


CHAPTER 1

Out of Africa

EX AFRICA SEMPER ALIQUID NOVI

 UT OF AFRICA, THERE IS ALWAYS SOMETHING NEW, wrote historian Pliny the Elder more than 2,000 years ago. He was quite right. As early as 1984, just three years after the first description of the new disease, it was suspected that HIV, its recently discovered aetiological agent, originated in central Africa. This was mainly because the first studies in Africa, conducted in Zaire and Rwanda, showed that AIDS was common in Kinshasa and Kigali, where nearly 90% of sex workers were infected. These field studies were prompted by the observation that, of the first few hundred cases of AIDS diagnosed in Europe, about half occurred among patients coming from central Africa, mostly from Zaire. Notably, all were non-drug-using heterosexuals. Over the following years, the epidemiology of HIV-1 infection in Kinshasa would be described in great detail by a group of American, Belgian and Congolese researchers known as *Projet Sida*, based at Hôpital Mama Yemo (Mama Yemo was dictator Mobutu's mother, a former sex worker). *Projet Sida* came to an abrupt end in 1991, when the city of Kinshasa was completely looted by millions of people from disadvantaged socio-economic backgrounds. This was a forerunner of the regime's eventual fall or, at the very least, a sign of its profound decadence. Concerned by what might happen next, the American sponsors of *Projet Sida* hastily decided to shut it down. During the same period and until the 1994

genocide, similar epidemiological work was conducted in the hills of Kigali, 1,700 km east of Kinshasa.^{1,2}

In retrospect, this early vision of central Africa as the source of HIV-1 was rather naive. Researchers assumed that since this was at the time the region with the highest prevalence (proportion of the population that is infected) among groups representative of the general adult population, the virus must have originated there, giving it more time to spread. There were at least two problems with this assumption.

First, there was an obvious bias, as little information on HIV prevalence was available from other parts of the continent, especially East and Southern Africa. Belgian researchers, the most prominent being Peter Piot from the Institute of Tropical Medicine in Antwerp (who subsequently became the founding executive director of UNAIDS, the United Nations programme specifically dedicated to the control of HIV/AIDS), had naturally initiated HIV research in the former Belgian colonies where their institutions had maintained networks and contacts over the preceding decades. Much to their credit, Zaire and Rwanda were open to AIDS research from the start; this was not the case with other countries such as Burundi and some English-speaking countries of East Africa, where there was a strong temptation to keep AIDS under wraps, thinking that if it was ignored perhaps it would go away. This proved to be a very poor strategy.

Second, the relationship between HIV prevalence and the duration of the epidemic is not straightforward: it all depends on the incidence (proportion of previously uninfected individuals who acquire HIV each year). We now know that in Kinshasa the HIV incidence among the general adult population was probably never higher than 1% per year. However, in some countries of Southern Africa, the annual incidence reached the extraordinary level of 5% in the 1990s (i.e. one seronegative adult out of twenty was newly infected with HIV each year). A prevalence of 10% could reflect an annual incidence of 1% continuing for more than ten years, or an incidence of 5% over just a couple of years.

However, even if these assumptions about a central African origin of HIV/AIDS were naive, they eventually proved to be correct, showing that in the scientific domain intuition can sometimes be trusted; however, proving or disproving hypotheses can be quite complicated, as we shall see.

ARCHIVAL SAMPLES

Additional and quite strong support for a central African origin of HIV-1 came from the testing of archival samples of blood. In the mid- and late 1980s, after reliable serological tests had become available, to understand the dynamics of HIV in the recent past, researchers tried to locate collections of sera obtained earlier for other purposes and kept frozen. This was the simplest, most obvious way to address this question. Scientists tend to clean out their freezers once in a while to make room for new samples, or their samples are destroyed when they retire or move on to other positions; however, sometimes samples are forgotten for a long time or deliberately conserved by forward-thinkers.

