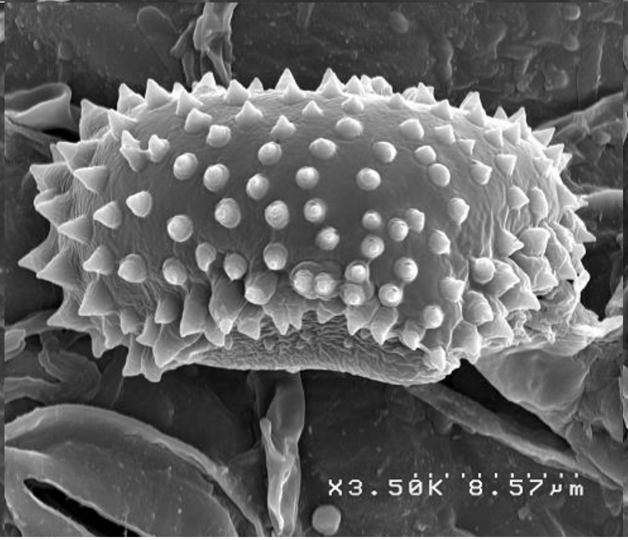
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30073			
TSH-R30080			

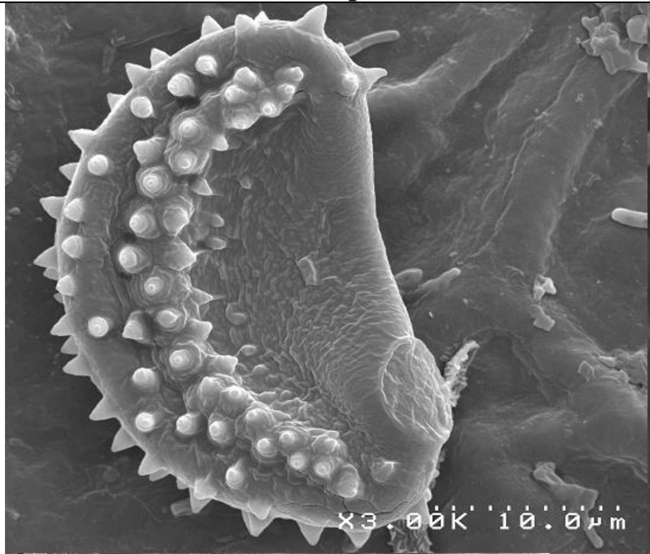
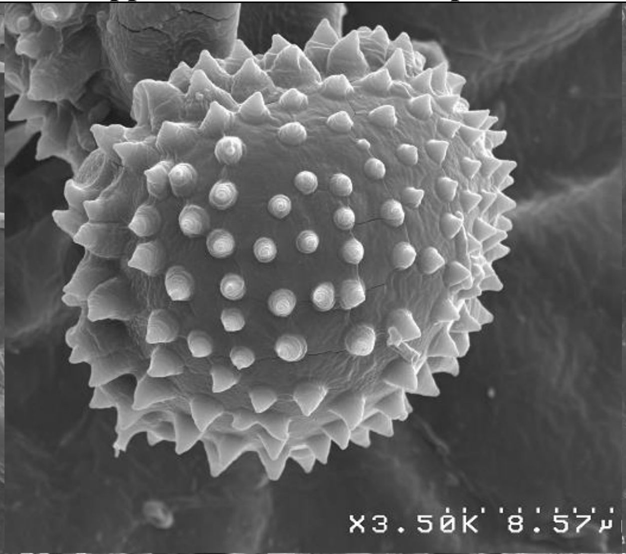
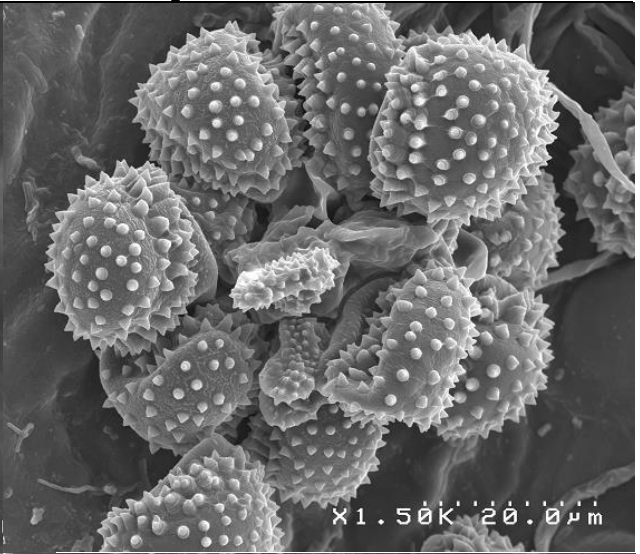
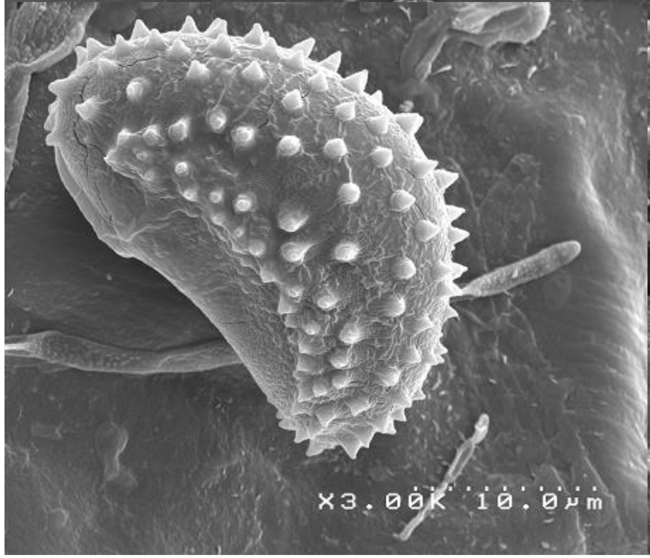
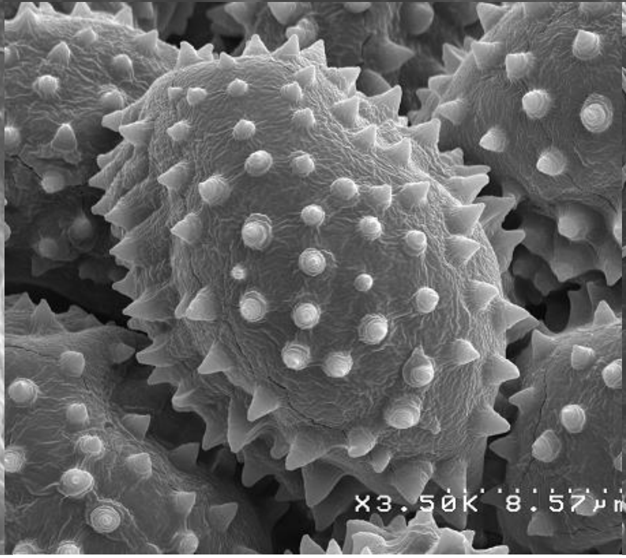
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30085			
TSH-R30088			

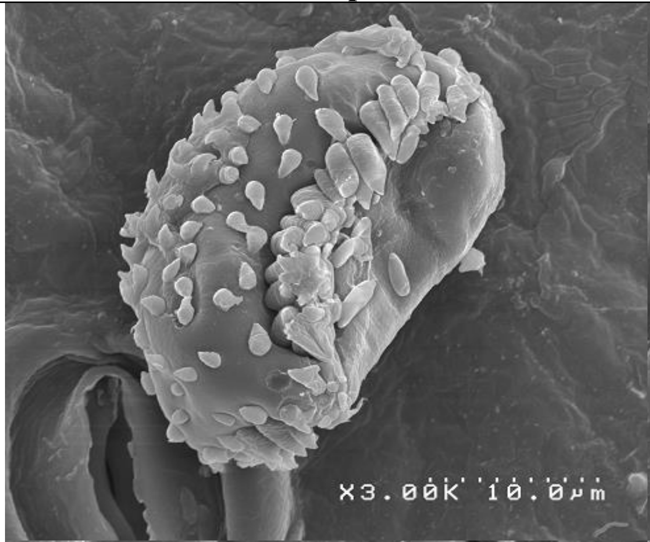
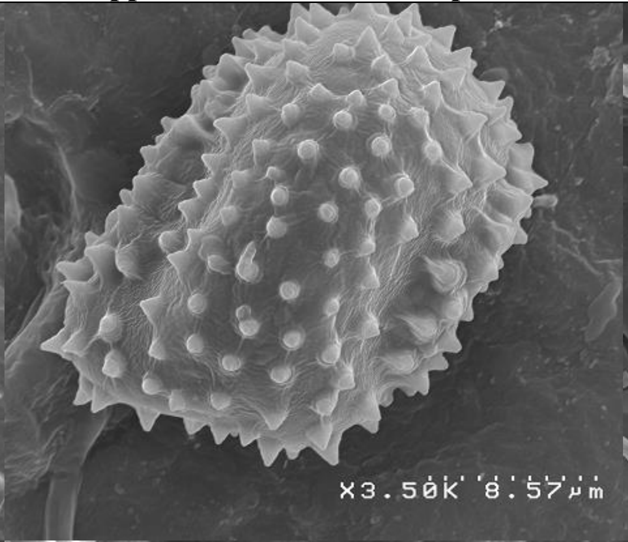
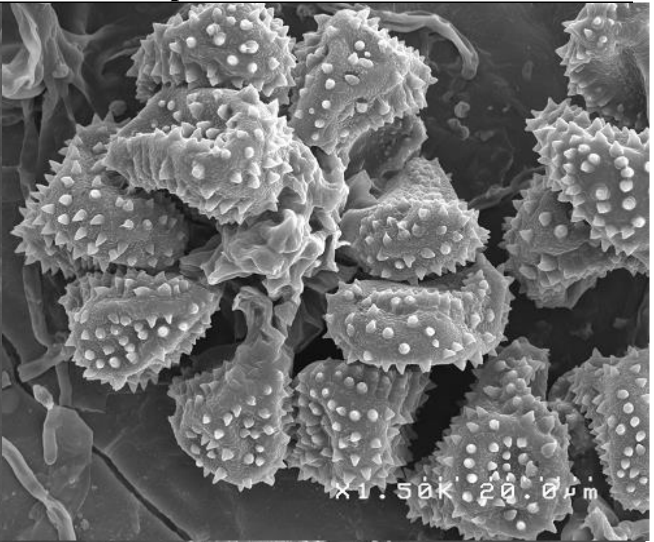
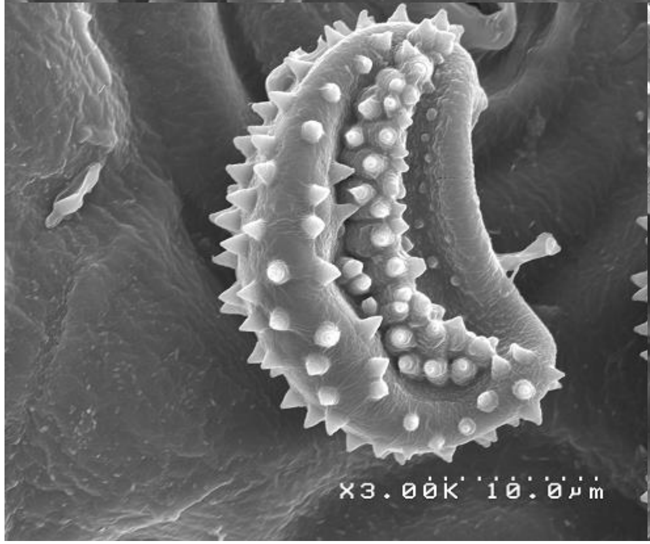
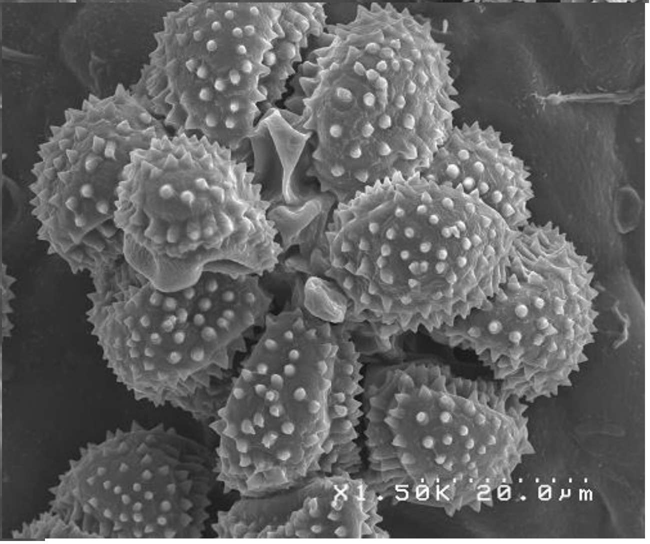
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30098			
TSH-R30101			

