Southeast Asia is one of the world's highest-risk EID hotspots, and the origin of the SARS pandemic, repeated outbreaks of novel influenza strains and the spillover of dangerous viral pathogens such as Nipah virus. It is a wildlife 'megadiversity' region, where a rapidly expanding human population is increasing contact with wildlife, and increasing the risk of zoonotic disease outbreaks. The overarching goal of this proposal is to launch the Emerging Infectious Diseases - South East Asia Research Collaboration Hub (EID-SEARCH) to analyze the diversity of key viral pathogens in wildlife, the frequency and causes of their spillover, and to identify viral etiologies of undiagnosed 'cryptic' outbreaks in people. EID-SEARCH includes leaders in the field of emerging viral pathogens at key US institutions, and in Thailand, Singapore, and the 3 major Malaysian administrative regions, whose collaborative networks span >50 clinics, laboratories, and research institutes across almost all SE Asian countries. This hub, and the network, will act as an early warning system for outbreaks - a way to exchange information, reagents, samples and technology, and a collaborative power-house for translational research. The long-term collaboration among the key personnel, and multidisciplinary skillsets from epidemiology, clinical management, lab analysis, through wildlife biology and data analysis **will act as significant assets when deployed to help counter outbreaks in the region.** The research goals of this EIDRC follow three specific aims:

Specific Aim 1: Identify, characterize and rank spillover risk of high zoonotic potential viruses from wildlife. We will: 1) analyze some of the tens of thousands of archived wildlife samples at our disposal, conduct geographically- and taxonomically-targeted field surveillance in wild mammals (bats, rodents, primates), and use serological & PCR assays to identify known high-profile zoonotic pathogens, or close relatives with potential to infect people; 2) biologically characterize novel viruses that our analyses suggest have high spillover and pandemic potential; and 3) conduct *in vitro* receptor binding assays and cell culture experiments, and *in vivo* animal model infections using humanized mice and the collaborative cross mouse to assess their potential to infect people and cause disease.

Specific Aim 2: Identify evidence and analyze risk factors for viral spillover in high-risk communities using novel serological assays. Assessing the spillover of rare or novel zoonotic agents will require targeted surveillance of high-risk communities and approaches that can deal with the low statistical probability of identifying rare events. To achieve this, we will 1) conduct targeted cross-sectional serological surveys of human communities with extremely high geographic and cultural, occupational and behavioral exposure to wildlife-origin viruses; 2) design and deploy novel serological assays to identify baseline spillover of known or novel CoVs, PMVs and FVs in these populations; and 3) analyze and test hypotheses on the occupational, cultural and other risk factors for spillover (e.g. hunting wildlife).

Specific Aim 3: Identify and characterize viral etiology of 'cryptic' outbreaks in clinical cohorts. Our prior work provides substantial evidence of spillover leading to undiagnosed illness in people in the region. To test if these represent 'cryptic' outbreaks of novel viruses, we will conduct syndromic surveillance at regional clinics for the communities sampled in SA2. We will: 1) enroll and collect biological samples, and detailed survey data on risk factors, from patients presenting with influenza-like illness, severe respiratory illness, encephalitis, and other specific symptoms; 2) conduct molecular and follow-up serological diagnostic assays to test causal links between their syndromes and known and novel viral agents identified in SA1. Where viruses are identified, we will attempt to isolate and characterize them, then use the survey data, ecological and phylogenetic analyses to identify likely reservoir hosts/spillover pathways and inform intervention programs.

This research will advance our understanding of the risk of novel viral emergence in a uniquely important region. **It will strengthen in-country research capacity** by linking local infectious disease scientists **with an international collaborative network that has proven capacity to conduct this work and produce significant findings.** These include: testing of tens of thousands of samples from wildlife, humans and livestock in the region; discovery of hundreds of novel viruses from zoonotic viral families in wildlife; outbreak investigations in rural communities across SE Asia; discovery of the bat-origin of SARSr-CoVs; discovery of a novel bat-origin SADS-CoV killing >25,000 pigs in S. China; and development of novel serological and molecular assays for high-impact viruses, and state-of-the-art *in vitro* and *in vivo* assays to characterize viral pathogenic potential. This body of collaborative research provides proof-of-concept that EID-SEARCH has the background, collaborative network, experience, and skillset to act as a **unique early warning system for novel EIDs of any etiology threatening to emerge in this hottest of the EID hotspots.**