In this way, it was possible to establish that in Kinshasa, among mothers attending a well-baby clinic in the Lemba district in the heart of the capital, HIV-1 prevalence was 0.25% in 1970 and 3.0% in 1980. In the remote catholic mission of Yambuku and surrounding communities of the province of Zaire, 0.8% of 659 samples collected in 1976 during an investigation of an epidemic of Ebola fever were found to be HIV-1 seropositive when tested ten years later. This proved that the virus had existed in this part of the world for some time, but not necessarily that it originated there. Testing of archived samples of serum from American gay men who participated in epidemiological studies of hepatitis B also retrospectively documented cases of HIV-1 in the late 1970s, and possibly even earlier for drug addicts.³⁻⁶

The Yambuku epidemic of Ebola fever, which had led to the discovery of its causal agent, had largely been 'iatrogenic' (healthcare-related; from the Greek *iatros*, physician). In this small rural hospital without any physicians, syringes and needles were scarce and constantly re-used after simply rinsing them in hot water. This fuelled transmission of the blood-borne Ebola virus between unfortunate patients attending the hospital for other reasons: uncomplicated malaria, a routine case of gonorrhoea, etc. The first medical teams sent to the area understood this when they noticed that the disease was especially common among pregnant women, who systematically received unnecessary injections of

drugs like calcium and vitamin B complex at the prenatal clinic. Like HIV, Ebola is a virus that is present in the blood and therefore transmissible by contact with it. To avoid waste and theft of the equipment, however, the Flemish nuns who ran this medical facility gave the nurses only five syringes and five needles each morning. This had to suffice not only for ambulatory patients and the prenatal clinics, but also for dozens of hospitalised patients. With this high rate of use, it was obviously impossible to sterilise injection equipment between patients. The medical staff, like their patients, firmly believed that injectable drugs were intrinsically superior to tablets. In the Congo, injections were considered to have an almost supernatural power, to transfer life force; the word *nkisi* means talisman, as well as medication. Three-quarters of the first 100 cases of Ebola in Yambuku were infected from injections received at the hospital, the others from taking care of sick relatives who themselves had been infected in the same facility. This was the exact opposite of *first, do no harm*.^{7,8}

The epidemic came to an end after the hospital was closed following the deaths of several nurses and nuns, infected by their patients, and thanks also to the local villagers. Most of them were illiterate and did not have even the most basic understanding of microbiology, yet they wisely decided to block all travel on the dirt roads leading to Yambuku. The aim was to prevent evil spirits from entering – the wrong diagnosis, but an excellent treatment! To understand the terror that reigned in Yambuku and also infected the heroic medical teams dispatched to try to combat an evil that was clearly very contagious but the cause of which was unknown, it is worth reading the romanticised version of events written by William Close, an American physician who lived in Zaire for many years.⁹

The tragedy of Yambuku showed that even noble intentions to provide underprivileged people with health care could have disastrous consequences if the risk of transmission of blood-borne viruses through contaminated injections was not appreciated. This unfortunate situation was not new or specific to the Yambuku hospital and had already had infinitely more dramatic consequences (although no one at the time knew this) than these few hundred deaths from Ebola fever. The researchers in Zaire at the time decided, while enjoying a beer after a hard day's work, to

call this new disease after a nearby river, rather than the Yambuku mission, to avoid further stigmatisation after everything it had gone through. A quick glance at the first available map resulted in the choice of Ebola.

The contrast between the two diseases is an excellent illustration of the genius of HIV and why it succeeded in causing a pandemic. People infected with the Ebola virus quickly fall ill and die within a matter of days. This causes a spectacular epidemic, which triggers a massive (and usually successful) international response to control it. People infected with HIV, on the other hand, can live and quietly pass on the virus for ten years or more, and it takes even longer before physicians, even the most discerning, recognise the emergence of this new disease. Since the incubation period is so long and variable (from two to fifteen years), instead of being grouped together over a short period of time like Ebola and thus attracting attention, the first cases of AIDS were randomly spread over a very long time span. In addition, the symptoms could not have been less specific: weight loss, diarrhoea, prolonged fever, all of which are consistent with many other tropical diseases.^{10,11}

The control of Ebola fever, with its short incubation period and dramatic manifestations, has so far been quite effective with two exceptions: the 2018–20 epidemic in the DRC (3,463 cases) and the 2013–16 crisis in West Africa (almost 29,000 cases). In the DRC, Ebola control was complicated by the chronic instability of the eastern part of this impoverished country, where violence and terror remain rampant, brought about by a plethora of armed movements, and where the people have long since lost confidence in all forms of authority. In West Africa, where the disease had previously been absent, Ebola was identified belatedly, after it had spread to many communities in three contiguous countries with porous borders. While Ebola outbreaks in isolated villages are relatively easy to contain, it is a different story when the virus reaches cities or densely populated rural areas. This is consistent with what has long been known regarding classical infectious diseases, such as measles, whooping cough and many others: the higher the population density, the faster they spread. Unfortunately, this also applies to HIV, and its urbanisation played a key role in its unrelenting propagation, as we will see later.

