

# Unravelling the determinants of human health in French Polynesia: the MATAEA project

**Running title:** Health determinants in French Polynesia

Iotefa Teiti<sup>1†</sup>, Maite Aubry<sup>1†</sup>, Sandrine Fernandes-Pellerin<sup>2</sup>, Etienne Patin<sup>3</sup>, Yoann Madec<sup>4</sup>, Pauline Boucheron<sup>4</sup>, Jessica Vanhomwegen<sup>5</sup>, Jérémie Torterat<sup>6</sup>, Stéphane Lastère<sup>7</sup>, Sophie Olivier<sup>8</sup>, Anthony Jaquaniello<sup>3,9</sup>, Maguelonne Roux<sup>3,10</sup>, Vincent Mendiboure<sup>4</sup>, Christine Harmant<sup>3</sup>, Aurélie Bisiaux<sup>3</sup>, Gaston Rijo de León<sup>3</sup>, Dang Liu<sup>3</sup>, Hervé Bossin<sup>11</sup>, Françoise Mathieu-Daudé<sup>11,12</sup>, Clémence Gatti<sup>13</sup>, Edouard Suhas<sup>14</sup>, Kiyojiken Chung<sup>1</sup>, Bertrand Condat<sup>15</sup>, Pierre Ayotte<sup>16</sup>, Nicolas Prud'homme<sup>6</sup>, Eric Conte<sup>17</sup>, Nathalie Jolly<sup>2</sup>, Jean-Claude Manugerra<sup>5</sup>, Anavaj Sakuntabhai<sup>18</sup>, Arnaud Fontanet<sup>4,19§</sup>, Lluís Quintana-Murci<sup>3,20§</sup>, Van-Mai Cao-Lormeau<sup>1§\*</sup>

<sup>1</sup> Laboratory of research on emerging viral diseases, Institut Louis Malardé, Papeete, Tahiti, French Polynesia

<sup>2</sup> Center for Translational Sciences, Institut Pasteur, Paris, France

<sup>3</sup> Institut Pasteur, Université Paris Cité, CNRS UMR2000, Human Evolutionary Genetics Unit, F-75015 Paris, France

<sup>4</sup> Institut Pasteur, Université Paris Cité, Epidemiology of emerging diseases unit, 75015 Paris, France

<sup>5</sup> Environment and Infectious Risk Unit, Laboratory for Urgent Response to Biological Threats, Institut Pasteur, Paris, France

<sup>6</sup> Institut de la Statistique de la Polynésie française, Papeete, Tahiti, French Polynesia

<sup>7</sup> Clinical Laboratory, Centre hospitalier de Polynésie française, Pirae, Tahiti, French Polynesia

<sup>8</sup> Clinical Laboratory, Institut Louis Malardé, Papeete, Tahiti, French Polynesia

<sup>9</sup> Institut Pasteur, Data Management Core Facility, Paris, France

<sup>10</sup> Institut Pasteur, Université de Paris, Bioinformatics and Biostatistics Hub, F-75015 Paris, France

<sup>11</sup> Laboratory of research in medical entomology, Institut Louis Malardé, Papeete, Tahiti, French Polynesia

<sup>12</sup> MIVEGEC, University of Montpellier, CNRS, IRD, Montpellier, France

<sup>13</sup> Laboratory of Marine Biotoxins, UMR241-Ecosystèmes Insulaires Océaniques (EIO) (IFREMER, ILM, IRD, UPF), Institut Louis Malardé, Papeete, Tahiti, French Polynesia

<sup>14</sup> Unit on non-communicable diseases, UMR241-Ecosystèmes Insulaires Océaniques (EIO) (IFREMER, ILM, IRD, UPF), Institut Louis Malardé, Papeete, Tahiti, French Polynesia

<sup>15</sup> Department of Gastroenterology, Centre hospitalier de Polynésie française, Pirae, Tahiti, French Polynesia

<sup>16</sup> Centre de Toxicologie du Québec, Institut National de Santé Publique du Québec, Québec, Canada

<sup>17</sup> Maison des Sciences de l'Homme du Pacifique, Université de la Polynésie française, Punaauia, Tahiti, French Polynesia

<sup>18</sup> Functional Genetics of Infectious Diseases Unit, Department of Global Health, Institut Pasteur, Paris Cedex 15, France

<sup>19</sup> PACRI unit, Conservatoire National des Arts et Métiers, Paris, France

<sup>20</sup> Chair Human Genomics and Evolution, Collège de France, Paris 75005, France

## \* Correspondence:

Van-Mai Cao-Lormeau

[mlormeau@ilm.pf](mailto:mlormeau@ilm.pf)

† These authors have contributed equally to this work and share first authorship

§ These authors have contributed equally to this work and share last authorship

**Keywords:** French Polynesia, cross-sectional survey, non-communicable diseases, communicable diseases, population genetics.

## Abstract

**Background:** French Polynesia is a French overseas collectivity in the Southeast Pacific, comprising 75 inhabited islands across five archipelagoes. The human settlement of the region corresponds to the last massive migration of humans to empty territories, but its timeline is still debated. Despite their recent population history and geographical isolation, inhabitants of French Polynesia experience health issues similar to those of continental countries. Modern lifestyles and increased longevity have led to a rise in non-communicable diseases (NCDs) such as obesity, diabetes, hypertension, and cardiovascular diseases. Likewise, international trade and people mobility have caused the emergence of communicable diseases (CDs) including mosquito-borne and respiratory diseases. Additionally, chronic pathologies including acute rheumatic fever, liver diseases, and ciguatera, are highly prevalent in French Polynesia. However, data on such diseases are scarce and not representative of the geographic fragmentation of the population.

**Objectives:** The MATAEA project aims to estimate the prevalence of several NCDs and CDs in the population of the five archipelagoes, and identify associated risk factors. Moreover, genetic analyses will contribute to determinate the sequence and timings of the peopling history of French Polynesia, and identify causal links between past genetic adaptation to island environments, and present-day susceptibility to certain diseases.

**Methods:** This cross-sectional survey is based on the random selection of 2,100 adults aged 18-69 years and residing on 18 islands from the five archipelagoes. Each participant answered a questionnaire on a wide range of topics (including demographic characteristics, lifestyle habits and medical history), underwent physical measurements (height, weight, waist circumference, arterial pressure, and skin pigmentation), and provided biological samples (blood, saliva, and stool) for biological, genetic and microbiological analyses.

**Conclusion:** For the first time in French Polynesia, the MATAEA project allows to collect a wide range of data to explore the existence of indicators and/or risk factors for multiple pathologies of public health concern. The results will help health authorities to adapt actions and preventive measures aimed at reducing the incidence of NCDs and CDs. Moreover, the new genomic data generated in this study, combined with anthropological data, will increase our understanding of the peopling history of French Polynesia.

## 79 INTRODUCTION

80 French Polynesia is located in the Southeast Pacific and within the Polynesian triangle, which is  
81 bounded by the islands of Hawaii to the North, Aotearoa (New Zealand) to the West and Rapa Nui  
82 (Easter Island) to the East. This French overseas collectivity extends over a maritime area as large as  
83 Europe and includes 121 islands, of which 75 are inhabited, distributed within five archipelagoes:  
84 Society, Marquesas, Tuamotu, Gambier and Austral (Figure 1). According to the previous national  
85 population census conducted in 2017, there are about 276,000 inhabitants distributed in the various  
86 archipelagoes, of which nearly 75% are concentrated in the islands of Tahiti and Moorea (Windward  
87 Islands, Society archipelago) (1).

88 In the course of the history of human migrations around the world, Pacific islands, and  
89 Polynesia in particular, were the last empty lands to be settled (2). Although archaeological and  
90 linguistic data have provided evidence for the geographical and temporal origin of the populations of  
91 the Eastern Pacific islands, the sequence and timings of settlement of the Polynesian islands are still  
92 debated (3). To help elucidate these gaps, a few genetic studies including samples from Pacific  
93 populations have been conducted over the past five years (2-6). However, for the studies involving the  
94 use of samples from the population of French Polynesia, only a limited number of samples collected  
95 in a few islands were available. Some studies also aimed to identify genomic variations related to  
96 adaptation of the populations to island environments. In a study including genomes from 20  
97 populations from the Pacific region, candidate variants for positive selection were found in genes  
98 relating to immunity and metabolism, suggesting genetic adaptation to pathogens and food sources  
99 specific to Pacific islands (4). However, further investigations are needed to identify causal links  
100 between past genetic adaptation and present-day disease risk in Pacific settings. Indeed, most studies  
101 investigating genetic links with vulnerabilities to non-communicable (NCDs) and communicable  
102 (CDs) diseases have focused on populations in Africa, Europe, Asia, and USA (7-10). The same  
103 observation is made for the research investigating the link between human gut microbiome composition  
104 and susceptibility to diseases (7). Despite the high burden of various NCDs and CDs, Pacific islands  
105 remain highly underrepresented in genomic studies.