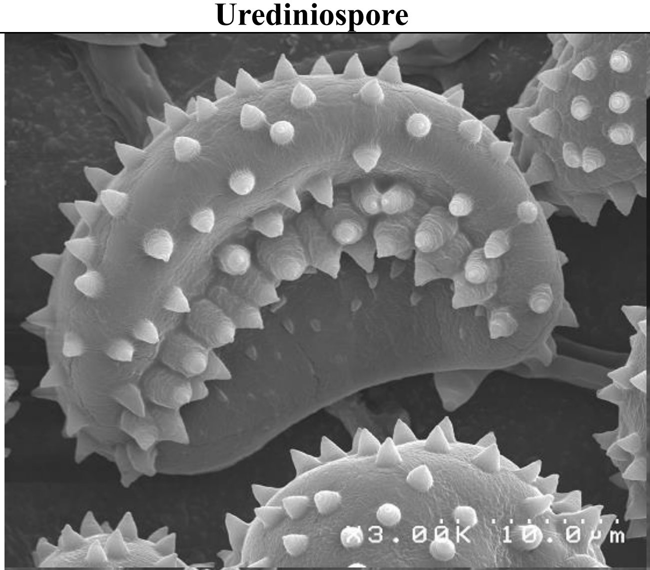
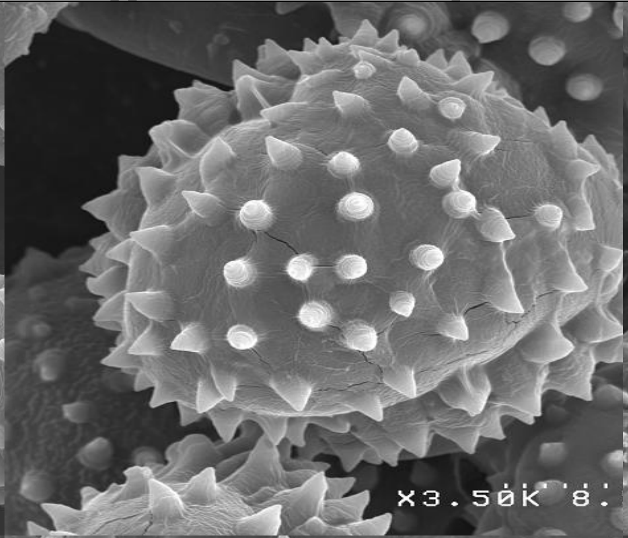
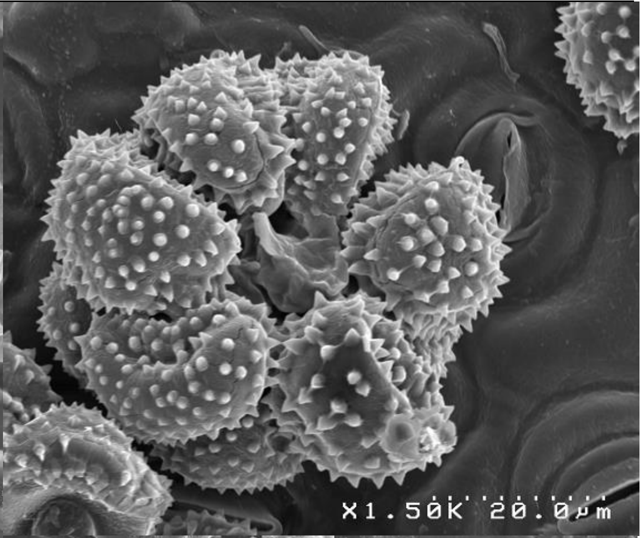
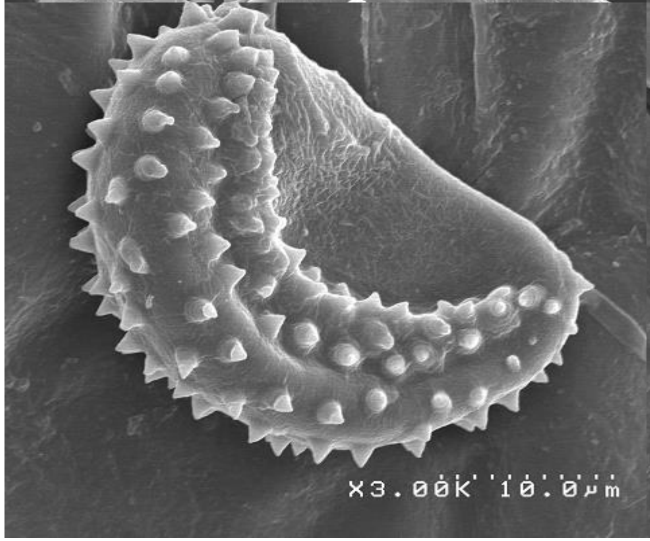
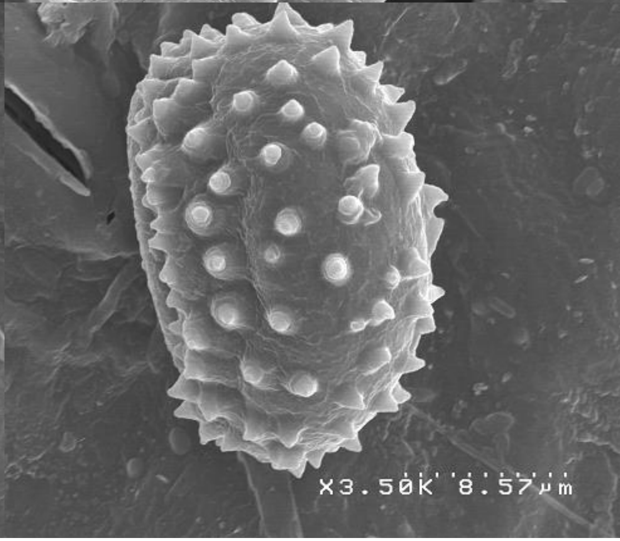
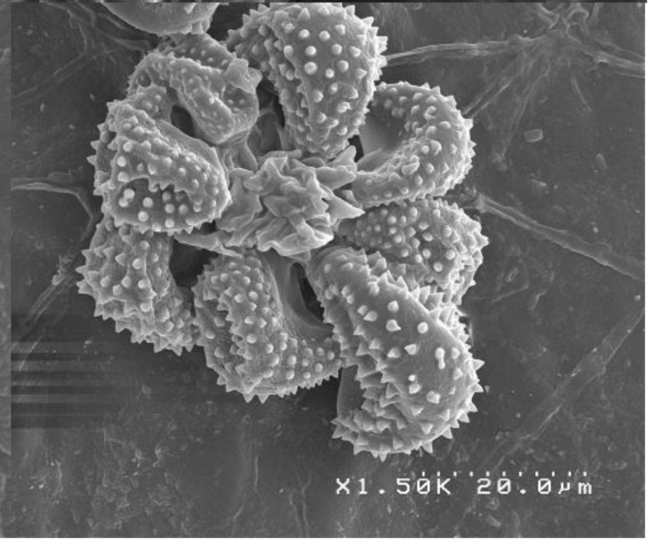
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30104			
TSH-R30107			