Elsewhere in Africa, no trace of HIV was found before the 1980s, which increasingly pointed to a central African origin of this 'new' virus. In West Africa, out of more than 6,000 samples obtained in Nigeria, Liberia, Ivory Coast, Togo, Senegal, Sierra Leone, Mali, Niger and Ghana in the 1960s and 1970s, not a single case of HIV-1 infection was elicited. A few cases of HIV-2 infection were documented, however.¹⁰

Meanwhile, in East and Southern Africa, in samples obtained from low-risk groups between 1959 and 1981, the scenario was the same and just as convincing: HIV was not found in Mozambique, Zimbabwe, Zambia, Uganda, Tanzania or northern Kenya, nor in mine workers in South Africa (who originated from Mozambique, Malawi, Lesotho, Botswana, Angola, Swaziland and South Africa itself). The earliest evidence of HIV in East Africa comes from Nairobi in 1980–1, where 1% of patients with venereal diseases and 5% of sex workers were HIV-1-infected.^{10,11}

Early cases of full-blown AIDS were also documented retrospectively. No conclusions can be drawn from isolated cases of apparently immunocompetent patients found to have been diagnosed many years ago with a condition now frequently associated with AIDS, such as *Pneumocystis* pneumonia, if this is not substantiated by a specimen positive for HIV in the patient or his/her spouse. This is because there are rare non-infectious diseases of the immune system which lead to very low counts of CD4 lymphocytes (the cells that are the main target for destruction by HIV), and subsequently to any of a long list of opportunistic infections. Short of an archived specimen positive for HIV, the clustering of cases, geographically or temporally, or within a couple, is more suggestive of AIDS but never conclusive.¹²

Valuable journalistic information about some documented early cases can be found in *The River* and *And the Band Played On*. The most fascinating is that of an unfortunate Norwegian family (father, mother and nine-year-old daughter), who all tragically died in 1976. AIDS was unknown at the time and doctors could not explain what had destroyed this family. However, the situation was so unusual that they deliberately saved some blood samples, hoping to find the solution to this mystery one day. This was an excellent idea: these specimens were tested twelve years later, and

not only were all three HIV-positive, they had all been infected with HIV-1 group O, a rare strain confined to central Africa. The child was born in 1967, which suggests that the mother was already HIV-infected by then. The father had been a sailor and, while still a teenager, he had visited a number of ports in Africa in the early 1960s, where he developed sexually transmitted diseases, presumably after contacts with prostitutes. He probably acquired HIV-1 group O in Nigeria or Cameroon, where his boat stopped for a few days in 1961–2. Life really does hang by a thread: a single ill-advised, unprotected act of sexual intercourse was sufficient to wipe out this family.^{13–16}

A Danish surgeon died of AIDS in 1977, after working at the Abumonbazi rural hospital in Zaire in 1972–5 and in Kinshasa in 1975–7, following an earlier stint in the same country around 1964. It was possible to confirm the infection years after her death in a Copenhagen hospital. An eight-year-old Zairean child, infected perinatally in 1974–5, died in Sweden in 1982, and AIDS was serologically proven later on. A very unfortunate Canadian pilot was involved in a plane crash in 1976 in northern Zaire. He had to wait twelve hours in the wreckage of his plane before being found by some villagers, who took him to Kisangani, where he had surgery and received a blood transfusion. He died of AIDS in 1980; his serum was later found to be HIV-1-positive (transfusion-acquired infection progresses rapidly to AIDS because of the huge quantity of viruses in the blood bag).^{17–19}