106 In addition to genetic and microbial factors, social and environmental factors such as poverty,  
107 education, and access to healthcare are thought to be major drivers of disease burden in the Pacific  
108 islands (7). In French Polynesia, prevention measures against NCDs and CDs have been enhanced, and  
109 access to healthcare has been improved in all archipelagoes over the past 3 decades (11). This has  
110 certainly contributed to the extension of life expectancy at birth, which has increased from 67.1 years  
111 to 74.7 for men and 72.8 years to 79.1 for women between 1990 and 2020 (12). However, modern  
112 lifestyles (including overconsumption of food and physical inactivity) and increased longevity have  
113 led to a rise in the incidence of NCDs (13). The most recent data collected in French Polynesia in 2010  
114 indicated that the prevalence of obesity and hypertension in adults was 40.4% and 26.7%, respectively.  
115 It was estimated that only 1.7% of the population of French Polynesia presented a low risk of  
116 developing a NCD whereas 45% were at high risk (14). According to the latest reports from the health  
117 authorities of French Polynesia, the leading causes of death are cardiovascular diseases (26.1%) mainly  
118 associated with lifestyle, and cancers (25.1%) (15, 16). Among the Pacific island countries and  
119 territories (PICTs), French Polynesia displayed one of the highest age-standardized cancer incidence  
120 (267.9 for males and 214.0 for females per 100,000) and age-standardized cancer mortality (157.2 for  
121 males and 106.8 for females per 100,000) in 2020 (17-19). The most frequently reported cancers are  
122 lung, prostate, breast, thyroid and colorectal cancers (19). Moreover, the incidence of hepatocellular  
123 carcinoma recorded in the Austral archipelago between 2013 and 2017 is among the highest in the  
124 world (43.1/100,000 inhabitants), this pathology being associated with a high prevalence of infection  
125 with hepatitis B virus (HBV) and obesity (20).

Dietary habits, especially high fish consumption, also expose the population of French Polynesia to seafood-related NCDs. Among them, ciguatera poisoning, a disease caused by the consumption of reef fishes or marine invertebrates contaminated by neurotoxins (ciguatoxins) produced by the microalgae *Gambierdiscus*, is highly prevalent in French Polynesia where it constitutes an important public health issue (21). In 2021, the overall incidence rate of ciguatera, based on the number of symptomatic cases reported by healthcare workers or by the patients themselves, was 6.9/10,000 inhabitants (22). However, as (i) ciguatera is not on the list of notifiable diseases, (ii) a large proportion of ciguatera-affected people do not consult a healthcare worker, and (iii) cases requiring emergency care or hospitalization in the main hospital of French Polynesia (*Centre hospitalier de Polynésie française*, Tahiti) are not systematically recorded, the true incidence rate of this disease is underestimated. Another source of seafood-related contaminants are heavy metals including mercury and lead. Indeed, high levels of mercury were detected in adolescents and adults from two atolls of the Tuamotu archipelago (Hao and Makemo) (23), in adults from the Leeward islands (Tahiti and Moorea) (24), and in cord blood samples of delivering women from all archipelagoes, suggesting that Polynesian newborns may be prenatally exposed to high doses of mercury (25). Moreover, concentrations of lead exceeding toxicological reference values were found in a significant proportion of the residents of Hao and Makemo islands tested to assess their exposure to toxic metals and polychlorinated biphenyls (PCBs) (23). Although levels of PCBs found in both atolls were low compared to international standards, there was a higher body burden of PCBs in residents of Hao which was suggested to be potentially linked to some contamination of the environment by military support activities in this island. Another source of environmental contamination that may have an impact on the population health are pesticides. To our knowledge, there are no data concerning the level of exposure to pesticides of the population of French Polynesia. However, surveys on the pollution of French Polynesia coral reefs by pesticides found the presence of herbicides and insecticides in fishes which are commonly consumed by the local inhabitants (26-28).

In addition to NCDs, CDs represent major health problems in French Polynesia. As a result of the prevention, screening, treatment and vaccination programs implemented by the health authorities, the incidence of several CDs has decreased (29). Despite these efforts, some CDs caused by bacterial infections persist, including leprosy (mean incidence of 1.6 cases per 100,000 inhabitants reported from 2000 to 2017) and tuberculosis (incidence of 19.4 cases per 100,000 inhabitants in 2018) (30, 31). Moreover, in 2013, the prevalence of Bancroftian lymphatic filariasis, an infection caused by the nematode worm *Wuchereria bancrofti* transmitted by mosquitoes, still had not reached the elimination threshold (*i.e.* filarial antigenemia <1%) defined by the World Health Organization (WHO) in Tahiti, Huahine (Leeward Islands, Society archipelago) and three islands in the Marquesas archipelago (Tahuata, Hiva Oa and Fatu Hiva), despite several years of community treatment (32-35). Furthermore, HBV remains a public health problem, despite the availability of an effective vaccine. Following the finding of hepatitis B surface antigens (HBsAg) among 10.48% inhabitants of the Austral archipelago randomly selected in 1987 (36), vaccination of children at birth started to be implemented in 1992 to decrease the burden of chronic HBV infection in French Polynesia (37), and became mandatory in 1995 (38). A nationwide serosurvey conducted in 2014 in 6 years old schoolchildren from the five archipelagoes showed a three-dose vaccination coverage of 98% and a prevalence of HBsAg of 0% (39). However, screening in 2019 of all adult residents of Rapa, the most isolated island of the Austral archipelago, revealed a prevalence of HBsAg of 11% in people born before 1992 (20).

Among the CDs considered as important public health issues in French Polynesia, those caused by arthropod-borne viruses (arboviruses) represent the highest risk for the population. Indeed, the four serotypes of dengue virus (DENV-1 to -4) have caused multiple outbreaks over the last decades, with deaths frequently reported in children (40, 41). Moreover, Zika virus (ZIKV) and chikungunya virus (CHIKV) caused explosive outbreaks in 2013-2014 (42) and 2014-2015 (43), respectively. During the ZIKV outbreak, severe neurologic complications in adults (44) and central nervous system

malformations in newborns and fetuses (45) were recorded. In addition to arboviral diseases, CDs caused by respiratory viruses have become major health concerns over the past decades. Several epidemics of variable severity caused by influenza A and/or B viruses have been reported in French Polynesia (29). While influenza was not considered a public health priority until the early 2000s, an epidemic caused by the influenza A virus subtype H1N1 highly affected the population of French Polynesia in 2009, with 13 severe cases requiring hospitalization in a critical care service and 7 deaths recorded (46). More recently, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) reached French Polynesia in March 2020 and caused several epidemic waves, with a total of 78,864 confirmed cases, 3,315 hospitalizations and 783 deaths reported as of August 31, 2022. A serosurvey conducted at the end of 2021, from a representative sample of the adult population of Tahiti, found that 57.7% of the study participants had been infected by SARS-CoV-2 during the first two epidemic waves (47).

Bacterial infections also represent a significant risk for the population of French Polynesia. The incidence of leptospirosis, a rodent-borne disease caused by bacteria belonging to the order *Spirochaetales*, exploded in 2017 with a rate of 72/100,000 inhabitants (48). Moreover, the incidence of *Staphylococcus aureus* (*S. aureus*) infections was estimated at 38/100,000 inhabitants in 2008, and this bacterium was shown to be responsible for one third of the community bacterial infections documented in French Polynesia. The proportion of *S. aureus* strains resistant to methicillin increased from less than 10% in 1995 to more than 35% in 2001 (49). French Polynesia is also characterized by the unusual high prevalence of acute rheumatic fever associated with infection by Group A streptococcus, with an incidence of 0.06% and a prevalence of 1.29% reported in 2015 (29, 50).

Although data exist on the levels of exposure of the population and the risk factors associated with NCDs and CDs in French Polynesia, these data are often old, fragmentary and concern a few islands. Moreover, genomic data, based on whole-genomes, from the Polynesian populations are lacking, precluding a comprehensive assessment of the settlement history of the various islands and archipelagoes, and of the history of adaptation of these populations to island environments - an information that should be of great benefit for the study of the genetic basis of certain diseases that present high incidence in the region. Here we describe the MATAEA project, a wide cross-sectional population-based survey designed to investigate the determinants of health in the adult population of French Polynesia, using a combination of data obtained from a questionnaire, physical measurements, and biological analyses in 2100 inhabitants of 18 different islands distributed among the five archipelagoes.