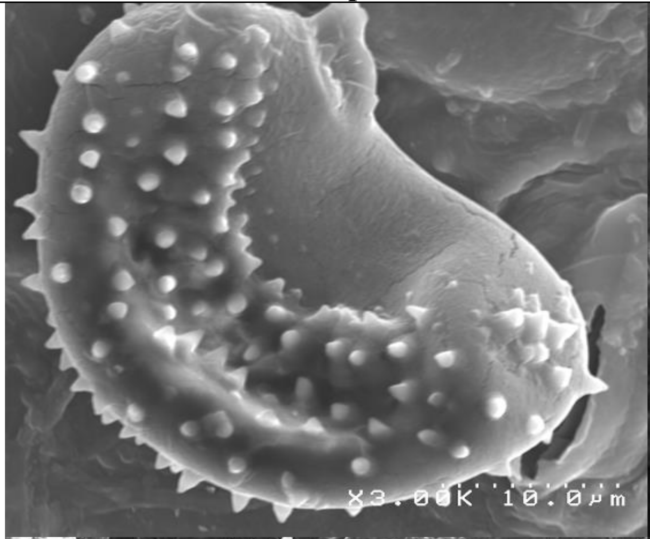
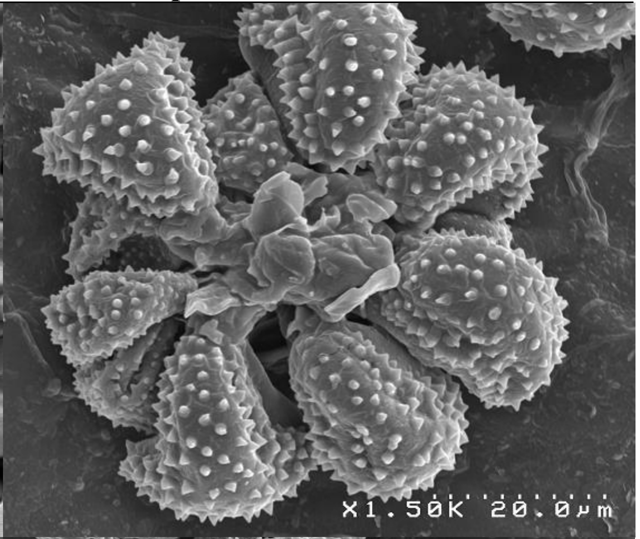
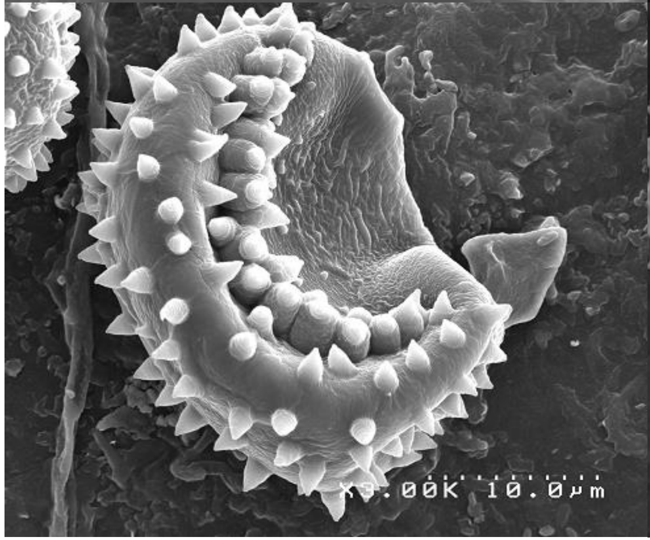
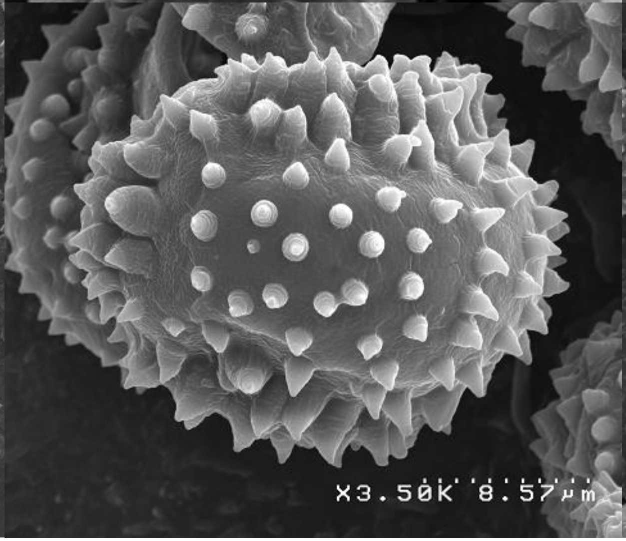
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30109			
TSH-R30118			

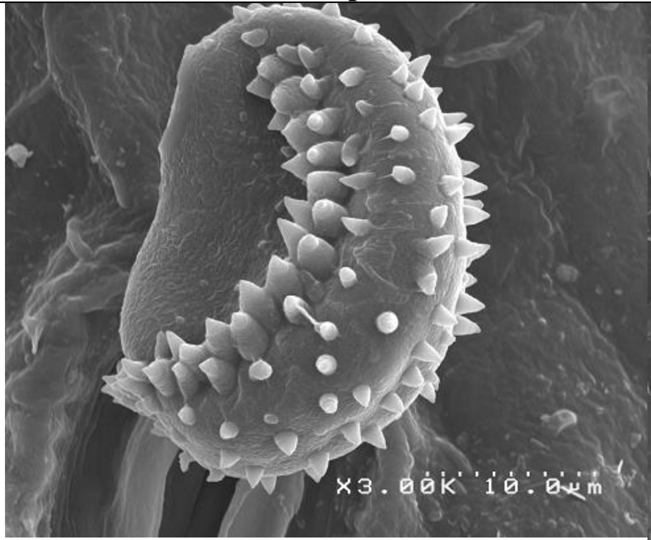
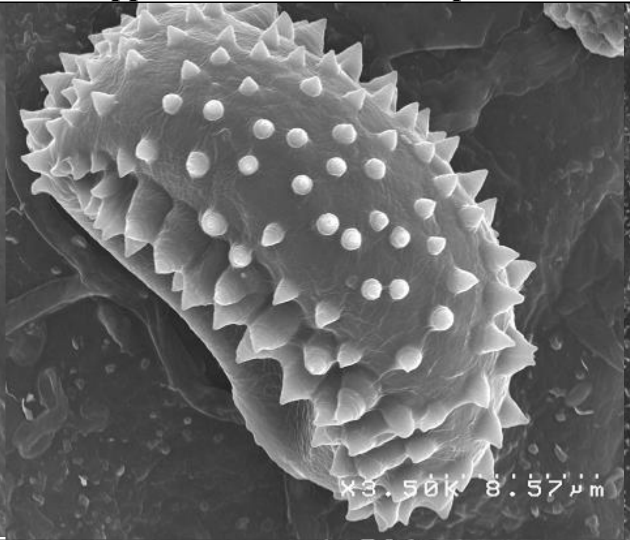
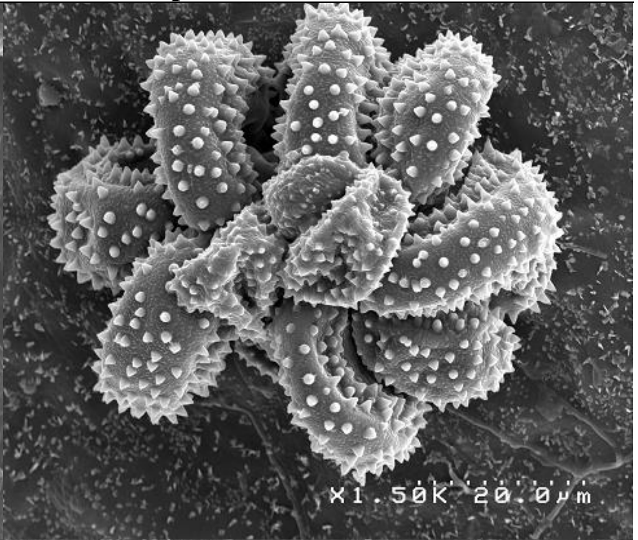
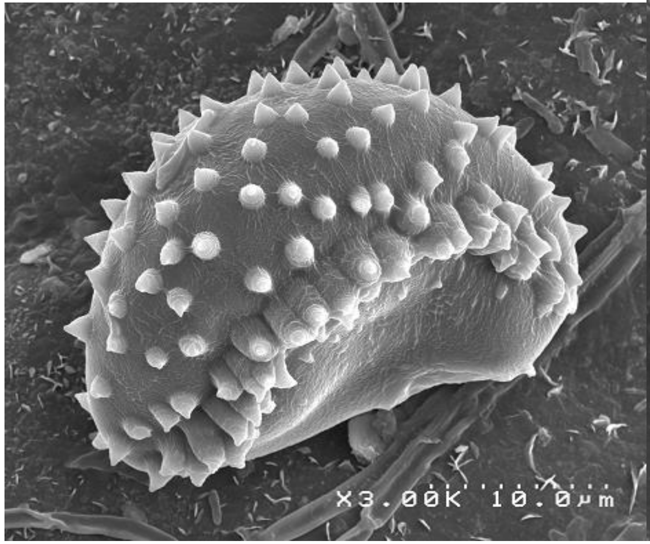
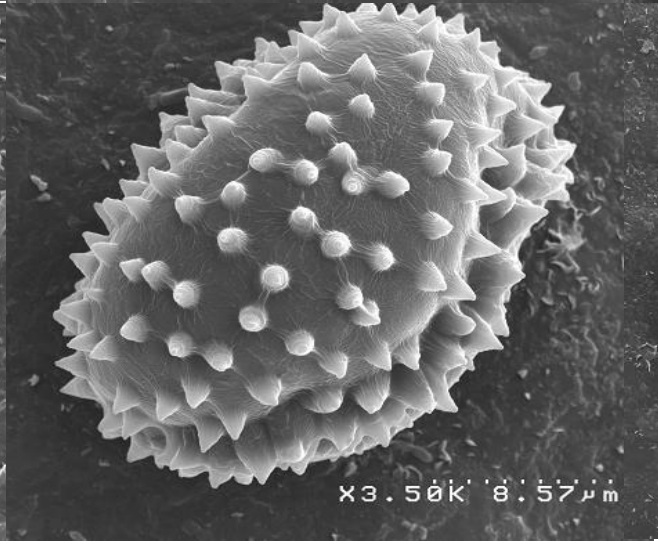
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30125			
TSH-R30129			