Jean Sonnet, formerly a physician at the university hospital in Kinshasa, reported seven cases of AIDS diagnosed retrospectively, four of them confirmed serologically, which had been acquired sexually in the DRC (or Burundi in one case) in the late 1960s or the 1970s, mostly among Belgian nationals. Then in 1979, some Brussels physicians tried unsuccessfully to treat a disease that was a mystery at the time, but which in hindsight was certainly AIDS. The patients were Zairians from the tiny political-financial elite rich enough to go to Europe in search of better medical care. The corrupt oligarchy that had destroyed the national health system had to leave the country if they had a serious illness, and no price was too high.^{20,21}

Thus, although sketchy, testing of archival samples suggested that HIV-1 was present in the 1960s and 1970s, albeit at a low prevalence, in

several locations in central Africa, but not in West, East or Southern Africa, at least on a scale that makes retrospective detection possible. The next step came from the documentation of the earliest case of HIV-1 infection in a sample obtained in the Belgian Congo around 1959, during a study on genetic diseases of red blood cells. Of 672 samples collected in Léopoldville and other locations and miraculously kept (probably forgotten) in a freezer, twenty-six years later one was found to contain antibodies against HIV-1, a male adult from Léopoldville. HIV genetic material was obtained from this sample, and analyses confirmed that this was indeed the oldest HIV-1 isolate ever documented. It was named ZR59.^{22,23}

It took twenty years for another ancient specimen containing HIV-1 to be located. Finding old tissue blocks collected between 1958 and 1960 and kept in the pathology department of the University of Kinshasa, scientists discovered HIV-1 in a lymph node biopsy obtained in 1960 from an adult woman. It was given the name DRC60. Twenty-six other specimens did not contain HIV. DRC60 and ZR59 differed by about 12%. It was calculated that they shared a common ancestor around 1908–21, as we will discuss later. Although the exact time of its introduction into human populations remains debated, there is no doubt that HIV-1 was present in Léopoldville by 1959–60. If HIV could be found decades later in specimens from these two individuals, the capital of the Congo must have had at the very least a few hundred infected people at that time. The search for the origins of HIV was already honing in on a specific place, namely the former Belgian Congo and Léopoldville in particular.²⁴

VIRAL DIVERSITY: THE PRINCIPLES

Now we will examine how the genetic diversity of HIV-1 in different parts of the world helped scientists not only trace back the origins of the virus but also reconstitute certain crucial steps in how it spread. Everything we call ‘genetic’ refers to DNA or RNA, the genetic code of all living organisms. First, we need to quickly review what ‘sequencing’ is all about. The aim here is not to understand these methods in detail, just the underlying principles, to get a general idea of how certain conclusions concerning the origins of HIV were reached. For readers

who haven't done biology since high school, the rest of this chapter and part of the next two may be a bit challenging, but in Chapter 4 we will return to mainly historical or socio-political concepts. Please be patient!

Sequencing is the identification in their proper order of the series of 'nucleotides' that constitute a gene. It has become the acid test in many fields of biology, because how can we go farther than the code of life? A gene consists of a series of nucleotides that encode instructions leading to the production of a specific protein. The genome is the set of all the genes in an organism, its complete code. There are four types of nucleotides: adenine (A), thymine (T), guanine (G) and cytosine (C). Nucleotides are like letters of the alphabet, and the genome of any living organism is a long list of these four letters. The total length of the genome is proportional to the complexity of the organism; for example, the HIV genome contains just under 10,000 nucleotides, the Ebola virus genome close to 19,000, the agent of COVID-19 nearly 30,000, and the human genome around three billion. Viruses (from the Latin word *virus* meaning poison) are living organisms that are disconcertingly simple, out of all proportion to their pathogenic power and potential for destruction. Because of this simplicity, to multiply they need to take over the cells of more complex living organisms; in the case of HIV, the main targets are human CD4 lymphocytes that involuntarily become virus production factories. As collateral damage, this process results in the destruction of these cells that are essential to the immune system.

When scientists compare viruses, the similarity between sequences is called 'homology', and non-similarity 'divergence'. If 90% of the sequences between two isolates are the same, they have 90% homology or 10% divergence. This degree of divergence is used to decide whether two isolates constitute subtypes of one viral species, or two distinct species. For instance, sequences of HIV-1 and HIV-2 differ by more than 50%. Homology and divergence are the scientific terms, but I will use 'similarity' and 'difference', which are easier to remember for the non-expert.