## STUDY OBJECTIVES

MATAEA is a Tahitian compound word (“MATA” means “eye” and “EA” means “health”) which can be translated by “take a look at health”. The first objective of the MATAEA project was to collect a wide range of data from a large number of inhabitants to assess the health status of the adult population of the five French Polynesian archipelagoes, and identify risk factors for NCDs and CDs related to the specific living context (including the lifestyle, place of residence, history of infections) and intrinsic characteristics of individuals (such as genetics, age, sex, oral and gut microbiome). Results will be made available to health authorities for them to adapt actions and preventive measures to better reduce diseases incidence. The second objective of the project was to use genomic data obtained from the residents to increase our understanding of the history of the settlement of the French Polynesian islands, and to identify the major demographic and adaptive events that may explain the current incidence of some NCDs, and the higher susceptibility to certain NCDs and CDs. Although designed as a general-purpose survey with a broad scope, topics identified as major public health issues will be given priority treatment by specialized laboratories involved in the project, as described below.

First, prevalence data for metabolic diseases (including obesity and diabetes) dating from 2010 will be updated (14), and their associated risk factors (such as dietary habits, physical activity, blood

pressure, high blood sugar, excess body fat around the waist, abnormal cholesterol or triglyceride level, genetic susceptibility factors, microbiome composition) will be investigated.

Second, the seroprevalence of mosquito-borne diseases known to have previously circulated in French Polynesia (including the four serotypes of dengue, Zika and chikungunya) will be measured as previously published (51, 52), to assess the risk for new outbreaks to occur. In addition, the prevalence of filarial antigenemia will be evaluated to check whether the elimination threshold recommended by the WHO has been reached in all archipelagoes, subsequent to the strengthening of community treatment in islands most affected by Bancroftian lymphatic filariasis (35). Risk factors for those mosquito-borne diseases will also be sought. Among them, the level of exposure of the population to mosquito bites will be measured by analyzing the presence of antibodies targeting mosquito salivary proteins in the blood of participants, as previously described (53).

Third, screening for HBsAg and anti-HBV antibodies will be performed to assess the prevalence of current and past HBV infections among the inhabitants of the five archipelagoes, especially those with increased risk factors such as being borne before 1995 and living in the Austral archipelago. As there is no available data regarding the prevalence of hepatitis C virus (HCV) infections in French Polynesia, the presence of anti-HCV antibodies will also be tested.

Fourth, concentrations of several heavy metals (including mercury and lead) and pesticides will be measured in the blood of inhabitants of the five archipelagoes to assess the degree of environmental contamination and the risk for human health. Data collected in the questionnaire, such as the island of residence and dietary habits, will be used to identify risk factors associated with contamination with toxic metals and pesticides, as well as ciguatera poisoning. As there is no confirmatory test available to detect current or past ciguatera poisoning, estimate of the prevalence of this disease in the population of the five archipelagoes will be based on the information self-reported by the participants in the questionnaire.

Fifth, as SARS-CoV-2 was introduced into French Polynesia during the participants' inclusion period and a large number of cases with mild or no symptoms might have not been reported, the presence of specific antibodies will be analyzed in the blood samples of participants to evaluate the real proportion of the population infected in the different archipelagoes after the emergence of the virus.

Sixth, serological analyzes will be conducted for a panel of viruses known to regularly cause infections in French Polynesia (such as influenza viruses, human immunodeficiency virus, rotaviruses and enteroviruses), but for which there is little data on the immune status of the population.

Finally, the seroprevalence of rodent-borne pathogens including Hantaviruses will be determined, as the high incidence of leptospirosis suggests that other diseases transmitted by rodents may circulate in French Polynesia. The only serosurvey of hantavirus infections in French Polynesia was conducted in 1989 and found specific antibodies in a few samples from rats and humans (54). Consequently, data collected in the present study will allow to detect possible silent circulation of Hantaviruses.

## **MATERIALS AND METHODS**

### **Sampling Design**

The MATAEA project is a cross-sectional population-based survey conducted on a random sample of the general population aged from 18 to 69 years, representative of the five archipelagoes of French Polynesia. The sampling protocol is based on the previous population census dating from 2017 (1). A sample with a target size of 2100 participants, equally divided into three age groups (18-29, 30-44 and 45-69), with a sex ratio of 1:1, was established. To represent all archipelagoes, French Polynesia was stratified into three geographical strata: Windward Islands (Society archipelago), Leeward Islands (Society archipelago), and other archipelagoes (Tuamotu, Gambier, Marquesas and Austral), which

respectively include 75%, 13% and 12% of the population (Table 1). Since the population size in the Windward Islands is much higher than in the Leeward Islands and in the other archipelagoes, the sample size was amplified in the latter two strata to keep sufficient precision around prevalence estimates for the health conditions under consideration. Consequently, the same sample size of 700 subjects was chosen for the three strata, allowing a precision never exceeding +/- 3.7% around prevalence estimates within each strata. In each strata, for convenient reasons, the most populated islands with a health care center and an airport with regular flight rotations were selected for the study, and the number of participants to be recruited was proportional to the population size of the islands as detailed in Table 1. In each island, households were randomly selected by the *Institut de la Statistique de la Polynésie française* (ISPF, Tahiti), and in each household one resident meeting the criteria of age and sex sought for the study was recruited.

### Study timeline

The recruitment of participants in the 18 islands selected for the study was initially planned from November 2019 to June 2020. However, following the introduction of SARS-CoV-2 in French Polynesia in March 2020, several interruptions occurred during the recruitment phases and the sampling period lengthened (Figure 2). The first break was caused by the population lockdown implemented from March 20 to May 21, 2020. Inclusions resumed in June 2020 but stopped the following month due to the occurrence of a first epidemic wave caused by the original Wuhan strain. Inclusions resumed once again in April 2021 but stopped during August and September 2021, because of a second epidemic wave caused by the Delta variant. The last recruitments were performed from October to December 2021.

### Inclusion/exclusion criteria

The study enrolled men and women aged from 18 to 69 years at the time of enrollment, who had been residing in French Polynesia for at least six months, were affiliated to the French Polynesia social security or assimilated regimens, and were able to provide a written informed consent. Were excluded from the study the minors, pregnant or parturient women, persons deprived of their liberty by a judicial or administrative decision, adults being subject to a legal protection measure or unable to express their consent, people undergoing psychiatric care, people not affiliated to the French Polynesia social security or assimilated regimens, disabled people, and the homeless.

### Participant recruitment and informed consent

Each household randomly selected for the study was visited by trained staff from the *Institut Louis Malardé* (ILM, Tahiti), and only one resident was recruited in the household. Detailed information on the purpose and procedures of the study were provided, to the individual selected, by the ILM staff. Then, signed written informed consent was obtained from the participant. Participation in the study involved agreeing to answer a questionnaire, to undergo physical measurements (height, weight, waist circumference, arterial pressure, and skin pigmentation), and to provide biological samples (blood, saliva, and optionally stool), as detailed in the next sections. The participant answered the questionnaire during face-to-face interview, and answers were reported on an electronic tablet device. Then, the participant was invited to provide a stool sample and was given a collection tube in case of acceptance. Finally, an appointment with a nurse was scheduled one to three days later, in the morning, to collect physical measurements and biological samples. The participant was asked to fast for at least 8 hours and to collect and store the stool sample in a fridge within 24 hours before the visit of the nurse.

Participants received no financial compensation. However, they freely benefited from the results of the anthropometric, blood pressure and fasting blood glucose measurements collected by the nurse, and of the hematology and biochemistry analyses performed from their blood samples at the

clinical laboratory of ILM. Moreover, and importantly, participants were informed if any atypical result was detected and were subsequently guided through a visit to a physician.

### **Questionnaire**

Case report forms (CRF) were built with ODK Collect (Get ODK Inc., USA) which is an open-source mobile client. The CRF consisted of 216 items divided into three major sections: socio-demographic status, lifestyle behaviors, and general health and medical history. It contained questions from the standardized STEPS questionnaire of the WHO designed to monitor the main NCDs risk factors (55), and additional questions specific to the context of French Polynesia. Questions related to COVID-19 were added in the questionnaire in April 2021. The detailed CRF is available in Supplemental Data 1.