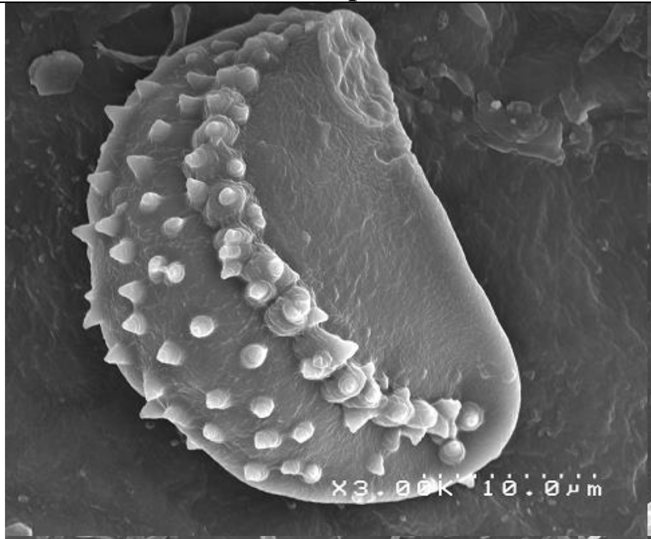
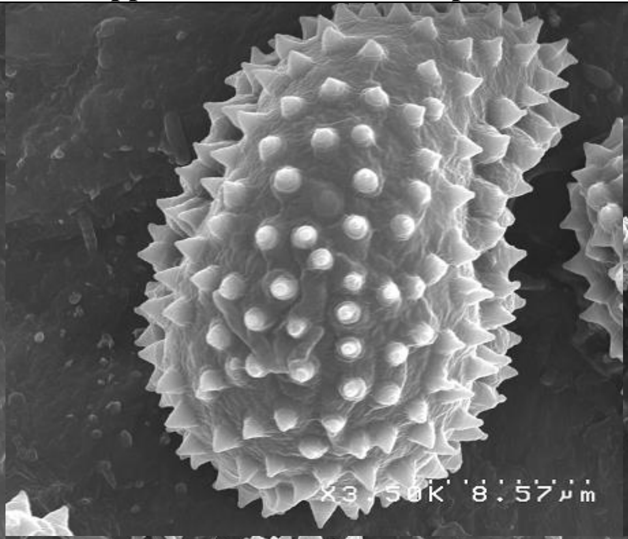
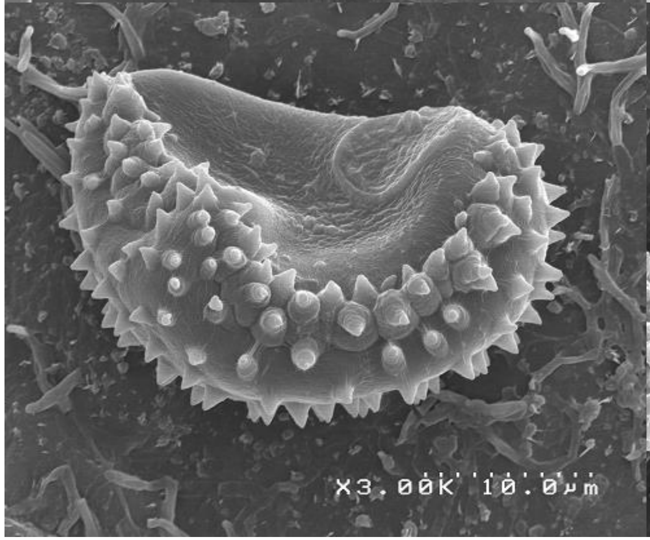
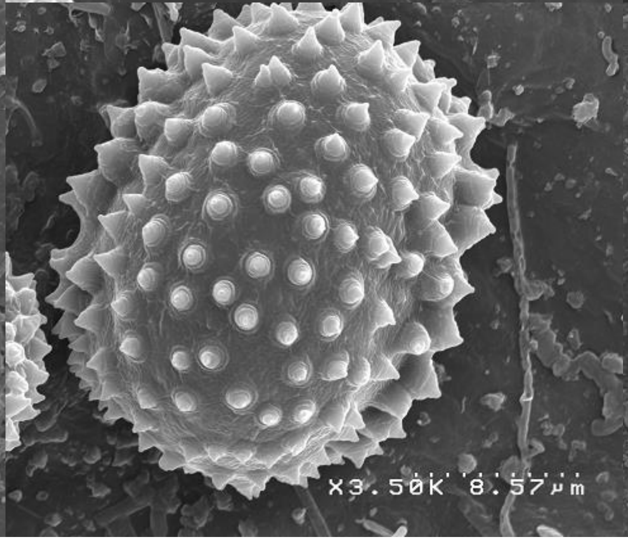
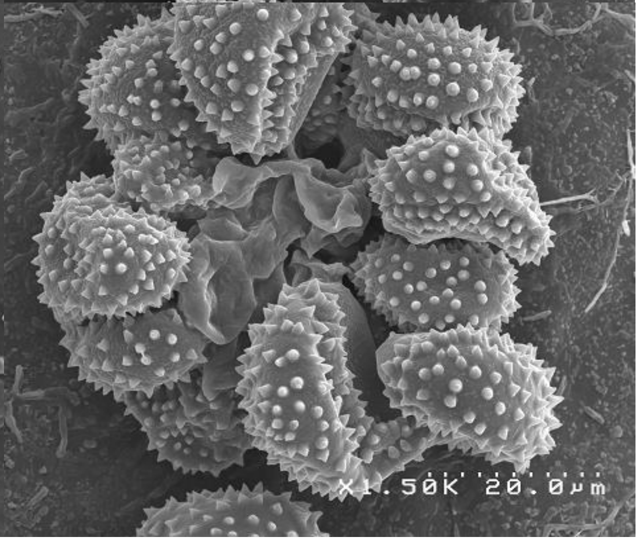
(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30133			
TSH-R30148			

(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30155			
TSH-R30160			

(Continued)

Specimens	Urediniospore	Upper surface of urediniospore	Suprastomatal uredinium
TSH-R30167			
TSH-R30172			

Appendix 3: Alignment dataset of Internal Transcribed Spacer sequences

```
>TSH-R59309
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACTGCGGC-
AATTTTTTGCTTACCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAT-CTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGAATATATGAAATGTCATTC
ATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGCTGACTTTATAAACACAT
GTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAAATAAAAAATTAAAGGGA
AAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATATGTCCCTTG-
TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R59420
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCGGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGAATGGGTCCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGGGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
```

TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30050
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACTGCGGC-
AATTTTTTGTTTACCGAATTTAACCCCTGCGGTTAGGCATATATAT-TTCTCTCG--
GAGGGTTGTATGGTTTTTAATTCTTGTTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30051
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTAATTTCCACCCAACGCC-
TTCGGGACACTGCGGC-
AATTTATGGTTTAGCGAATTTAACCCCGGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGTAGGGGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCCCAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAAAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-
CTTGCCATTATTAAGTTTGATGGGCTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-

TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-
TTCTCTTTTTATTTTGATTATCTTTATTTTTTTATATTTTAATTTTAATTTTTATGTTT
T-AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30052
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCCCAAATTTATACATTTGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCGGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTGGTTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAACCTCAAA
>TSH-R30055
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT-G-
AGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACCCAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGGGAA
TCATCGAATTTTTGAACGCATATTGGCCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTTTTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-

ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30059
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACTTC-
TTCGGGACATTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGAATG-TGTTCAATTCTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30063
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACGGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGGGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-

```
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCCTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGGTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30067
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTAATTTCCCCCCACCGCC-
TTCGGAACCCGGCGGC-ATTTAATGGTTAACCGAATTTAACCCCGGCGGTAAGGCAAA-
ATAA-TTCTCTCC--GAAGGGTGGAAGGGGTCCAATCCTTGTTTTTTTTATTTT-
CAAACACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTGGCTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30069
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
```

```
CTCACCTTAAATTTATAAGTTGTTTTTTT-
TTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGG
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30071
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTAATTTCCCCCCAACGTC-
TTCGGAACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-
TTCTCTCTTGAGGGGTGGTATGTGTTCTAATCCTTGTTTTTTTTATTTT-
CAACCCCAAATTTATACATATGTATATATGTTTTTATTTACTATCAAGTAAATAAATAT
TAAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAAAACCCAGTGAAATGGGAAAAGTAATGGGAATTGCAAAATTCAGGGAA
TCATCGAATTTTTGAACCCATATTGGCCCTTTTGGCTTTTCCAAAAGGTACCCCTGTT
TGAAAATATGAAAGGTCAGGGGGTTGAAGAAGTTTTTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTT
AGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30073
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAT-CTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGA-
```

TTTTTTTTTTCTCTTAGCATCTTGGATATTGGGTG-
CTTGCCATTATTAAGTTTGATGG-CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30080
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTAATTTACACCCAACGTC-
TTCGGGACACTGCGGCCAATTTATGGCTTAGCGAATTTAACCCCGGCGGTTAGGCAT
ATATAA-TTCTCTCT--GAGGGTGGTAGGGGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30085
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGCC-
TTCGGAACACTGCGGC-
AATTTATGGCTTACCGAATTTAACCCCGGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGAAGGTGTCCAAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-

CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30088
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTCCACCCAACGTC-
TTCGGAACACTGCGGC-
AATTAATGGCTTAGCGAATTTAACCCCGGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGTAGGGGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30098
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTTTTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTACCGAATTTAACCCCTGCGGTTAGGCATATATAT-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAATCTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-

GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30101
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTAATTTACCCCCAACGCC-
TTCGGGACACTGCGGC-
AATTTATGGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGAAGG-GTTCCAAACCTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGATCCTCAAA
>TSH-R30104
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGAAGGGGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT

```
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30107
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30109
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGAACACTGCGGC-
AATTTATTGCTTACCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTC-AATCCTTGTTTTTTTTATTTT-
CAACCCCAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGGGAATTGCAGAATTCAGTGAA
```

```
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGGGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATGGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30118
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTAATTTACCCCCAACGCC-
TTCGGGACACGGCGGC-
ATTTAATGGCTTAGCGAATTTAACCCCGGCGGTAAGGCATATATAA-TTCCCTCG--
GAGGGTGGTATGGGTTCTAACCTT-GTTTTTTTTATTTT-
CACCCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30120
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAT-
CTTGTTTTTTTTATTTTTCAACCACAAATTTATACATATGTATATATGTATT-
ATTTACTATCAAGTAAATAAATAT-AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
```

```
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30125
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACCCAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACCCATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30129
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCAAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
```

AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30133
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAT-CTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGA-
TTTTTTTTTTCTCTTAGCATCTTGGATATTGGGTG-
CTTGCCATTATTAAGTTTGATGG-CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30137
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-

```
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30145
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAT-CTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGGGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGAATTTTTTTTTTCTCTT
AGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30148
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
```

GAGGGTTGTATGTGTTCTAATC-TTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGGCCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTT
AGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30155
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAT-CTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACCCCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTT
AGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGCTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30160
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-

```
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTTTAAGTTTTATTTTATTTTTTACTTCTTTATAATGA
ATCTCAAA
>TSH-R30167
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGCCACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30172
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
```

TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30018
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAAATTCTCTCT--
GAGGGTGGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30022

TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30024
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCACCGTC-
TTCGGACACCGGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTAGGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAACCTCAAA

```
>TSH-R30028
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATCTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAACCTCAAA
>TSH-R30031
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTACGGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGTATGTGTTC-AAATCTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
```

AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30034
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTAATTTAC-
CCCAACGTCCTTCGGGACACTGCGGC-
AATTTATTGCTAACGAATTTTACCCCCTGCGGTAAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTC-AAATCTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACTTGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30036
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAT-CTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCCCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACCCTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTT
AGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-

TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30042
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACGGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGGGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAAAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-AAGTTTTATTTTATTTTTTACTTCTT-
ATAATGAATCTCAAA
>TSH-R30043
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACCCTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT

```
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>TSH-R30045
TTAAGGG-GCACTTAATTGTGGCTTGAA--TTTTCTTTATTTACCCCCAACGTT-
TTGGGGACATTG-GGC-AATTTATTGCTTAGCGAATTTAACCCCTG-
GGTTGGCCATAT--AA-TTCTCTCT-GGAGGGTTGTGTGTGTTCTAAT-
TTTGTTTTTTTTATTTT-CAACCACAAATTTATACATATGTATATATGTATT-
ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCAACTTCGATGAAGAACACAGTGAAATG
TGATAAGTAATGTGAATTGCAGAATTCAGTGAATCATCGAATTTTTGAACCCATATTG
CCCCTTTTGGCTATTCCAAAAGGCACCCCCTGTTGGGAGTATGAAAGGTCCGGGTGT
TGAGAGAGTTTTTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTTGAAGTTTTATTTTATTATTTACTTCTTTATAATGA
AACTCAAA
>TSH-R30046
TTAAGAG-GCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
```

```
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-TAGTTTTAA-----
GTTTTAATAGAACCCAATCGATATCAAA
>Mexico-1
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Mexico-2
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
```

```
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Mexico-3
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Mexico-4
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAA-
AAAAAAGAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
```

TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Mexico-5
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCGGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTT-TTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAA-
AAAAAAGAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Mexico-6
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACCCCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAAT-CTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-

```
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Mexico-7
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCGGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTT-TTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAA-
AAAAAAGAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Mexico-8
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
```

CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Mexico-9
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-1
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGTAGGTGTTCTAATCTTGGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGGGTTATTTGGTCTTGGATTTTTTTTT--

CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-2
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTGGTATGTGTTCTAATCTTGGTTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATCAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-3
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTGTTTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA

```
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-4
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-5
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
```

```
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGTAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTT-
TTTTTTCTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-6
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAA-
AAAAAGAAAAAGGTAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTT-
TTTTTTCTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAA-TAAA----
--
AAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATATGTC
CCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-7
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
```

CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGTAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTT-
ATTTTTTCTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-8
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGTAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTT-
ATTTTTTCTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAA-------
AATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-9
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-

```
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGTAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTT-
ATTTTTTCTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Colombia-10
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGTAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTT-
ATTTTTTCTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAA-------
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>CIFC-1
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
```

CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTT--
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>CIFC-2
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-TTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTT-
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>CIFC-3
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--

```
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAA-
GAAAAAGGCAAAGTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTT--
CTCTTAGCATCTTGGATATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTT-
ATTAAGGATGAAAAGTCTTTTGATGGCTTGATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-1
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAGAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-2
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
```

AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTATTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-3
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGTAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTATTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-4
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-

```
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTAAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-5
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-6
```

TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-7
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTCTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAGAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA

```
>Brazil-8
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGTGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-9
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGTAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTATTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAGTAAAAA
ATAAAAAATTAAAGAGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
```

```
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-10
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGTACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAGACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTTTTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-11
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAGAGAAGTAAAAAGAATGGGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
```

TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-12
TTAAGAGTGCACTTAATTGTGGCTCGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTGAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-13
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCAGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATTAAAA
ATTAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT

```
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-14
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTATTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAAATAAAA
AATAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-15
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTT-ATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAGAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAAATAAAA
```

AATAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-16
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGC
TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAAATAAAA
AATAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-17
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGCGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTTTGTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTTTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT

TGACTTTATAAACACATGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAAGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAATA
TGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA
>Brazil-18
TTAAGAGTGCACTTAATTGTGGCTTGAAATTTTACTTTATTTACACCCAACGTC-
TTCGGGACACTGCGGC-
AATTTATTGCTTAGCGAATTTAACCCCTGTGGTTAGGCATATATAA-TTCTCTCT--
GAGGGTTGTATGTGTTCTAATCTT-GTTTTTTTTATTTT-
CAACCACAAATTTATACATATGTATATATGTATT-ATTTACTATCAAGTAAATAAATAT-
AAAACTTTTAACAATGGATCTCTTGGCTCTCA-
CATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAA
TCATCGAATTTTTGAACGCATATTGCGCCTTTTGGCTATTCCAAAAGGTACACCTGTT
TGAGAGTATGAAAGGTCAGGGTGTTGAGAGAGTTATTAAAAAAAAAGAAAAAGGCAAA
GTAACACTTTAAGTGTTATTTTGTCTTTGATTTTTTTTTTTCTCTTAGCATCTTGGAT
ATTGGGTG-CTTGCCATTATTAAGTTTGATGG-
CTCACCTTAAATTTATAAGTTGTTTTTTTATTAAGGATGAAAAGTCTTTTGATGGCTT
GATGTTATTGA-
TATATGAAATGTCATTCATCAAGAAATCAGGGGGGTGACTAACCTTGATGAGAAATGT
TGACTTTATAAACACAAGTTAAAAGATAAATAAAAAGTAAAAGATAAAAAAAATAAAAA
ATAAAAAATTAAAGGGAAGGAAGTAAAAAGAATGAGAGGTATGACATATTTTGAAAAT
ATGTCCCTTG-TTCTCTTTTTATTTTGATTATCTTTATTTTTTTAT-
TTTTAATTTTAATTTTTATGTTTT-
AAGTTTTATTTTATTTTTTACTTCTTTATAATGAATCTCAAA