Based on such analyses, HIV-1 is now divided into four 'groups': group M (main), which is responsible for the pandemic and causes more than 99% of all HIV-1 infections in the world; group O (outlier), group

N (neither M nor O) and group P (called P simply because it followed O), which did not spread outside central Africa. There was no particular reason to use the letters M, N, O and P to designate these four groups of HIV-1: they were simply chosen by the people working on it.

HIV-1 group M is divided into nine ‘subtypes’ (this time, it was decided to follow alphabetical order): A, B, C, D, F, G, H, J and K (E and I are missing because they were found not to be real subtypes). HIV-1 often makes small mistakes when replicating (replication means generating two copies from one original). These mistakes or ‘mutations’ are usually substitutions, which means that two copies have one nucleotide out of 10,000 which differs from the original (for example, a letter A becomes a G or a C becomes a T). This phenomenon is compounded by the enormous viral production throughout the long natural history of the infection. Up to ten billion copies of the virus are produced every day, with a commensurate potential for errors in replication, especially considering that the average patient survives for a decade between infection and death due to AIDS. This means that approximately 3.8×10^{17} copies of the virus are generated each day worldwide, and over 10^{20} each year.

Over the long term, among the population infected with HIV living in a given location, the progressive accumulation of all these replication errors slowly but surely leads to HIV’s genetic diversity. ‘Diversity’ is the accepted term, but it could also be called variability or differentiation. ‘Evolution’ is the term for the entire process of progressive change in an organism’s genetic code as copying mistakes accumulate, while ‘diversity’ is the result of this process. It goes from 1 error out of 10,000 nucleotides to 2, 10, 100, 1,000, and more. Virologists arbitrarily agreed that, when between 17% and 35% of the sequences of nucleotides in the initial virus have undergone replication errors (= 17–35% difference), there is a new subtype of HIV. Below this threshold, the two viruses belong to the same subtype. Above this threshold, there could be a distinct group, even another kind of virus.²⁵

High-risk individuals (especially sex workers in Africa) can be infected with a first subtype, and later with a second subtype, which can recombine within some cells of this human host. Such recombinants have part of their genome derived from the first subtype and part from the second.

Mixed-race, as it were. Over 100 recombinants have been recognised to date and more are regularly added to the list. Their names correspond to the two subtypes involved in the recombination, their progenitors; for example, CRF02_AG is a recombination of subtypes A and G. Some recombinants have generated their own epidemic in specific countries or regions.²⁶

Subtypes differ with regard to their propensity to cause AIDS (individuals infected with subtypes C and D die faster) but also in their transmissibility. Shedding of HIV-1 in the genital tract of women infected with subtype C is higher than for other subtypes, which implies more effective female-to-male sexual transmission. With subtype C, the high degree of 'viraemia' (the quantity of virus in the blood) that characterises acute infections lasts longer (months rather than weeks) than with other subtypes, increasing its infectiousness. Subtype C spread like wildfire in Southern Africa, even though other subtypes had arrived at the same time. This higher transmissibility of subtype C explains its current worldwide preponderance.²⁷⁻²⁹

Some subtypes are associated in specific locations with particular modes of transmission. This represents a founder effect within specific risk groups: a subtype originally introduced in a group of intravenous drug users will continue to be transmitted preferentially to other addicts; another subtype originally introduced in sexual networks of homosexuals will spread preferentially to other gay men, and so on. A good example of this is South Africa, where subtype B is found in 96% of white homosexuals (after it was imported from the USA), while subtype C accounts for 99% of infections of black heterosexuals. In this rainbow nation there was an astonishing overlapping of two epidemics that evolved completely independently of each other, with subtype B being limited to a minority (homosexuals) within a minority (whites), while subtype C, which was more contagious, decimated heterosexual adult Bantu, who were infinitely more numerous, a majority within the majority.³⁰

To date, there is no evidence that some subtypes are intrinsically transmitted better by one route than another, but this is difficult to study. As the relative contribution of some modes of transmission varies over time, depending on the effectiveness of control efforts targeting

a specific risk group, the distribution of subtypes within a given country can also change.