### **Physical measurements**

#### ***Anthropometric measurements***

Anthropometric measurements included height, weight and waist circumference. Standing height was measured in centimeters (cm) with a portable stadiometer (SECA, model 213, France). Weight was measured in kilograms (kg) using a portable electronic weighting scale (SECA, model 813, France). Waist circumference was measured in cm using a constant tension tape (SECA, model 203 measuring tape, France) at the midpoint between the lowest rib and above the iliac crest. Body mass index (BMI) was calculated as weight (in kg) divided by height (in m) squared. BMI values were categorized according to standardized criteria (56).

#### ***Blood pressure measurements***

Blood pressure was measured on the left arm at heart level using an automatic digital blood pressure monitor (Omron model HBP-1320, Omron Healthcare Co., Japan). The first measurement was performed after the participant rested for 15 min in a sitting position, then two additional measurements were performed 3 minutes apart. Results of the systolic and diastolic pressures were expressed in millimeter of mercury (mmHg).

#### ***Skin pigmentation measurements***

Skin pigmentation was measured on the inner arm between biceps and triceps using the DSM III Skin colorimeter (Cortex technology, Denmark). Four measurement replicates were performed for each participant. Colors were quantified using the L\*, a\*, and b\* parameters defined by the International commission on illumination (CIE, Vienna, Austria), where L\* measures the degree of brightness or whiteness, a\* the degree of redness, and b\* the yellowness. Skin erythema 'E' and melanin content 'M' were also measured (57, 58).

### **Biological sample collection and biobanking**

#### ***Blood***

A capillary whole blood sample was collected at the finger tip of each fasting participant using microtainer contact-activated lancets (Becton Dickinson, USA), and was immediately used for the measurement of blood glucose concentration (mg/dL) with the Freestyle Optium Neo H point-of-care device (Abbott Diabetes Care, France).

Blood samples were also collected by venipuncture in four 5 mL serum separation tubes, and in two 5 mL EDTA tubes (Becton Dickinson, USA). Serum separation tubes and EDTA tubes were maintained at 4 °C until centrifugation during 15 minutes at 3,500 and 2,000 rpm, respectively. Part of the serum and plasma samples were used by the clinical laboratory of ILM to measure several hematological and biochemical parameters as described in Table 2. The remaining samples of serum,

plasma and red blood cell pellets were aliquoted in several tubes and stored at -20 °C until processing as detailed in Table 3.

### **Saliva**

A 2 mL sample of saliva was self-collected by each participant in an Oragene DNA tube pre-filled with 2 mL of stabilization liquid following the manufacturer's protocol (DNA Genotek, Canada). After gentle mix, the sample tube was stored at 4 °C until processing (Table 3).

### **Stool**

Stool samples were self-collected by participants in tubes pre-filled with 8 mL of DNA stabilization buffer following the manufacturer's instructions (Canvax Biotech, Spain), and stored in a fridge within 24 hours before the appointment with the nurse. Then, stool samples were stored at 4 °C until shipment to ILM. Upon reception at ILM, stool samples were stored at -20 °C until processing (Table 3).

### **Data management**

The data flow and processing of biological samples are detailed in Figure 3.

Each participant was assigned a unique identification number to anonymize the information collected in the CRF and biological samples. Personal identifying information such as last name, first name, and phone number only appeared on the consent form, which has been kept in a secure location accessible only to the biologists of the clinical laboratory of ILM.

Data collected in the CRF were transferred from electronic tablets to a password-protected ODK Aggregate server (Version 2.0.5, Get ODK Inc., USA). Then, these data were uploaded to a password-protected Redcap server (Version 9.5.11, Vanderbilt University, USA) used to aggregate all data from each participant. Results of hematological and biochemical analyses performed by the clinical laboratory of ILM were also uploaded to the Redcap server, and linked to questionnaire data using the unique identification number of each participant. The results of the analyses that will be performed later from the biological samples of the participants will also be uploaded to the Redcap server. All collaborators of the MATAEA project were given a personal and unique login and password to access data in the Redcap server.

The staff of ILM received training on ethical procedures and measures to protect confidentiality needed, and was required to sign a confidentiality agreement form. The Redcap database was validated according to the guidelines of the French National Commission of Data Protection and Individual Liberties, allowing collaboration with external researchers and collaborators.

### **Data Analysis**

According to the data retrieved from the participants, indicators will be defined to identify subjects suffering from, or having a history, of any pathology targeted by the study.

For continuous data, summary statistics will include the arithmetic mean with 95% confidence interval (CI) or standard deviation (SD), or median with inter quartile range. For categorical data, frequency counts, percentages, and 95% CIs will be presented. The characteristics of the subjects will be compared according to sex, age category, and geographical area using the chi-square test for the categorical variables, and using the analysis of variance (ANOVA) for the continuous variables.

Non-standardized and standardized (weighted) results will be provided. Weights will be estimated by archipelago (except the Society archipelago for which the Windward Islands and Leeward Islands will be treated separately) and will correspond to the inverse of the probability for each subject to be selected, where the probability in a given archipelago is the number of subjects in each of the 6 age/sex strata divided by the total number population in that strata.

Risk factors associated with each pathology (if the prevalence is high enough to justify a statistical analysis) will be identified using a logistic regression model. The factors investigated will

include age, sex, and all other relevant factors. These analyses will account for the sampling plan through consideration of survey weights.

### **Return of individual results to participants**

Results of physical measurements performed by a nurse were provided to each participant during the appointment. If any abnormal result was detected, the nurse advised the participant to consult a doctor immediately.

Results of hematological and biochemical analyses performed by the clinical laboratory of ILM from blood samples of each participant were available on a secured server. Participants could retrieve these results online using individual login valid for seven days, or anytime at ILM providing proof of identity. If any abnormal result was detected, the participant or the participant's attending doctor (depending on the participant's choice as indicated in the informed consent) was notified by a biologist of the clinical laboratory of ILM.

Participants will not be informed of individual research results that will be obtained by the specialized laboratories involved in the MATAEA project. However, if any research result suggests an underlying pathology, the participant or the participant's attending doctor (depending on the participant's choice as indicated in the informed consent) will be notified by a biologist of the clinical laboratory of ILM to carry out additional investigations.

### **STRENGTHS AND LIMITATIONS**

The MATAEA project is the first cross-sectional survey designed to investigate a broad range of health issues in the five archipelagoes of French Polynesia. The high number of participants and their wide geographical distribution guarantee a good representativeness of the population of French Polynesia. The large amount of data obtained for each participant from the CRF, physical measurements and biological analyses will make it possible to update the information on the prevalence of numerous NCDs and CDs known as major public health concerns in French Polynesia, and possibly identify specific indicators and/or risk factors for these diseases. The study will also provide new epidemiological information on poorly studied or unknown pathologies. Finally, the MATAEA project will generate a high amount of genetic data for a population that has been so far poorly represented in human genetic studies.

Nevertheless, the MATAEA project has some limitations that should be acknowledged. First, the population aged under 18 years was not included to the survey, which might have an impact on the estimate of the prevalence of certain pathologies. Second, participants were recruited in only 18 of the 75 inhabited islands while differences may exist between the inhabitants of distinct islands within the same archipelago. Third, there might be an underrepresentation of hard-to-reach subjects, such as heavy drinkers or socially excluded persons, which might be more prone to some pathologies. Fourth, the number of participants recruited for the study may be not sufficient to investigate rare pathologies. Fifth, the cross-sectional nature of the survey does not allow to examine the temporal relationship between exposures and diseases under study. Finally, the sampling period has been interrupted three times because of the COVID-19 pandemic and more than two years elapsed between the first and last inclusion (Figure 2).

### **COMMUNITY ENGAGEMENT**

The main objectives of the MATAEA project have already been presented to the Health and Research authorities of French Polynesia as well as, in several occasions, to the general French Polynesian populations, in particular in the form of conferences principally conducted by Van-Mai Cao-Lormeau and Lluís Quintana-Murci. Furthermore, a video presenting the MATAEA project is already accessible online (<https://youtube.com/playlist?list=PLVLsoPZrkK4ipioiTOhGh0z05XZSHPC6H>). All the results generated in the framework of the MATAEA project will be also made available to the

population of French Polynesia through various communication channels. The results will be presented in an accessible way for the public, through conferences, seminars, and roundtables, and will be sent to the mayors of the islands that participated in the study to be relayed to the community. Finally, the public will be informed of the results published in scientific journals through popularized summaries translated into both French and Tahitian.

## ETHICS STATEMENT

This study involving human participants was approved by the *French Comité de protection des personnes* (CPP - OUEST III no. 19.08.60 / SI CNRIPH 19.07.02.38421) and the *Comité d’Ethique de la Polynésie française* (Avis no. 80 CEPF-03/09/2019). Written informed consent was obtained from all participants. People who were unable to express their consent or answer the questionnaire were excluded from the study. Data collection, storage, transfer and process were executed according to the European general data protection regulation procedures.

## CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## AUTHOR CONTRIBUTIONS

AF, LQM, AS, and VMCL designed and conceptualized the project; AF, LQM, EP, AS, JV, NP, JT, SO, SL, HB, PB, CG, ES, BC and PA participated in the design of the study protocol; IT, MA, and VMCL wrote the study protocol; SFP and NJ managed the ethical aspects of the project; NP and JT carried out the draw of participants; IT, MA, and VMCL supervised and coordinated the project; IT, EP, AJ, MR, VM, and KC performed the data management; CH and AB processed biological samples; GRL, DL and EP designed genetic analyses; VMCL acquired funding for the project; IT and MA drafted the manuscript. All authors read and approved the submitted version of the manuscript.

## FUNDING

The MATAEA project received financial support from *La Délégation à la Recherche de la Polynésie française* (Convention no. 03557/MED/REC du 29 mai 2019) and the *Contrat de Projet Etat-Pays* 2015-2020 (Arrêté no. HC/372/DIE/BPT du 18 mai 2018; Convention no. 03298/MTF/REC du 17 mai 2018). The costs related to the shipping and analysis of the samples will be borne by the laboratories involved in the project.

## ACKNOWLEDGEMENTS

We thank the staff of all the laboratories that contributed to the realization of this study. We also acknowledge the municipal staff and the guides in the islands selected for the study for their support in recruiting participants. Finally, we especially thank all the individuals who willingly agreed to participate in this study

**Table 1. Geographical distribution of the 2,100 participants randomly selected for the MATAEA project**

| Strata               |                      |             | Archipelago | Island    |                      |             |
|----------------------|----------------------|-------------|-------------|-----------|----------------------|-------------|
| Name                 | Population size (%)* | Sample size | Name        | Name      | Population size (%)* | Sample size |
| Windward Islands     | 207,333 (75%)        | 700         | Society     | Tahiti    | 189,517 (91%)        | 637         |
|                      |                      |             |             | Moorea    | 17,816 (9%)          | 63          |
| Leeward Islands      | 35,393 (13%)         | 700         | Society     | Raiatea   | 12,249 (36%)         | 252         |
|                      |                      |             |             | Bora Bora | 10,549 (31%)         | 217         |
|                      |                      |             |             | Tahaa     | 5,234 (15%)          | 105         |
|                      |                      |             |             | Huahine   | 6,075 (18%)          | 126         |
| Others archipelagoes | 33,192 (12%)         | 700         | Tuamotu     | Rangiroa  | 3,657 (16%)          | 112         |
|                      |                      |             |             | Fakarava  | 1,637 (7%)           | 49          |
|                      |                      |             |             | Hao       | 1,258 (5%)           | 35          |
|                      |                      |             |             | Makemo    | 1,508 (6%)           | 42          |
|                      |                      |             | Gambier     | Mangareva | 1,535 (7%)           | 49          |
|                      |                      |             | Marquesas   | Nuku Hiva | 2,951 (13%)          | 91          |
|                      |                      |             |             | Hiva Oa   | 2,243 (10%)          | 70          |
|                      |                      |             |             | Ua Pou    | 2,213 (9%)           | 63          |
|                      |                      |             | Austral     | Tubuai    | 2,217 (9%)           | 63          |
|                      |                      |             |             | Rurutu    | 2,466 (10%)          | 70          |
|                      |                      |             |             | Raivavae  | 903 (4%)             | 28          |
|                      |                      |             |             | Rimatara  | 872 (4%)             | 28          |
| <b>Total</b>         | 275,918              | 2,100       |             |           |                      | 2,100       |

\* data from the population census of 2017 (59)

514 **Table 2. List of hematological and biochemical parameters analyzed for the MATAEA project**  
515

| Group                | Measured parameter                        | Units    | Test                 | Analyzer                            |
|----------------------|-------------------------------------------|----------|----------------------|-------------------------------------|
| Complete blood count | Hemoglobin                                | g/dL     | Spectrophotometry    | XT-4000 or XN-1000, Sysmex (France) |
|                      | Hematocrit                                | %        | Electrical impedance | XT-4000 or XN-1000, Sysmex (France) |
|                      | Leucocytes count                          | G/L      | Fluorocytometry      | XT-4000 or XN-1000, Sysmex (France) |
|                      | Neutrophils count                         | G/L or % | Fluorocytometry      | XT-4000 or XN-1000, Sysmex (France) |
|                      | Eosinophils count                         | G/L or % | Fluorocytometry      | XT-4000 or XN-1000, Sysmex (France) |
|                      | Lymphocytes count                         | G/L or % | Fluorocytometry      | XT-4000 or XN-1000, Sysmex (France) |
|                      | Monocytes count                           | G/L or % | Fluorocytometry      | XT-4000 or XN-1000, Sysmex (France) |
|                      | Platelets count                           | G/L      | Electrical impedance | XT-4000 or XN-1000, Sysmex (France) |
|                      | Mean platelet volume                      | fL       | Electrical impedance | XT-4000 or XN-1000, Sysmex (France) |
|                      | Red blood cells count                     | T/L      | Electrical impedance | XT-4000 or XN-1000, Sysmex (France) |
|                      | Red blood cell distribution width         | -        | Electrical impedance | XT-4000 or XN-1000, Sysmex (France) |
|                      | Mean corpuscular volume                   | fL       | Electrical impedance | XT-4000 or XN-1000, Sysmex (France) |
|                      | Mean corpuscular hemoglobin               | pg/cell  | Spectrophotometry    | XT-4000 or XN-1000, Sysmex (France) |
|                      | Mean corpuscular hemoglobin concentration | g/dL     | Spectrophotometry    | XT-4000 or XN-1000, Sysmex (France) |
| Serum aspect         | Serum aspect                              | -        | Visual assessment    | -                                   |
|                      | Serum index                               | UI       | Spectrophotometry    | Cobas c501, Roche (Switzerland)     |

|                               |                             |                            |                                                                          |                                             |
|-------------------------------|-----------------------------|----------------------------|--------------------------------------------------------------------------|---------------------------------------------|
| <b>Diabetes</b>               | Glycated hemoglobin A1c     | %                          | Capillary eletrophoresis                                                 | Capillarys 2 Flex Piercing, SEBIA (France)  |
|                               | Fasting blood glucose level | mg/dL                      | Fingertip lancing technique                                              | Optium Neo H, Abbott Diabetes Care (France) |
| <b>Lipids</b>                 | Total cholesterol           | mmol/l                     | Enzymatic colorimetric assay                                             | Cobas c501, Roche (Switzerland)             |
|                               | HDL-cholesterol             | mmol/l                     | Homogenous enzymatic colorimetric assay                                  | Cobas c501, Roche (Switzerland)             |
|                               | LDL-cholesterol             | mmol/l                     | Friedewald formula (60)                                                  | Cobas c501, Roche (Switzerland)             |
|                               | Triglycerides               | mmol/l                     | Enzymatic colorimetric assay                                             | Cobas c501, Roche (Switzerland)             |
| <b>Kidney function</b>        | Creatinine                  | μmol/L                     | Enzymatic colorimetric assay                                             | Cobas c501, Roche (Switzerland)             |
|                               | Glomerular filtration rate  | mL/min/1.73 m <sup>2</sup> | Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) formula (61) | Cobas c501, Roche (Switzerland)             |
| <b>Liver and inflammation</b> | Aspartate aminotransferase  | UI/L                       | Enzymatic colorimetric assay                                             | Cobas c501, Roche (Switzerland)             |
|                               | Alanine aminotransferase    | UI/L                       | Enzymatic colorimetric assay                                             | Cobas c501, Roche (Switzerland)             |
|                               | Gamma glutamyl transferase  | UI/L                       | Enzymatic colorimetric assay                                             | Cobas c501, Roche (Switzerland)             |
|                               | Total bilirubin             | μmol/L                     | Photometry assay                                                         | Cobas c501, Roche (Switzerland)             |
|                               | C-reactive protein          | mg/L                       | Particle-enhanced immunoturbidimetric assay                              | Cobas c501, Roche (Switzerland)             |

**Table 3. List of laboratory analyses planned in the MATAEA project**

| Sample type                    | Analysis                               | Laboratory                          |
|--------------------------------|----------------------------------------|-------------------------------------|
| Serum and plasma               | Hematological and biochemical analyses | LABM, ILM, Tahiti, French Polynesia |
| Serum                          | Arbovirus serology                     | LIV, ILM, Tahiti, French Polynesia  |
| Serum                          | Mosquito saliva serology               | LEM, ILM, Tahiti, French Polynesia  |
| Serum                          | Hantavirus serology                    | EIRU, IP, Paris, France             |
| Serum                          | VirScan (62)                           | HEGU, IP, Paris, France             |
| Plasma                         | Hepatitis B and C serologies           | CHPF, Tahiti, French Polynesia      |
| Plasma                         | Pesticides quantitation                | INSPQ, Montreal, Canada             |
| Red blood cells pellet         | Heavy metals quantitation              | INSPQ, Montreal, Canada             |
| Red blood cells pellet, saliva | Genetic analysis                       | HEGU, IP, Paris, France             |
| Saliva                         | Oral microbiome analysis               | HEGU, IP, Paris, France             |
| Stool                          | Intestinal microbiome analysis         | HEGU, IP, Paris, France             |

LABM: Clinical Laboratory; ILM: Institut Louis Malardé; LIV: Laboratory of research on emerging viral diseases; LEM: Laboratory of research in medical entomology; IP: Institut Pasteur; EIRU: Environment and Infectious Risk Unit ; HEGU : Human Evolutionary Genetics Unit ; CHPF: Centre Hospitalier de la Polynésie française; INSPQ: Institut National de Santé Publique du Québec.