Appendix 4: Input file for PopART software

```
#NEXUS
[File generated by DnaSP Ver. 6.12.03;		Feb 25, 2021]
[Haplotype Data from Data File: ITSSEQCOM.phy]
[Number of sequences: 83	Number of sequences used: 83]
[Number of Sites in the complete Data File: 886]
[Selected Region:	1-886]
[Sites with alignment gaps:	not considered]
[Invariable sites:	removed]

BEGIN TAXA;
DIMENSIONS NTAX=52;
TAXLABELS
Hap_1
Hap_2
Hap_3
Hap_4
Hap_5
Hap_6
Hap_7
Hap_8
Hap_9
Hap_10
Hap_11
Hap_12
Hap_13
Hap_14
Hap_15
Hap_16
Hap_17
Hap_18
Hap_19
Hap_20
Hap_21
Hap_22
Hap_23
Hap_24
Hap_25
Hap_26
Hap_27
Hap_28
Hap_29
Hap_30
Hap_31
Hap_32
Hap_33
Hap_34
Hap_35
Hap_36
Hap_37
Hap_38
Hap_39
Hap_40
Hap_41
Hap_42
Hap_43
Hap_44
Hap_45
Hap_46
Hap_47
Hap_48
Hap_49
Hap_50
Hap_51
```

```
Hap_52;
END;

[Hap#    Freq. Sequences]
[Hap_1:    1        1]
[Hap_2:    1        2]
[Hap_3:    1        3]
[Hap_4:    1        4]
[Hap_5:    1        5]
[Hap_6:    1        6]
[Hap_7:    1        7]
[Hap_8:    15        8 15-16 27 35-37 41 43-45 59 63-64 71]
[Hap_9:    2        9 13]
[Hap_10:    1        10]
[Hap_11:    1        11]
[Hap_12:    1        12]
[Hap_13:    1        14]
[Hap_14:    1        17]
[Hap_15:    1        18]
[Hap_16:    1        19]
[Hap_17:    9        20-24 26 46 57-58]
[Hap_18:    1        25]
[Hap_19:    1        28]
[Hap_20:    1        29]
[Hap_21:    1        30]
[Hap_22:    1        31]
[Hap_23:    1        32]
[Hap_24:    2        33-34]
[Hap_25:    1        38]
[Hap_26:    1        39]
[Hap_27:    2        40 61]
[Hap_28:    1        42]
[Hap_29:    1        47]
[Hap_30:    1        48]
[Hap_31:    1        49]
[Hap_32:    7        50-54 56 60]
[Hap_33:    1        55]
[Hap_34:    1        62]
[Hap_35:    1        65]
[Hap_36:    1        66]
[Hap_37:    1        67]
[Hap_38:    1        68]
[Hap_39:    1        69]
[Hap_40:    1        70]
[Hap_41:    1        72]
[Hap_42:    1        73]
[Hap_43:    1        74]
[Hap_44:    1        75]
[Hap_45:    1        76]
[Hap_46:    1        77]
[Hap_47:    1        78]
[Hap_48:    1        79]
[Hap_49:    1        80]
[Hap_50:    1        81]
[Hap_51:    1        82]
[Hap_52:    1        83]

[Hap#    Freq. Sequences]
[Hap_1:    1        H1_11]
[Hap_2:    1        S2_14]
[Hap_3:    1        H2_6]
[Hap_4:    1        H3_12]
[Hap_5:    1        N1_11]
[Hap_6:    1        S1_18]
[Hap_7:    1        18TH13]
```

```
[Hap_8:    15      Mexico_6 N2_4 N2_6 Brazil_8 Brazil_1 Mexico_4 Mexico_8 Colombia_3 Brazil_5 Brazil_6 Colombia_4
Brazil_14 S1_5 H3_11 N1_9]
[Hap_9:    2       N1_4 H2_7]
[Hap_10:   1       N2-1]
[Hap_11:   1       N1_8]
[Hap_12:   1       N1_10]
[Hap_13:   1       N1_7]
[Hap_14:   1       N2_5]
[Hap_15:   1       H1_10]
[Hap_16:   1       S1_11]
[Hap_17:   9       Mexico_3 Mexico_2 Mexico_1 Mexico_9 Colombi_11 Colombi_12 Colombi_13 Brazil_15 Brazil_16]
[Hap_18:   1       Brazil_12]
[Hap_19:   1       Brazil_18]
[Hap_20:   1       H2_5]
[Hap_21:   1       Brazil_7]
[Hap_22:   1       Brazil_11]
[Hap_23:   1       Brazil_17]
[Hap_24:   2       Mexico_5 Mexico_7]
[Hap_25:   1       N1_5]
[Hap_26:   1       Colombia_1]
[Hap_27:   2       Colombia_2 S1_2]
[Hap_28:   1       Brazil_4]
[Hap_29:   1       Brazil_2]
[Hap_30:   1       Brazil_10]
[Hap_31:   1       Brazil_13]
[Hap_32:   7       Colombia_5 Brazil_3 Colombia_7 Colombia_8 Colombi_10 Colombia_9 Colombia_6]
[Hap_33:   1       Brazil_9]
[Hap_34:   1       S2-15]
[Hap_35:   1       S1_14]
[Hap_36:   1       S1_15]
[Hap_37:   1       H3_8]
[Hap_38:   1       H2_2]
[Hap_39:   1       S1_9]
[Hap_40:   1       S2_6]
[Hap_41:   1       S2_7]
[Hap_42:   1       H1_8]
[Hap_43:   1       H1_12]
[Hap_44:   1       H3_10]
[Hap_45:   1       19TH46]
[Hap_46:   1       H1_9]
[Hap_47:   1       H3_3]
[Hap_48:   1       H2_8]
[Hap_49:   1       H2_4]
[Hap_50:   1       H3_14]
[Hap_51:   1       H3_9]
[Hap_52:   1       H2_9]

BEGIN CHARACTERS;
DIMENSIONS NCHAR=123;
FORMAT DATATYPE=DNA   MISSING=? GAP=- MATCHCHAR=. INTERLEAVE ;
CHARLABELS
 [1]   nt_6     [2]   nt_25    [3]   nt_34    [4]   nt_39    [5]    nt_43

 [6]   nt_48    [7]   nt_50    [8]   nt_52    [9]   nt_53    [10]   nt_54

 [11]  nt_58    [12]  nt_59    [13]  nt_61    [14]  nt_62    [15]   nt_63

 [16]  nt_64    [17]  nt_65    [18]  nt_66    [19]  nt_74    [20]   nt_77

 [21]  nt_78    [22]  nt_80    [23]  nt_82    [24]  nt_84    [25]   nt_86

 [26]  nt_87    [27]  nt_88    [28]  nt_90    [29]  nt_95    [30]   nt_100

 [31]  nt_106   [32]  nt_107   [33]  nt_109   [34]  nt_112   [35]   nt_118

 [36]  nt_123   [37]  nt_132   [38]  nt_135   [39]  nt_136   [40]   nt_138
```

```
[41]   nt_139   [42]   nt_140   [43]   nt_141   [44]   nt_143   [45]    nt_144
[46]   nt_145   [47]   nt_146   [48]   nt_148   [49]   nt_150   [50]    nt_172
[51]   nt_173   [52]   nt_175   [53]   nt_189   [54]   nt_201   [55]    nt_219
[56]   nt_237   [57]   nt_264   [58]   nt_273   [59]   nt_274   [60]    nt_277
[61]   nt_288   [62]   nt_291   [63]   nt_300   [64]   nt_309   [65]    nt_317
[66]   nt_338   [67]   nt_346   [68]   nt_347   [69]   nt_358   [70]    nt_364
[71]   nt_369   [72]   nt_372   [73]   nt_375   [74]   nt_376   [75]    nt_377
[76]   nt_381   [77]   nt_382   [78]   nt_384   [79]   nt_396   [80]    nt_400
[81]   nt_406   [82]   nt_407   [83]   nt_408   [84]   nt_413   [85]    nt_433
[86]   nt_442   [87]   nt_450   [88]   nt_458   [89]   nt_464   [90]    nt_478
[91]   nt_594   [92]   nt_655   [93]   nt_666   [94]   nt_669   [95]    nt_671
[96]   nt_672   [97]   nt_729   [98]   nt_732   [99]   nt_733   [100]   nt_749
[101]  nt_753   [102]  nt_767   [103]  nt_780   [104]  nt_790   [105]   nt_827
[106]  nt_840   [107]  nt_845   [108]  nt_852   [109]  nt_853   [110]   nt_859
[111]  nt_860   [112]  nt_865   [113]  nt_867   [114]  nt_868   [115]   nt_869
[116]  nt_870   [117]  nt_872   [118]  nt_873   [119]  nt_877   [120]   nt_878
[121]  nt_879   [122]  nt_880   [123]  nt_881
;

MATRIX
[                 10         20         30         40         50]
[                  *          *          *          *          *]
Hap_1     ATAAACAGTCCGGACACTATATCTGCGAATTAGTATGTTTATGGTTCATA
Hap_2     G.T......TG.....T.............GC.......G.........
Hap_3     ....C.......A.........................G...........
Hap_4     ............A...........C.........................
Hap_5     ..................................................
Hap_6     ..................................................
Hap_7     ....................T..C.........................
Hap_8     ..................................................
Hap_9     ..................................................
Hap_10    ..................................................
Hap_11    ..................................................
Hap_12    ..................................................
Hap_13    .....A............................................
Hap_14    .............C....................................
Hap_15    ..................................................
Hap_16    ..................................................
Hap_17    ..................................................
Hap_18    .C................................................
Hap_19    ..................................................
Hap_20    ..................................................
Hap_21    ..................................................
Hap_22    ..................................................
Hap_23    ..................................................
Hap_24    .............................G....................
Hap_25    ..................................................
Hap_26    ......................................G..G........
Hap_27    ......................................G...........
Hap_28    ..................................................
Hap_29    ..................................................
Hap_30    ..................................................
Hap_31    ...........A......................................
```

```
Hap_32    ..................................................
Hap_33    ..................................................
Hap_34    ..................................................
Hap_35    ........................CG............G.........A.
Hap_36    .......................ACGATC.A.................A.
Hap_37    ...T....................C.........T.............A.
Hap_38    .................G................................
Hap_39    ......C.....ACAC.G.......................G........
Hap_40    .................G................................
Hap_41    ...............C..................................
Hap_42    ....................T.T.C.........T........T..T...
Hap_43    .......T........T........................A..TG.T..
Hap_44    ......................................GA.G........
Hap_45    .............................G........GA.....C....
Hap_46    ....C...C............GT......G......AG.GTA..G.TTA.
Hap_47    ....C.......A......A.G.......G......AG.GTA..G.TTA.
Hap_48    .....................G.......G......AG.GTA..G.TTA.
Hap_49    ....C.C.C...A..C.GTA.GTAC....GA..A..AG.G.A..G.....
Hap_50    ........C........GTA.G.......GA....CAG.GTAT.G.TTAC
Hap_51    ........C............G..............AG.G.A......A.
Hap_52    ........C...A........G..C....G......AG.G.A.TG...A.

[                        60        70        80        90       100]
[                         *         *         *         *         *]
Hap_1     CAAAATAGACTTTGGGGCAATCTGTAGGATGAGTCATTTCTTAAATGAAA
Hap_2     ......T..A....TCC...C.CTGG..C............C........
Hap_3     .C.T...A..GAGA.C..T.......AA.GAGA..C..............
Hap_4     .C.......A..G.T.CG...A...........A.CG...G.........
Hap_5     .........A....T..................A................
Hap_6     .........A....T.C....A...........A.C..............
Hap_7     .........A....T.CG...A...........A.C.....C........
Hap_8     .........A....T.CG...A...........A................
Hap_9     .........A....T.CG...............A................
Hap_10    .........A....T.CG...............A.C..............
Hap_11    .........A....T.CG...............A.C..............
Hap_12    .........A....T.CG...........G...A.C..............
Hap_13    .........A....T.CG...A...........A................
Hap_14    .........A....T.CG...A...........A................
Hap_15    .CT......A....T.CG...........G...A.C.....C........
Hap_16    .........A....T.CG...A...........A......C.........
Hap_17    .........A....T.CG...A...........A......C.........
Hap_18    ....G....A....T.CG...A...........A......C.........
Hap_19    .........A....T.CG...A...........A...........A..G.
Hap_20    .........A....T.CG...A...........A......G.........
Hap_21    .....C...A....T.CG.G.A...........A................
Hap_22    .........A....T.CG...A...........A.............G.G
Hap_23    .........A....T.CG...A...........A.....T..........
Hap_24    .........A....T.CG...A...........A................
Hap_25    ..............T.CG...A...........A........C.......
Hap_26    .........A....T.CG...A...........A..GGG...........
Hap_27    .........A....T.CG...A...........A................
Hap_28    .........A....T.CG...A...........A................
Hap_29    .........A....T.CG...A...........A................
Hap_30    ........TA....T.CG...A...........A.......G........
Hap_31    .........A....T.CG...A...........A................
Hap_32    .........A....T.CG...A...........AT...............
Hap_33    .........A....T.CG...A...........AT...........A...
Hap_34    .........A....T.CG...A...........A................
Hap_35    .........A....T.CG...A...........A................
Hap_36    .........A....T.CG...A...........A.........T......
Hap_37    .........A....T.CG...A...........A................
Hap_38    .........A..G.T.CG...A...........A................
Hap_39    .........A....T.CG...A...........A................
Hap_40    .........A...AT.CG...A...........A................
Hap_41    .........A....T.CG...A...........A......C.........
Hap_42    .........A....T.CG...A...........A................
```

```
Hap_43   .........A....T.CG...A..........A................
Hap_44   .........A....T.CG..............A................
Hap_45   .........A..G.T.CG...A..........A.C..............
Hap_46   .C.......A...AT.CG...A..........A.C..............
Hap_47   .........A....T.CG...A..........A................
Hap_48   .........A....T.CG..............A................
Hap_49   A........A....T.CG..............A................
Hap_50   .........A....T.CG...A..........A................
Hap_51   .........A....T.CG...A..........A................
Hap_52   .........A....T.CG...A..........A................

[                110        120  ]
[                  *          *  ]
Hap_1    GGTTTTAATTTCTCTTATGAATC
Hap_2    ..........A..........A.
Hap_3    .......................
Hap_4    .......................
Hap_5    .......................
Hap_6    .......................
Hap_7    .......................
Hap_8    .......................
Hap_9    .......................
Hap_10   C......................
Hap_11   .......................
Hap_12   .......................
Hap_13   .......................
Hap_14   .......................
Hap_15   .....................C.
Hap_16   ....C................C.
Hap_17   .......................
Hap_18   .......................
Hap_19   .......................
Hap_20   .......................
Hap_21   .......................
Hap_22   .......................
Hap_23   .......................
Hap_24   .......................
Hap_25   .......................
Hap_26   .......................
Hap_27   .......................
Hap_28   .A.....................
Hap_29   ...A...................
Hap_30   .......T...............
Hap_31   .......................
Hap_32   .......................
Hap_33   .......................
Hap_34   ......T.AG.AAGAACCCG..A
Hap_35   .......................
Hap_36   .......................
Hap_37   .......................
Hap_38   ..C..G.................
Hap_39   .....................C.
Hap_40   .......................
Hap_41   .......................
Hap_42   .......................
Hap_43   .......................
Hap_44   .......................
Hap_45   .......................
Hap_46   .......................
Hap_47   .......................
Hap_48   .......................
Hap_49   .......................
Hap_50   .......................
Hap_51   ....................TC.
Hap_52   .......................
;END;
```

```
BEGIN DnaSP;
    Genome= Diploid;
    ChromosomalLocation= Autosome;
    VariationType= Haplotypes;
    DnaSPversion= Ver. 6.12.03;
END;
BEGIN TRAITS;
Dimensions NTRAITS=8;
Format labels=yes missing=? separator=Comma;
TraitLabels Southeast CentralHighlands Northwest Thailand Brazil Colombia Mexico CIFC;
Matrix
Hap_1 0,1,0,0,0,0,0,0
Hap_2 1,0,0,0,0,0,0,0
Hap_3 0,1,0,0,0,0,0,0
Hap_4 0,1,0,0,0,0,0,0
Hap_5 0,0,1,0,0,0,0,0
Hap_6 1,0,0,0,0,0,0,0
Hap_7 0,0,0,1,0,0,0,0
Hap_8 1,1,3,0,5,2,3,0
Hap_9 0,1,1,0,0,0,0,0
Hap_10 0,0,1,0,0,0,0,0
Hap_11 0,0,1,0,0,0,0,0
Hap_12 0,0,1,0,0,0,0,0
Hap_13 0,0,1,0,0,0,0,0
Hap_14 0,0,1,0,0,0,0,0
Hap_15 0,1,0,0,0,0,0,0
Hap_16 1,0,0,0,0,0,0,0
Hap_17 0,0,0,0,2,0,4,3
Hap_18 0,0,0,0,1,0,0,0
Hap_19 0,0,0,0,1,0,0,0
Hap_20 0,1,0,0,0,0,0,0
Hap_21 0,0,0,0,1,0,0,0
Hap_22 0,0,0,0,1,0,0,0
Hap_23 0,0,0,0,1,0,0,0
Hap_24 0,0,0,0,0,0,2,0
Hap_25 0,0,1,0,0,0,0,0
Hap_26 0,0,0,0,0,1,0,0
Hap_27 1,0,0,0,0,1,0,0
Hap_28 0,0,0,0,1,0,0,0
Hap_29 0,0,0,0,1,0,0,0
Hap_30 0,0,0,0,1,0,0,0
Hap_31 0,0,0,0,1,0,0,0
Hap_32 0,0,0,0,1,6,0,0
Hap_33 0,0,0,0,1,0,0,0
Hap_34 1,0,0,0,0,0,0,0
Hap_35 1,0,0,0,0,0,0,0
Hap_36 1,0,0,0,0,0,0,0
Hap_37 0,1,0,0,0,0,0,0
Hap_38 0,1,0,0,0,0,0,0
Hap_39 1,0,0,0,0,0,0,0
Hap_40 1,0,0,0,0,0,0,0
Hap_41 1,0,0,0,0,0,0,0
Hap_42 0,1,0,0,0,0,0,0
Hap_43 0,1,0,0,0,0,0,0
Hap_44 0,1,0,0,0,0,0,0
Hap_45 0,0,0,1,0,0,0,0
Hap_46 0,1,0,0,0,0,0,0
Hap_47 0,1,0,0,0,0,0,0
Hap_48 0,1,0,0,0,0,0,0
Hap_49 0,1,0,0,0,0,0,0
Hap_50 0,1,0,0,0,0,0,0
Hap_51 0,1,0,0,0,0,0,0
Hap_52 0,1,0,0,0,0,0,0
;END;
```