Because of its frequent replication errors, HIV-1 evolves (in the biological sense of ‘gradual transformation’) one million times faster than animal or human DNA. This means that in just over a decade, HIV-1 will change as much as all the genetic changes and the ensuing diversity that accrued among the common predecessors of *Homo sapiens*, chimpanzees and gorillas over ten million years. The longer HIV-1 has been present somewhere, the more opportunities it will have had to undergo a series of mutations that will eventually allow it to evolve into different subtypes, and the more likely it is that recombinants will be created. Conversely, if we were to examine all of the HIV-1 isolates in a city where the very first case occurred only a year earlier, for example in a population of drug addicts, we would find little genetic variation. The viruses of all the individuals infected after this first case would still belong to the same original subtype, the founder.

Because HIV, like all living organisms, evolves in only one direction, from a single model of the virus (that of the very first patient – ‘patient zero’) to an increasingly complex differentiation into numerous subtypes and recombinants, we can reconstruct the sequence of its progress in a particular region or country by examining the local distribution of subtypes. The more different subtypes there are, the longer the virus has been ravaging the area.

Starting in the early 1990s, as new tools made it easier to examine nucleotide sequences from a large number of HIV-1 isolates obtained in various locations, an additional and very convincing argument emerged that supported a central African origin: the extreme genetic diversity of HIV-1 isolates from this part of the world.

VIRAL DIVERSITY: THE FACTS

Worldwide, subtype C now accounts for 47% of all HIV-1 infections, followed by subtypes B (12%), A (10%), recombinants CRF02_AG (8%) and CRF01_AE (5%), subtypes G (5%) and D (3%), while F, H, J and K have undergone limited transmission (all combined: 1% of

cases). The remaining 9% is made up of a long list of recombinants. However, the contribution of each subtype varies dramatically from region to region, which can be used to reconstruct the history of the dissemination of the virus.^{25,31}

In North America and Western Europe, 85% of HIV-1 infections correspond to subtype B, which is clearly the one that was originally introduced into these two continents, the founder strain. Subtypes other than B are usually found in migrants who acquired HIV-1 in their countries of origin and brought this unpleasant travelling companion with them. In contrast, in Eastern Europe and Central Asia, subtype A accounts for more than half of HIV-1 infections: this epidemic had a different origin than that in Western Europe, and it spread initially through needles rather than gay sex. During the iron curtain era, the barriers that isolated the two halves of Europe from each other were relatively hermetic and the movements of populations – and therefore viruses – were very limited.^{25,31}

In Latin America and the Caribbean, subtype B accounts for respectively 80% and 95% of strains, reflecting their proximity to the United States and its fifty million Latinos. This constantly growing mass of migrants, and the endless coming-and-going between the superpower and its backyard, facilitated the re-exportation of American strains of the virus. Cuba, on the other hand, stands out as the country with the lowest HIV prevalence in the Americas, but also the greatest diversity: about half of Cuban isolates are either non-B subtypes or recombinants.^{31–33}

Cuba deserves close scrutiny, because it is the exception that proves the rule: in this case, diversity does not mean that the virus was there earlier; rather, it reflects the acquisition of multiple subtypes of HIV-1 by some of the *internationalistas*, the soldiers sent to fight alongside the leftist regime during the civil war in Angola. Recent studies documented great diversity in HIV-1 isolates in Angola, where all non-B subtypes found in Cuba are present. In addition to the needless deaths of several thousand Cuban soldiers (the regime has never admitted exactly how many were lost), victims of the megalomania of their all-powerful, entrenched dictator, the Cuban adventure in Angola illustrates how political and military events, even ideologies, had a measurable impact

on the transcontinental dissemination of HIV. Epidemiology is also political.