## References

1. Institut de la Statistique de la Polynésie française. Points fort de la Polynésie française, Le recensement de la population en Polynésie française en 2017. (2018). Available online at: <https://data.ispf.pf/docs/default-source/publi-pf-bilans-et-etudes/pf-etudes-7-2018-rp-2017.pdf?sfvrsn=8> (Accessed March 20, 2023).
2. Gosling AL, Matisoo-Smith EA. The evolutionary history and human settlement of Australia and the Pacific. *Curr Opin Genet Dev.* 2018;53:53-9. doi: 10.1016/j.gde.2018.06.015.
3. Ioannidis AG, Blanco-Portillo J, Sandoval K, Hagelberg E, Barberena-Jonas C, Hill AVS, et al. Paths and timings of the peopling of Polynesia inferred from genomic networks. *Nature.* 2021;597(7877):522-6. doi: 10.1038/s41586-021-03902-8.
4. Choin J, Mendoza-Revilla J, Arauna LR, Cuadros-Espinoza S, Cassar O, Larena M, et al. Genomic insights into population history and biological adaptation in Oceania. *Nature.* 2021;592(7855):583-9. doi: 10.1038/s41586-021-03236-5.
5. Hudjashov G, Endicott P, Post H, Nagle N, Ho SYW, Lawson DJ, et al. Investigating the origins of eastern Polynesians using genome-wide data from the Leeward Society Isles. *Sci Rep.* 2018;8(1):1823. doi: 10.1038/s41598-018-20026-8.
6. Kirch PV. Modern Polynesian genomes offer clues to early eastward migrations. *Nature.* 2021;597(7877):477-8. doi: 10.1038/d41586-021-01719-z.
7. Horwood PF, Tarantola A, Goarant C, Matsui M, Klement E, Umezaki M, et al. Health Challenges of the Pacific Region: Insights From History, Geography, Social Determinants, Genetics, and the Microbiome. *Front Immunol.* 2019;10:2184. doi: 10.3389/fimmu.2019.02184.
8. Barreiro LB, Quintana-Murci L. Evolutionary and population (epi)genetics of immunity to infection. *Hum Genet.* 2020;139(6-7):723-32. doi: 10.1007/s00439-020-02167-x.
9. Quach H, Rotival M, Pothlichet J, Loh YE, Dannemann M, Zidane N, et al. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. *Cell.* 2016;167(3):643-56.e17. doi: 10.1016/j.cell.2016.09.024.
10. Quintana-Murci L. Human Immunology through the Lens of Evolutionary Genetics. *Cell.* 2019;177(1):184-99. doi: 10.1016/j.cell.2019.02.033.
11. Gouvernement de la Polynésie française. La politique de santé de la Polynésie française - Orientations stratégiques 2016-2025. (2016). Available online at: <https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2017/06/La-politique-de-Sante%CC%81-en-PF-2016-2025-web.pdf> (Accessed March 20, 2023).
12. Institut de la Statistique de la Polynésie française. Bilan démographique 2020. (2021). Available online at: [https://www.ispf.pf/content/uploads/1256\\_Bilan\\_demographique\\_2020\\_f9055f2afa.pdf](https://www.ispf.pf/content/uploads/1256_Bilan_demographique_2020_f9055f2afa.pdf) (Accessed March 20, 2023).
13. Reeve E, Lamichhane P, McKenzie B, Waqa G, Webster J, Snowdon W, et al. The tide of dietary risks for noncommunicable diseases in Pacific Islands: an analysis of population NCD surveys. *BMC Public Health.* 2022;22(1):1521. doi: 10.1186/s12889-022-13808-3.
14. Bertrand S, Berry A-L. Enquête santé 2010 en Polynésie française : surveillance des facteurs de risque des maladies non transmissibles. *Bull Épidémiol Hebd.* 2013;28-29:326-32.

- 565 15. Direction de la Santé. Statistiques des causes de décès en Polynésie française 2011-2015.  
566 (2019). Available online at: [https://www.service-public.pf/dsp/wp-](https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2020/02/Rapport_StatistiquesCausesDC2011_15PF.pdf)  
567 [content/uploads/sites/12/2020/02/Rapport\\_StatistiquesCausesDC2011\\_15PF.pdf](https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2020/02/Rapport_StatistiquesCausesDC2011_15PF.pdf) (Accessed March  
568 20, 2023).
- 569 16. Institut d'Emission d'Outre-Mer. Rapport annuel 2017 de l'IEOM - Polynésie française.  
570 (2018). Available online at:  
571 [https://www.ieom.fr/IMG/pdf/rapport\\_annuel\\_2017\\_ieom\\_polynesie\\_francaise\\_-\\_juin\\_2018.pdf](https://www.ieom.fr/IMG/pdf/rapport_annuel_2017_ieom_polynesie_francaise_-_juin_2018.pdf)  
572 (Accessed March 20, 2023).
- 573 17. Direction de la Santé. Cancers en Polynésie française - Statistiques du Registre des cancers.  
574 (2016). Available online at: [https://www.service-public.pf/dsp/wp-](https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2021/12/Rapport-cancer-PF-2016-.pdf)  
575 [content/uploads/sites/12/2021/12/Rapport-cancer-PF-2016-.pdf](https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2021/12/Rapport-cancer-PF-2016-.pdf) (Accessed March 20, 2023).
- 576 18. Tervonen H, Foliaki S, Bray F, Roder D. Cancer epidemiology in the small nations of Pacific  
577 Islands. *Cancer epidemiology*. 2017;50(Pt B):184-92. doi: 10.1016/j.canep.2017.09.002.
- 578 19. World Health Organisation. Global Cancer Observatory - French Polynesia in 2020. (2021).  
579 Available online at: [https://gco.iarc.fr/today/data/factsheets/populations/258-french-polynesia-fact-](https://gco.iarc.fr/today/data/factsheets/populations/258-french-polynesia-factsheets.pdf)  
580 [sheets.pdf](https://gco.iarc.fr/today/data/factsheets/populations/258-french-polynesia-factsheets.pdf) (Accessed March 20, 2023).
- 581 20. Condat B, Lascols A, Lutringer-Magnin D, Chakhtoura F, Lastere S, Bronstein J.  
582 Épidémiologie du carcinome hépatocellulaire en Polynésie française. *Bull Épidémiol Hebd*.  
583 2022;16:280-7.
- 584 21. Chinain M, Gatti CM, Roué M, Darius HT. Ciguatera poisoning in French Polynesia: insights  
585 into the novel trends of an ancient disease. *New microbes and new infections*. 2019;31:100565. doi:  
586 10.1016/j.nmni.2019.100565.
- 587 22. Chinain M, Gatti CM. Surveillance de la ciguatera en Polynésie française - Bilan 2021.  
588 (2022). Available online at: [https://ciguatera.pf/images/BILAN\\_CIGUATERA\\_PF\\_2021\\_.pdf](https://ciguatera.pf/images/BILAN_CIGUATERA_PF_2021_.pdf)  
589 (Accessed March 20, 2023).
- 590 23. St-Jean A, Suhas E, De Pina JJ, Cordier S, Lucas M, Ayotte P. Exposure to toxic metals and  
591 polychlorinated biphenyls of adolescents and adults from two atolls in the Tuamotu Archipelago  
592 (French Polynesia). *The Science of the total environment*. 2019;695:133791. doi:  
593 10.1016/j.scitotenv.2019.133791.
- 594 24. Dewailly E, Chateau-Degat L, Suhas E. Fish consumption and health in French Polynesia.  
595 *Asia Pacific journal of clinical nutrition*. 2008;17(1):86-93.
- 596 25. Dewailly E, Suhas E, Mou Y, Dallaire R, Chateau-Degat L, Chansin R. High fish  
597 consumption in French Polynesia and prenatal exposure to metals and nutrients. *Asia Pacific journal*  
598 *of clinical nutrition*. 2008;17(3):461-70.
- 599 26. Roche H, Salvat B, Ramade F. Assessment of the pesticides pollution of coral reefs  
600 communities from French Polynesia. *Revue d'Ecologie (Terre Vie)*. 2011;66(1):3-10. doi:  
601 10.3406/revec.2011.1553.
- 602 27. Salvat B, Roche H, Berny P, Ramade F. Recherches sur la contamination par les pesticides  
603 d'organismes marins des réseaux trophiques récifaux de Polynésie Française. *Revue d'Ecologie*  
604 *(Terre Vie)*. 2012;67(2):129-47. doi: 10.3406/revec.2012.1629.
- 605 28. Salvat B, Roche H, Ramade F. On the occurrence of a widespread contamination by  
606 herbicides of coral reef biota in French Polynesia. *Environ Sci Pollut Res Int*. 2016;23(1):49-60. doi:  
607 10.1007/s11356-015-4395-9.