Appendix 5: Input file for SplitsTree software

```
#NEXUS
BEGIN DATA;
DIMENSIONS NTAX= 52 NCHAR=123;
FORMAT DATATYPE=DNA    MISSING=? GAP=- MATCHCHAR=. INTERLEAVE ;
MATRIX
[                    10         20         30         40         50]
[                     *          *          *          *          *]
Hap_1     ATAAACAGTCCGGACACTATATCTGCGAATTAGTATGTTTATGGTTCATA
Hap_2     G.T......TG.....T.............GC.......G.........
Hap_3     ....C.......A.........................G..........
Hap_4     ............A...........C.........................
Hap_5     ..................................................
Hap_6     ..................................................
Hap_7     ....................T..C.........................
Hap_8     ..................................................
Hap_9     ..................................................
Hap_10    ..................................................
Hap_11    ..................................................
Hap_12    ..................................................
Hap_13    .....A............................................
Hap_14    ............C.....................................
Hap_15    ..................................................
Hap_16    ..................................................
Hap_17    ..................................................
Hap_18    .C................................................
Hap_19    ..................................................
Hap_20    ..................................................
Hap_21    ..................................................
Hap_22    ..................................................
Hap_23    ..................................................
Hap_24    ............................G.....................
Hap_25    ..................................................
Hap_26    ....................................G..G........
Hap_27    ....................................G..........
Hap_28    ..................................................
Hap_29    ..................................................
Hap_30    ..................................................
Hap_31    ...........A......................................
Hap_32    ..................................................
Hap_33    ..................................................
Hap_34    ..................................................
Hap_35    .......................CG............G.........A.
Hap_36    ......................ACGATC.A.................A.
Hap_37    ...T...................C.........T.............A.
Hap_38    .................G................................
Hap_39    ......C.....ACAC.G.......................G........
Hap_40    .................G................................
Hap_41    ...............C..................................
Hap_42    ....................T.T.C.........T........T..T...
Hap_43    .......T........T......................A...TG.T...
Hap_44    .....................................GA.G........
Hap_45    .............................G........GA.....C....
Hap_46    ....C...C............GT......G......AG.GTA..G.TTA.
Hap_47    ....C.......A......A.G.......G......AG.GTA..G.TTA.
Hap_48    ......................G.......G......AG.GTA..G.TTA.
Hap_49    ....C.C.C...A..C.GTA.GTAC....GA..A..AG.G.A..G.....
Hap_50    ........C........GTA.G.......GA....CAG.GTAT.G.TTAC
Hap_51    ........C............G..............AG.G.A......A.
Hap_52    ........C...A........G..C....G......AG.G.A.TG...A.

[                    60         70         80         90        100]
[                     *          *          *          *          *]
Hap_1     CAAAATAGACTTTGGGGCAATCTGTAGGATGAGTCATTTCTTAAATGAAA
```

```
Hap_2     ......T..A....TCC...C.CTGG..C............C........
Hap_3     .C.T...A..GAGA.C..T.......AA.GAGA..C..............
Hap_4     .C.......A..G.T.CG...A...........A.CG...G.........
Hap_5     .........A....T..................A................
Hap_6     .........A....T.C....A...........A.C.............
Hap_7     .........A....T.CG...A...........A.C.....C........
Hap_8     .........A....T.CG...A...........A................
Hap_9     .........A....T.CG..............A................
Hap_10    .........A....T.CG..............A.C.............
Hap_11    .........A....T.CG..............A.C.............
Hap_12    .........A....T.CG...........G...A.C.............
Hap_13    .........A....T.CG...A...........A................
Hap_14    .........A....T.CG...A...........A................
Hap_15    .CT......A....T.CG..........G....A.C.....C........
Hap_16    .........A....T.CG...A...........A.......C........
Hap_17    .........A....T.CG...A...........A.......C........
Hap_18    ....G....A....T.CG...A...........A.......C........
Hap_19    .........A....T.CG...A...........A...........A..G.
Hap_20    .........A....T.CG...A...........A.......G........
Hap_21    .....C...A....T.CG.G.A...........A................
Hap_22    .........A....T.CG...A...........A.............G.G
Hap_23    .........A....T.CG...A...........A.....T..........
Hap_24    .........A....T.CG...A...........A................
Hap_25    ..............T.CG...A...........A.........C......
Hap_26    .........A....T.CG...A...........A..GGG...........
Hap_27    .........A....T.CG...A...........A................
Hap_28    .........A....T.CG...A...........A................
Hap_29    .........A....T.CG...A...........A................
Hap_30    ........TA....T.CG...A...........A........G.......
Hap_31    .........A....T.CG...A...........A................
Hap_32    .........A....T.CG...A...........AT...............
Hap_33    .........A....T.CG...A...........AT...........A...
Hap_34    .........A....T.CG...A...........A................
Hap_35    .........A....T.CG...A...........A................
Hap_36    .........A....T.CG...A...........A..........T.....
Hap_37    .........A....T.CG...A...........A................
Hap_38    .........A..G.T.CG...A...........A................
Hap_39    .........A....T.CG...A...........A................
Hap_40    .........A...AT.CG...A...........A................
Hap_41    .........A....T.CG...A...........A.......C........
Hap_42    .........A....T.CG...A...........A................
Hap_43    .........A....T.CG...A...........A................
Hap_44    .........A....T.CG..............A................
Hap_45    .........A..G.T.CG...A...........A.C.............
Hap_46    .C.......A...AT.CG...A...........A.C.............
Hap_47    .........A....T.CG...A...........A................
Hap_48    .........A....T.CG..............A................
Hap_49    A........A....T.CG...............A................
Hap_50    .........A....T.CG...A...........A................
Hap_51    .........A....T.CG...A...........A................
Hap_52    .........A....T.CG...A...........A................

[                  110          120    ]
[                    *            *    ]
Hap_1     GGTTTTAATTTCTCTTATGAATC
Hap_2     ..........A.........A.
Hap_3     ......................
Hap_4     ......................
Hap_5     ......................
Hap_6     ......................
Hap_7     ......................
Hap_8     ......................
Hap_9     ......................
Hap_10    C.....................
Hap_11    ......................
Hap_12    ......................
```

```
Hap_13    .......................
Hap_14    .......................
Hap_15    .....................C.
Hap_16    ....C................C.
Hap_17    .......................
Hap_18    .......................
Hap_19    .......................
Hap_20    .......................
Hap_21    .......................
Hap_22    .......................
Hap_23    .......................
Hap_24    .......................
Hap_25    .......................
Hap_26    .......................
Hap_27    .......................
Hap_28    .A.....................
Hap_29    ...A...................
Hap_30    .......T...............
Hap_31    .......................
Hap_32    .......................
Hap_33    .......................
Hap_34    ......T.AG.AAGAACCCG..A
Hap_35    .......................
Hap_36    .......................
Hap_37    .......................
Hap_38    ..C..G.................
Hap_39    .....................C.
Hap_40    .......................
Hap_41    .......................
Hap_42    .......................
Hap_43    .......................
Hap_44    .......................
Hap_45    .......................
Hap_46    .......................
Hap_47    .......................
Hap_48    .......................
Hap_49    .......................
Hap_50    .......................
Hap_51    ....................TC.
Hap_52    .......................
;
END;
BEGIN TRAITS;
Dimensions NTRAITS=8;
Format labels=yes missing=? separator=Comma;
TraitLabels Southeast CentralHighlands Northwest Thailand Brazil Colombia Mexico CIFC;
Matrix
Hap_1 0,1,0,0,0,0,0,0
Hap_2 1,0,0,0,0,0,0,0
Hap_3 0,1,0,0,0,0,0,0
Hap_4 0,1,0,0,0,0,0,0
Hap_5 0,0,1,0,0,0,0,0
Hap_6 1,0,0,0,0,0,0,0
Hap_7 0,0,0,1,0,0,0,0
Hap_8 1,1,3,0,5,2,3,0
Hap_9 0,1,1,0,0,0,0,0
Hap_10 0,0,1,0,0,0,0,0
Hap_11 0,0,1,0,0,0,0,0
Hap_12 0,0,1,0,0,0,0,0
Hap_13 0,0,1,0,0,0,0,0
Hap_14 0,0,1,0,0,0,0,0
Hap_15 0,1,0,0,0,0,0,0
Hap_16 1,0,0,0,0,0,0,0
Hap_17 0,0,0,0,2,0,4,3
Hap_18 0,0,0,0,1,0,0,0
Hap_19 0,0,0,0,1,0,0,0
```

```
Hap_20 0,1,0,0,0,0,0,0
Hap_21 0,0,0,0,1,0,0,0
Hap_22 0,0,0,0,1,0,0,0
Hap_23 0,0,0,0,1,0,0,0
Hap_24 0,0,0,0,0,0,2,0
Hap_25 0,0,1,0,0,0,0,0
Hap_26 0,0,0,0,0,1,0,0
Hap_27 1,0,0,0,0,1,0,0
Hap_28 0,0,0,0,1,0,0,0
Hap_29 0,0,0,0,1,0,0,0
Hap_30 0,0,0,0,1,0,0,0
Hap_31 0,0,0,0,1,0,0,0
Hap_32 0,0,0,0,1,6,0,0
Hap_33 0,0,0,0,1,0,0,0
Hap_34 1,0,0,0,0,0,0,0
Hap_35 1,0,0,0,0,0,0,0
Hap_36 1,0,0,0,0,0,0,0
Hap_37 0,1,0,0,0,0,0,0
Hap_38 0,1,0,0,0,0,0,0
Hap_39 1,0,0,0,0,0,0,0
Hap_40 1,0,0,0,0,0,0,0
Hap_41 1,0,0,0,0,0,0,0
Hap_42 0,1,0,0,0,0,0,0
Hap_43 0,1,0,0,0,0,0,0
Hap_44 0,1,0,0,0,0,0,0
Hap_45 0,0,0,1,0,0,0,0
Hap_46 0,1,0,0,0,0,0,0
Hap_47 0,1,0,0,0,0,0,0
Hap_48 0,1,0,0,0,0,0,0
Hap_49 0,1,0,0,0,0,0,0
Hap_50 0,1,0,0,0,0,0,0
Hap_51 0,1,0,0,0,0,0,0
Hap_52 0,1,0,0,0,0,0,0
;
END;
```