During the interminable conflict that followed the Portuguese colony's accession to independence, Cuba's military intervention in Angola, financed by the USSR, lasted sixteen years, from 1975 to 1991. Nearly 400,000 Cubans spent time there, mostly soldiers, but also physicians, engineers and technicians. At the peak of their intervention in 1986, 35,000 Cuban troops were stationed in Angola. For the Havana regime, it was a way to whip-up revolutionary fervour in its young people and make them worthy heirs of Che Guevara and the *Lider Maximo*. Unfortunately, Angola became one of the most corrupt and capitalist countries on the continent: President dos Santos, in power from 1979 to 2017, amassed a fortune of thirty billion dollars, half of Cuba's gross domestic product, while the vast majority of the Angolan population was mired in poverty and did not see any of the windfall from petroleum.^{34–36}

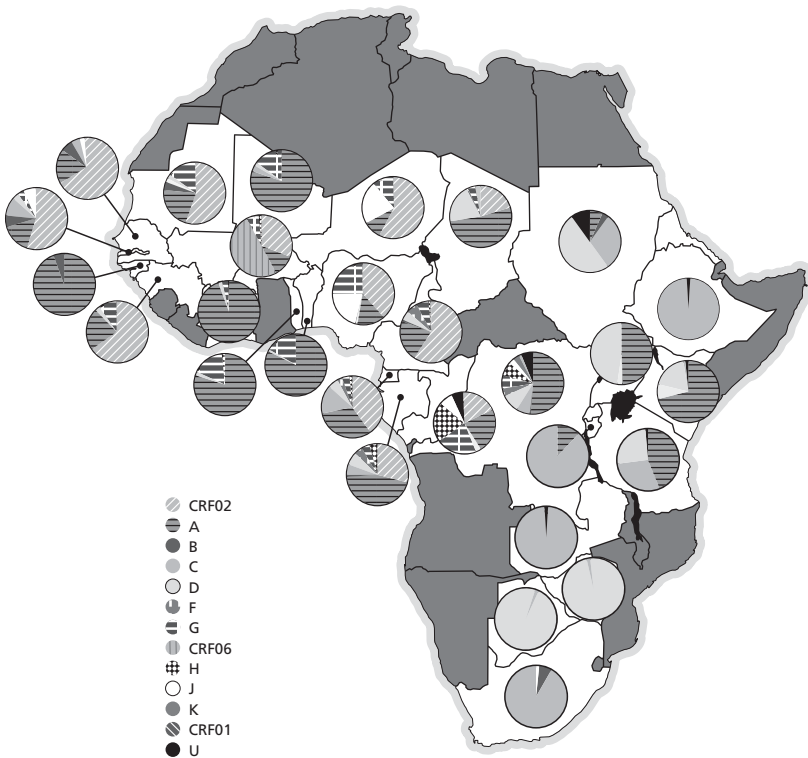
When the Cuban soldiers returned to their island nation, there were few opportunities to transmit the pathogen acquired during their time in Africa. To limit the spread of HIV back home, the whole Cuban population was screened in 1986–9. Starting in 1986, the *internationalistas* were also tested before they left Angola: one couldn't be too careful. HIV seropositives were quarantined for years in AIDS sanatoria ('*sidatoriums*') and brainwashed with preventive messages. The best-known sidatorium was called Los Cocos; it was located in a former leprosarium near a sanctuary dedicated to St Lazare – from leprosy to AIDS was just one small step!

According to the regime, the AIDS sanatoria were not places of death, far from it: living conditions were better than in the rest of society and seropositive patients, most of whom were asymptomatic, wanted to stay there as long as possible! Former inmates told a very different story: the sanatoria were prisons with wretched living conditions, especially during the 'special period', starting in 1990 when Soviet funds dried up and the economy nosedived as a result.

Cuba, a particularly long-lasting communist dictatorship and a veritable museum of utopias, was the only country that tried to control HIV like a classical infectious disease (identify those who are infected and isolate them to prevent the virus from spreading) rather

than making it a human rights issue. With some success, it must be said.

Now Africa. Map 1 summarises the distribution of HIV subtypes across Africa. Additional studies published since then pointed to small changes in the distribution of some subtypes, but without modifying the general pattern. In Southern Africa, subtype C corresponds to 99% of HIV-1 infections, the greatest homogeneity on Earth. This implies that the virus arrived relatively recently in this region that is now so afflicted by AIDS, a finding corroborated by epidemiological investigations. Recent introduction, followed by phenomenal transmission. Subtype C accounts for 99% of infections also in Ethiopia. In East Africa, the