29. Hubert B. Situation des maladies infectieuses en Polynésie française en 2001. 2001. doi: 10.13140/RG.2.1.3451.2804.
30. Direction de la Santé. La tuberculose en Polynésie française - Rapport annuel 2018. (2018). Available online at: <https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2019/04/Rapport-tuberculose-2018.pdf> (Accessed March 20, 2023).
31. Musso D, Roveery C, Loukil A, Vialette V, Nguyen NL. Leprosy in French Polynesia. New microbes and new infections. 2019;29:100514. doi: 10.1016/j.nmni.2018.10.010.
32. Ichimori K, Graves PM. Overview of PacELF-the Pacific Programme for the Elimination of Lymphatic Filariasis. Tropical medicine and health. 2017;45:34. doi: 10.1186/s41182-017-0075-4.
33. Kline K, McCarthy JS, Pearson M, Loukas A, Hotez PJ. Neglected tropical diseases of Oceania: review of their prevalence, distribution, and opportunities for control. PLoS neglected tropical diseases. 2013;7(1):e1755. doi: 10.1371/journal.pntd.0001755.
34. Mou Y, Plichart C, Legrand A, Mallet H, Cerf N, Nguyen L. Evaluation de la prevalence de la filariose lymphatique en 2008 en Polynésie française. Bull Epidemiol Hebd. 2009;48-49-50:504-7.
35. Segalin JM, Le Calvez E. Programme de lutte contre les maladies tropicales négligées en Polynésie française 2019-2023. (2019). Available online at: <https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2019/11/2019-03-14-PROGRAMME-MTN.pdf> (Accessed March 20, 2023).
36. Boutin JP, Sainte Marie FF, Cartel JL, Cardines R, Girard M, Roux J. Prevalence of hepatitis B virus infection in the Austral archipelago, French Polynesia: identification of transmission patterns for the formulation of immunization strategies. Trans R Soc Trop Med Hyg. 1990;84(2):283-7. doi: 10.1016/0035-9203(90)90288-p.
37. Moulia-Pelat J, Spiegel A, Excler J, Martin P, Roux J, Boutin J, et al. Lutte contre l'hépatite B en Polynésie française par un programme de vaccination systématique des nouveau-nés avec le vaccin Genhevac B®. Cahiers d'études et de recherches francophones / Santé. 1996;6(1):11-5.
38. Gouvernement de la Polynésie Française. Délibération n° 95-63 AT du 23 mai 1995 portant réglementation des vaccinations contre certaines maladies transmissibles chez l'enfant. (1995). Available online at: <https://lexpol.cloud.pf/document.php?document=190635&deb=1202&fin=1203&titre=RMOpbGliw6lyYXRpb24gbsKwIDk1LTZyZIEFUIGR1IDZLzA1LzE5OTU=> (Accessed March 20, 2023).
39. Patel MK, Le Calvez E, Wannemuehler K, Ségalin JM. Hepatitis B Vaccination Coverage and Prevalence of Hepatitis B Surface Antigen Among Children in French Polynesia, 2014. The American journal of tropical medicine and hygiene. 2016;94(6):1370-5. doi: 10.4269/ajtmh.15-0903.
40. Aubry M, Cao-Lormeau VM. History of arthropod-borne virus infections in French Polynesia. New microbes and new infections. 2019;29:100513. doi: 10.1016/j.nmni.2019.01.009.
41. Aubry M, Mapotoe M, Teissier A, Paoaafaite T, Dumas-Chastang E, Giard M, et al. Dengue virus serotype 2 (DENV-2) outbreak, French Polynesia, 2019. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2019;24(29). doi: 10.2807/1560-7917.Es.2019.24.29.1900407.
42. Cao-Lormeau VM, Roche C, Teissier A, Robin E, Berry AL, Mallet HP, et al. Zika virus, French polynesia, South pacific, 2013. Emerging infectious diseases. 2014;20(6):1085-6. doi: 10.3201/eid2006.140138.

43. Aubry M, Teissier A, Roche C, Richard V, Yan AS, Zisou K, et al. Chikungunya outbreak, French Polynesia, 2014. *Emerging infectious diseases*. 2015;21(4):724-6. doi: 10.3201/eid2104.141741.
44. Cao-Lormeau VM, Blake A, Mons S, Lastère S, Roche C, Vanhomwegen J, et al. Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet (London, England)*. 2016;387(10027):1531-9. doi: 10.1016/s0140-6736(16)00562-6.
45. Besnard M, Eyrolle-Guignot D, Guillemette-Artur P, Lastère S, Bost-Bezeaud F, Marcelis L, et al. Congenital cerebral malformations and dysfunction in fetuses and newborns following the 2013 to 2014 Zika virus epidemic in French Polynesia. *Euro surveillance : bulletin Européen sur les maladies transmissibles = European communicable disease bulletin*. 2016;21(13). doi: 10.2807/1560-7917.Es.2016.21.13.30181.
46. Barboza P, Baudon C, Chérié-Challine L, Gastellu-Etchegorry M, Gueguen J, La Ruche G, et al. Influenza A(H1N1)2009 in the French Pacific territories: assessment of the epidemic wave during the austral winter. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2010;16(4):304-8. doi: 10.1111/j.1469-0691.2010.03172.x.
47. Aubry M, Maset N, Chapman L, Simon A, Olivier S, Bos R, et al. Seroprevalence of SARS-CoV-2 Antibodies in French Polynesia and Perspective for Vaccine Strategies. *Preprints*. 2022. doi: 10.20944/preprints202212.0386.v1.
48. Direction de la Santé. La leptospirose en Polynésie française – Rapport annuel 2018. (2018). Available online at: [https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2022/04/190314\\_Rapport-leptospirose-2018.pdf](https://www.service-public.pf/dsp/wp-content/uploads/sites/12/2022/04/190314_Rapport-leptospirose-2018.pdf) (Accessed March 20, 2023).
49. Lévy M, Baudet J, Beau F, Devaud F, Sauget A. Étude de prévalence de la sensibilité aux antibiotiques de *Staphylococcus aureus* en Polynésie française, 2006-2008. *Bull Epidemiol Hebd Thématique*. 2009;48-49-50:526-28.
50. Leitwein K. Evaluation de la mise en place d'une méthode de dépistage échographique simplifiée des cardiopathies rhumatismales chroniques chez les élèves de CM2 des années 2012/2013 et 2013/2014 à Moorea en Polynésie Française 2017. Available online at: <https://dumas.ccsd.cnrs.fr/dumas-01534602/document> (Accessed March 20, 2023).
51. Aubry M, Teissier A, Huart M, Merceron S, Vanhomwegen J, Mapotoeke M, et al. Seroprevalence of Dengue and Chikungunya Virus Antibodies, French Polynesia, 2014-2015. *Emerging infectious diseases*. 2018;24(3):558-61. doi: 10.3201/eid2403.171149.
52. Aubry M, Teissier A, Huart M, Merceron S, Vanhomwegen J, Roche C, et al. Zika Virus Seroprevalence, French Polynesia, 2014-2015. *Emerging infectious diseases*. 2017;23(4):669-72. doi: 10.3201/eid2304.161549.
53. Mathieu-Daudé F, Claverie A, Plichart C, Boulanger D, Mphande FA, Bossin HC. Specific human antibody responses to *Aedes aegypti* and *Aedes polynesiensis* saliva: A new epidemiological tool to assess human exposure to disease vectors in the Pacific. *PLoS neglected tropical diseases*. 2018;12(7):e0006660. doi: 10.1371/journal.pntd.0006660.
54. Chungue E, Rollin PE, Saluzzo JF, Gentile B, Roux JF, Sureau P. Serological survey of Hantavirus infection in French Polynesia. *Trans R Soc Trop Med Hyg*. 1989;83(2):258. doi: 10.1016/0035-9203(89)90669-x.