Appendix 6: Input file for GENGIS software

Newick Tree file

```
((((((((((TSH-R30059:0.00615303,TSH-R30109:0.01059881)0.2150:0.00148073,(TSH-R30031:0.00450173,TSH-R30034:0.01893241)0.2450:0.00156462)0.0890:0.00148373,(TSH-R30050:0.00915058,TSH-R30098:0.00601757)0.4650:0.00365093)0.0080:0.00176082,(TSH-R30137:0.00000347,(TSH-R30043:0.00297248,(TSH-R30024:0.01210394,(TSH-R30063:0.00592900,TSH-R30042:0.00311465)0.4010:0.00000000)0.1710:0.00160205)0.0490:0.00000347)0.0360:0.00151358)0.0010:0.00176716,Mexico-6:0.00000347)0.0000:0.00177510,((TSH-R30073:0.00000347,TSH-R30120:0.00149603)0.1740:0.00000000,(((TSH-R30133:0.00000000,TSH-R30145:0.00439843)0.2250:0.00000347,TSH-R30155:0.00294384)0.1750:0.00000000,((TSH-R59309:0.00592906,TSH-R30036:0.00000347)0.1470:0.00144322,(TSH-R30148:0.00156235,(TSH-R30045:0.04426131,(TSH-R30055:0.00105594,TSH-R30071:0.03951139)0.4320:0.00341678)0.2100:0.00149482)0.2340:0.00282988)0.0040:0.00290419)0.0010:0.00145853)0.0010:0.00147879)0.0000:0.00000000,((((((((Mexico-9:0.00000000,Brazil-12:0.00293079)0.0830:0.00000347,CIFC-1:0.00000000)0.0080:0.00000347,Mexico-3:0.00000000)0.0140:0.00000347,Mexico-2:0.00000000)0.0200:0.00000347,Mexico-1:0.00000000)0.1550:0.00000000,(TSH-R30052:0.00738290,TSH-R30028:0.00146449)0.2550:0.00149405)0.0330:0.00000347,TSH-R30069:0.00301454)0.0120:0.00147591,(TSH-R30129:0.00151099,(TSH-R30167:0.00150287,(TSH-R30125:0.00301272,((Mexico-8:0.00000000,Brazil-18:0.00302052)0.6570:0.00148610,(Mexico-8:0.00000000,(Mexico-4:0.00000000,(TSH-R30172:0.00000000,(TSH-R30160:0.00000000,(Brazil-17:0.00148582,(Brazil-1:0.00148478,(Brazil-7:0.00297760,Brazil-11:0.00298576)0.0690:0.00000347)0.0290:0.00000347)0.0290:0.00000347)0.0010:0.00000347)0.0010:0.00000000)0.0020:0.00000347)0.0120:0.00000000)0.0070:0.00000347)0.0050:0.00000347)0.0090:0.00000347)0.0520:0.00000347)0.0010:0.00000347)0.0000:0.00000347,(Colombia-3:0.00458628,(CIFC-2:0.00150361,(Mexico-5:0.00000000,Mexico-7:0.00000000)0.8140:0.00151121)0.3920:0.00000347)0.3080:0.00149626)0.0000:0.00000347,(Colombia-5:0.00304452,(Colombia-6:0.00000347,(((Colombia-8:0.00000000,Brazil-9:0.00299104)0.2200:0.00000347,Colombia-7:0.00150516)0.3700:0.00000347,(Colombi-10:0.00000000,(Colombia-9:0.00000347,Brazil-3:0.00149841)0.4000:0.00105291)0.1510:0.00044325)0.3270:0.00149462)0.0610:0.00000347)0.3550:0.00149033,((((Brazil-15:0.00146517,Brazil-16:0.00000000)0.6600:0.00000347,Brazil-14:0.00298239)0.3690:0.00295135,CIFC-3:0.00000347)0.0410:0.00299116,((((((((TSH-R30046:0.02466789,Brazil-6:0.00000000)0.0670:0.00000347,Brazil-5:0.00000000)0.0070:0.00000347,Colombia-4:0.00000000)0.0040:0.00149557,TSH-R30022:0.00000000)0.0010:0.00000347,TSH-R30107:0.00000000)0.0130:0.00000347,Brazil-4:0.00300865)0.0040:0.00000000,(Brazil-2:0.00149115,(Brazil-10:0.00452523,Brazil-13:0.00149143)0.1480:0.00000347)0.0990:0.00149485)0.0520:0.00000347,(TSH-R30018:0.00000000,((Colombia-1:0.00753500,Colombia-2:0.00000347)0.4780:0.00297521,(TSH-R30104:0.00147137,(TSH-R59420:0.00666614,((TSH-R30085:0.00482631,TSH-R30101:0.01009049)0.4030:0.00373512,(TSH-R30080:0.00166263,(TSH-R30088:0.00397153,(TSH-R30051:0.00802187,(TSH-R30067:0.02688610,TSH-R30118:0.00705795)0.6480:0.00964668)0.1150:0.00161317)0.1020:0.00196726)0.1470:0.00337021)0.1920:0.00196291)0.1330:0.00237569)0.1980:0.00471557)0.0180:0.00000000)0.0450:0.00149875)0.0000:0.00000000)0.0000:0.00149336);
```

Location file:

Site Id	Latitude	Longitude	Group
TSH-R30120	21.239	103.86	NW_SonLa
TSH-R30125	21.259	103.869	NW_SonLa
TSH-R30129	21.354	103.812	NW_SonLa
TSH-R30133	21.518	103.644	NW_SonLa
TSH-R30137	21.524	103.638	NW_SonLa
TSH-R30145	21.506	103.653	NW_SonLa
TSH-R30148	21.539	103.62	NW_SonLa
TSH-R30155	21.565	103.465	NW_DienBien
TSH-R30160	21.515	103.206	NW_DienBien
TSH-R30167	21.504	103.197	NW_DienBien
TSH-R30172	21.503	103.202	NW_DienBien
TSH-R30050	12.553	108.181	CH_DakLak
TSH-R30051	12.551	108.191	CH_DakLak

TSH-R30052	12.621	108.192	CH_DakLak
TSH-R30055	12.667	108.196	CH_DakLak
TSH-R30059	12.704	108.196	CH_DakLak
TSH-R30063	12.568	107.824	CH_DakNong
TSH-R30067	12.495	107.74	CH_DakNong
TSH-R30069	12.472	107.692	CH_DakNong
TSH-R30071	12.182	107.638	CH_DakNong
TSH-R30073	12.156	107.64	CH_DakNong
TSH-R30080	12.062	107.677	CH_DakNong
TSH-R30085	11.984	107.7	CH_DakNong
TSH-R30088	11.737	107.979	CH_Lam Dong
TSH-R30098	11.776	108.154	CH_Lam Dong
TSH-R30101	11.785	108.184	CH_Lam Dong
TSH-R30104	11.763	108.345	CH_Lam Dong
TSH-R30107	11.624	108.256	CH_Lam Dong
TSH-R30109	11.628	108.141	CH_Lam Dong
TSH-R30118	11.559	107.831	CH_Lam Dong
TSH-R30018	10.984	107.254	SE_DongNai
TSH-R30022	10.947	107.313	SE_DongNai
TSH-R30024	10.858	107.163	SE_DongNai
TSH-R30028	10.973	107.077	SE_DongNai
TSH-R30031	11.242	107.391	SE_DongNai
TSH-R30034	11.267	107.379	SE_DongNai
TSH-R30036	10.77	107.264	SE_DongNai
TSH-R30042	11.706	107.053	SE_BinhPhuoc
TSH-R30043	11.742	107.146	SE_BinhPhuoc
TSH-R30045	11.801	107.217	SE_BinhPhuoc
TSH-R30046	11.796	107.217	SE_BinhPhuoc
TSH-R59309	18.32	98.32	Thailand
TSH-R59420	18.3	98.28	Thailand
Mexico-1	19.43	-96.9	Veracruz
Mexico-2	19.17	-96.96	Veracruz
Mexico-3	16.7	-92.83	Chiapas
Mexico-4	15	-92.15	Chiapas
Mexico-5	15.05	-92.07	Chiapas
Mexico-6	15.94	-96.46	Oaxaca
Mexico-7	15.95	-96.45	Oaxaca

Mexico-8	20.6	-97.69	Puebla
Mexico-9	20.37	-97.86	Puebla
Colombia-1	4.35857	-75.7405	Quindio_Buenavista
Colombia-2	4.35857	-75.7405	Quindio_Buenavista
Colombia-3	4.35857	-75.7405	Quindio_Buenavista
Colombia-4	4.98294	-75.6102	Caldas_Chinchina
Colombia-5	4.35857	-75.7405	Quindio_Buenavista
Colombia-6	4.35857	-75.7405	Quindio_Buenavista
Colombia-7	4.98294	-75.6102	Caldas_Chinchina
Colombia-8	4.98294	-75.6102	Caldas_Chinchina
Colombia-9	4.35857	-75.7405	Quindio_Buenavista
Colombia-10	4.98294	-75.6102	Caldas_Chinchina
CIFC-1	38.70152	-9.31236	CIFC-Portugal
CIFC-2	38.70152	-9.31236	CIFC-Portugal
CIFC-3	38.70152	-9.31236	CIFC-Portugal
Brazil-1	-20.8468	-42.8445	MG_Coimbra
Brazil-2	-20.8468	-42.8445	MG_Coimbra
Brazil-3	-20.8468	-42.8445	MG_Coimbra
Brazil-4	-20.8468	-42.8445	MG_Coimbra
Brazil-5	-20.8468	-42.8445	MG_Coimbra
Brazil-6	-20.8468	-42.8445	MG_Coimbra
Brazil-7	-20.3329	-41.1296	ES_Venda_Nova_do_Imigrante
Brazil-8	-20.3329	-41.1296	ES_Venda_Nova_do_Imigrante
Brazil-9	-20.3329	-41.1296	ES_Venda_Nova_do_Imigrante
Brazil-10	-20.3329	-41.1296	ES_Venda_Nova_do_Imigrante
Brazil-11	-18.6831	-49.5708	MG_Capinopolis
Brazil-12	-18.6831	-49.5708	MG_Capinopolis
Brazil-13	-18.6831	-49.5708	MG_Capinopolis
Brazil-14	-20.9178	-46.9997	MG_Sao_Sebastiao_do_Paraiso
Brazil-15	-20.9178	-46.9997	MG_Sao_Sebastiao_do_Paraiso
Brazil-16	-20.9178	-46.9997	MG_Sao_Sebastiao_do_Paraiso
Brazil-17	-20.7939	-43.3431	MG_Senhora_de_Oliveira
Brazil-18	-20.7939	-43.3431	MG_Senhora_de_Oliveira