Map 1 Genetic diversity of HIV-1 in sub-Saharan Africa. The circles show the distribution of HIV subtypes in various African countries (U stands for *unknown*). Adapted from Peeters.³⁷

picture is more nuanced and multicoloured, a result of the sequential introduction of several distinct strains. In West Africa, all the way from Nigeria to Senegal, CRF02_AG predominates, which means that this part of the continent was infected only after subtypes A and G had had the opportunity to recombine elsewhere, corroborating what had been revealed by the examination of old HIV-free blood samples before the 1980s.^{25,31,37}

As shown in Map 1, countries in central Africa display by far the widest diversity in HIV-1 subtypes. All subtypes and many recombinants have been found in this region. However, differences between countries help us to track past events. In the Central African Republic, there is a strong preponderance of subtype A. In Chad, there are many recombinants, suggesting that the virus disseminated there relatively recently.^{31,37,38}

In Cameroon, the CRF02_AG recombinant is by far the predominant strain, as in Nigeria to the north and Gabon and Equatorial Guinea to the south. This means that most of the HIV-1 transmission occurred relatively recently in these countries, since recombinants must be more recent than the two subtypes from which they are derived (just as children are younger than their parents). However, this observation does not in any way exclude the possibility that the very first case of human infection (by a virus which, by definition, was not a recombinant) occurred in Cameroon, the pathogen being then exported to a neighbouring country, where, finding circumstances more propitious for its propagation (and for the production of trillions of copies of the virus), a substantial part of its diversification took place. Subsequently, at least one recombinant strain was re-exported to Cameroon, where it caused much of its contemporaneous epidemic. Cross-border movements of the virus are two-way, in accordance with human migrations that fluctuate over time in tandem with economic or political factors.^{31,37,39}

In the DRC, HIV-1 isolates obtained in 1997 from Kinshasa, Bwamanda (a town in the province of Equateur) and Mbuji-Mayi (the capital of the diamond region of Kasai) were characterised. In Kinshasa, by descending order these were subtypes A (44%), D (13%), G (11%), H (10%), F (6%), K (3%), J (4%) and C (2%), while 7% could not be

properly identified. More recent studies had similar findings, except for a surge in subtype C, which is more transmissible. This broad distribution of subtypes proved similar to what was measured retrospectively in samples collected in Kinshasa in the mid-1980s by *Projet Sida*: HIV-1 diversity in Kinshasa thirty-five years ago was far more complex than among strains currently found in any other part of the world! Clearly, this sprawling city was at the heart of our story.^{31,40–42}

Only recently has HIV-1 diversity in Congo-Brazzaville been evaluated using a large number of isolates, obtained mainly in Brazzaville. A pattern similar to that in Kinshasa was found, namely a predominance of A and G but few CRF02_AG recombinants. Unlike its neighbour on the other side of the river, a few subtype B isolates were documented.^{31,43,44}

In the final analysis, the two Congos are the countries with by far the greatest diversity of HIV-1 subtypes. Thus, the oldest epidemic (the term ‘epidemic’ implies that a large number of people are infected, as opposed to an ‘outbreak’, which is more limited) did in fact start in the DRC and Congo-Brazzaville, which allowed numerous subtypes to develop and some to recombine. The virus arrived later in the Central African Republic and Chad, and even more recently in Gabon and Equatorial Guinea, after subtypes A and G had the opportunity to recombine. The recombination that created CRF02_AG took place in the DRC, from where this strain was exported to Cameroon and its neighbours around 1973, before relentlessly following its own path.⁴⁵

The genetic diversity of HIV-1 in a given location is influenced not only by how long it has been there, but also by how efficiently it propagated. An HIV-1 strain producing only one case of secondary infection every five years would present fewer variations after fifty years than the same strain that, introduced into another environment, generated a secondary case every three months, with each secondary case in turn producing a tertiary case every three months, and so on. The more people are infected, the more copies of the virus are produced each day, which increases the number of mutations and differentiation into subtypes.

What the great genetic diversity of HIV-1 in Kinshasa and Brazzaville means is that for the first time, in this large urban area, the strategic hub

of the huge Congo River basin, the virus found conditions conducive to its dissemination on a scale that enabled it to flourish and differentiate, after an initial phase of stagnation or very slow multiplication that could have occurred elsewhere, in any of the countries inhabited by its simian source.