692 55. World Health Organisation. Standard STEPS instrument. (2020). Available online at:  
693 <https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/steps/instrument>  
694 (Accessed March 20, 2023).

695 56. World Health Organisation. Obesity: preventing and managing the global epidemic. Report of  
696 a WHO consultation. World Health Organization technical report series. 2000;894:i-xii, 1-253.

697 57. Matias AR, Ferreira M, Costa P, Neto P. Skin colour, skin redness and melanin biometric  
698 measurements: comparison study between Antera(®) 3D, Mexameter(®) and Colorimeter(®). Skin  
699 research and technology : official journal of International Society for Bioengineering and the Skin  
700 (ISBS) [and] International Society for Digital Imaging of Skin (ISDIS) [and] International Society  
701 for Skin Imaging (ISSI). 2015;21(3):346-62. doi: 10.1111/srt.12199.

702 58. Wong LC, Jamil A, Md Nor N. The effect of pre-phototest sun exposure on minimal  
703 erythema dose and minimal melanogenic dose among skin phototypes III, IV and V.  
704 Photodermatology, photoimmunology & photomedicine. 2018;34(6):400-4. doi:  
705 10.1111/phpp.12410.

706 59. Institut de la Statistique de la Polynésie française. Recensement de la population 2017-  
707 Données détaillées - Population par Géographie administrative et par Age décennal. (2018).  
708 Available online at: [https://data.ispf.pf/bases/Recensements/2017/Donnees\\_detaillees.aspx](https://data.ispf.pf/bases/Recensements/2017/Donnees_detaillees.aspx) (Accessed  
709 March 20, 2023).

710 60. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density  
711 lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem.  
712 1972;18(6):499-502. doi: 10.1186/s12944-019-1178-1.

713 61. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new  
714 equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12. doi:  
715 10.7326/0003-4819-150-9-200905050-00006.

716 62. Xu GJ, Kula T, Xu Q, Li MZ, Vernon SD, Ndung'u T, et al. Viral immunology.  
717 Comprehensive serological profiling of human populations using a synthetic human virome. Science  
718 (New York, NY). 2015;348(6239):aaa0698. doi: 10.1126/science.aaa0698.

719

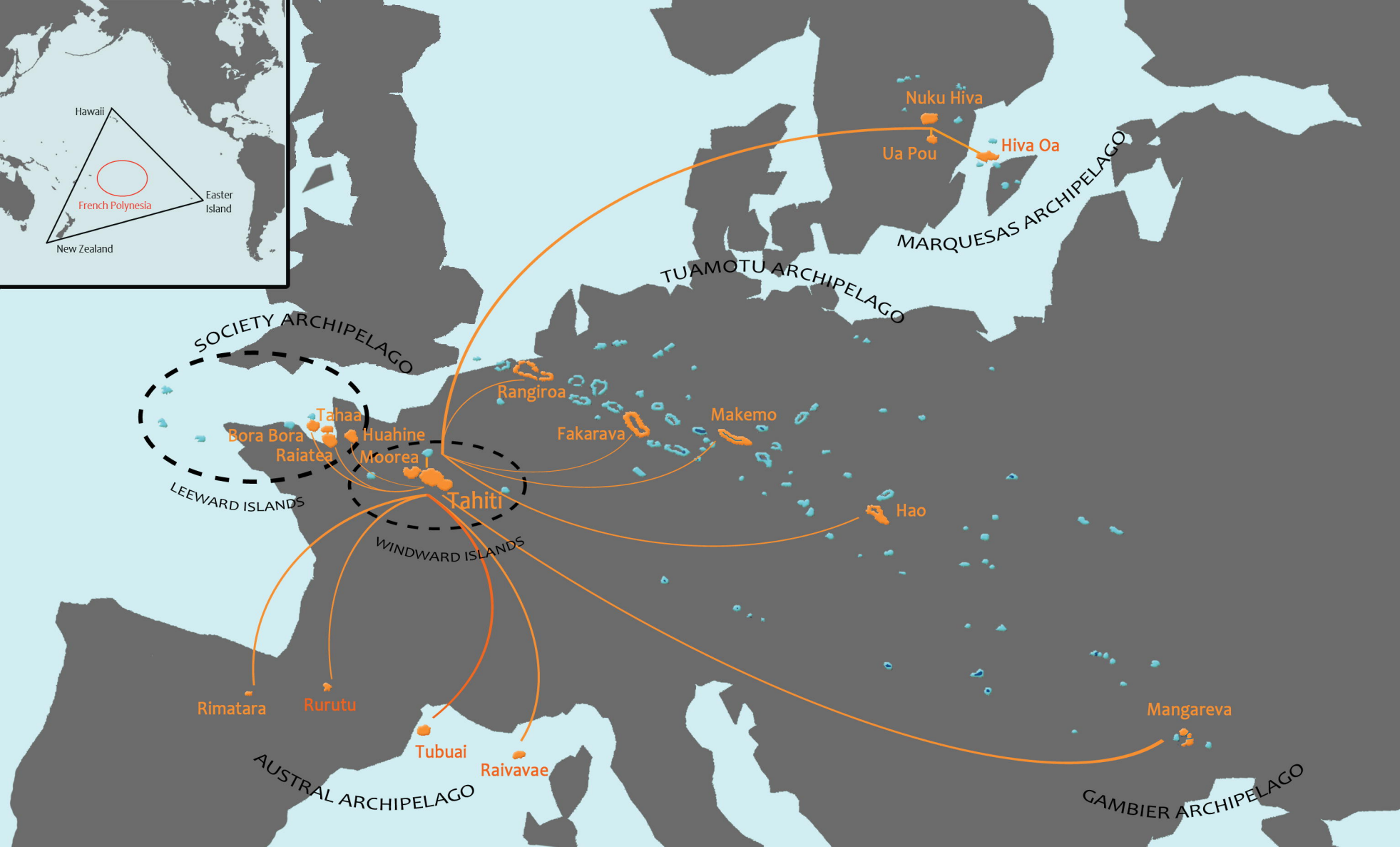
720

721 **Figure 1.** Map of French Polynesia showing the geographical distribution of the five archipelagoes  
 722 (Society, Tuamotu, Gambier, Marquesas and Austral) compared to the European continent (showed at  
 723 the same scale). Flight connections between Tahiti and the islands selected for the survey are indicated  
 724 by orange lines. Inset map at upper left shows location of French Polynesia (red circle) within the  
 725 Polynesian triangle (black triangle) bounded by the islands of Hawaii, New Zealand and Easter Island  
 726 in the Pacific Ocean.

728 **Figure 2.** Study timeline for the MATAEA project. Recruitment of participants in the 18 islands of  
 729 French Polynesia started on November 4, 2019, and ended on December 8, 2021. Three breaks  
 730 occurred during that period because of SARS-CoV-2 circulation: (1) population lockdown from  
 731 March 20 to May 21, 2020; (2) first epidemic wave due to the SARS-CoV-2 Wuhan strain from  
 732 August 2020 to February 2021; (3) second epidemic wave due the SARS-CoV-2 Delta variant from  
 733 August to September 2021.

734 **Figure 3.** Data flow and biological samples processing for the MATAEA project. LABM: Clinical  
 735 Laboratory; ILM: Institut Louis Malardé; LIV: Laboratory of research on emerging viral diseases; IP:  
 736 Institut Pasteur.

737



SOCIETY ARCHIPELAGO

TUAMOTU ARCHIPELAGO

MARQUESAS ARCHIPELAGO

GAMBIER ARCHIPELAGO

AUSTRAL ARCHIPELAGO

LEEWARD ISLANDS

WINDWARD ISLANDS

Nuku Hiva

Ua Pou

Hiva Oa

Rangiroa

Fakarava

Makemo

Hao

Mangareva

Tahiti

Tahaa

Bora Bora

Raiatea

Huahine

Moorea

Rimatara

Rurutu

Tubuai

Raivavae