II. Research Strategy:

1. Significance: Southeast Asia is a critical hotspot for emerging diseases due to its high biodiversity of wildlife and their viruses, and the presence of key ecological and socioeconomic drivers of emergence **(Fig. 1)** (*1*). These include dense human populations living in networks of rural and urban communities, with strong cultural and occupational connection to wildlife and livestock farming and trading, and intimate connection to global travel networks (*2*). The region is undergoing rapid environmental and demographic change, both of which increase the risk of disease emergence and spread. Pathogens that emerge in this region often spread and sometimes enter the USA (e.g. prior influenza pandemics, SARS) and threaten global health security.
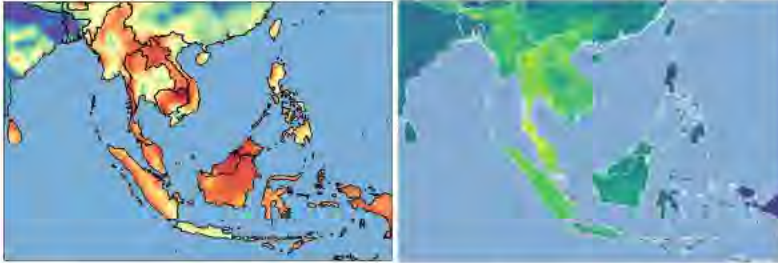

Fig. 1: Southeast Asia, and in particular, parts of Thailand and Malaysia are hotspots of risk for disease emergence **(left)**, and predicted high diversity of 'missing' or as-yet undiscovered viruses, yellow = highest diversity **(left)**. From (*1-3*).

Not surprisingly novel viruses, including near-neighbors of known agents, have emerged in the region leading to often unusual clinical presentations **(Table 1)**.This adds to a significant background burden of repeated Dengue virus outbreaks in the region (*20*), and over 30 known *Flavivirus* species circulating throughout S. and SE Asia (*21*). The events in **Table 1** are dominated by three viral families: coronaviruses (CoVs), paramyxoviruses (PMVs - particularly henipaviruses) and filoviruses (FVs), which are initial examples of our team's research focus and capabilities. These viral groups have led to globally important emerging zoonoses (*4, 5, 22-31*), causing tens of thousands of deaths, and costing billions of dollars (*32-34*). Furthermore, near-neighbors of known pathogens in these families have been reported throughout the region, including by members of our consortium: henipaviruses in bats in Thailand, Cambodia, Philippines, Laos, Indonesia, China, Malaysia, Bangladesh and India (*35-41*); a novel henipavirus, Mòjiāng virus, in wild rats in Yunnan (*14*); serological evidence of FVs in bats in Bangladesh (*42*) and Singapore (*43*); Nipah- and Ebola-like viruses in *Hipposideros*, *Cynopterus* and *Rhinolophus* species in Malaysia (Hughes *et al.*, in prep.); evidence of novel FVs in bats in China (*44-46*), including Měnglà virus

Viral agent	Site, date	Impact	Novelty of event	Ref.
Nipah virus	Malaysia, Singapore 1998-9	~246 human cases, ~40% fatal	2nd emergence of a zoonotic henipavirus, 1st large outbreak	(4-6)
Melaka & Kampar virus	Malaysia 2006	SARI in family group, individual	1st disease due to bat-origin reoviruses, also Singapore, Vietnam etc.	(7-10)
Reston filovirus	Philippines 2008	Seropositive people killed pigs,	No prior FVs in pigs	(11)
Thrombocytopenia Syndrome virus	E. Asia 2009	100s of deaths in people	Novel tick-borne zoonosis with large caseload	(12)
Reston filovirus	Shanghai 2011	PCR positive pigs	Further evidence of pig RESTV infection	(13)
Mòjiāng virus	Yunnan 2012	Death of 3 mineworkers	1st evidence of rodent origin henipavirus in people	(14)
Nipah-like virus	Philippines 2015	Killed horses	No prior horse infection for NiV, known for HeV	(15)
SARSr-CoV & HKU10-CoV	S. China 2015	Seropositive people	1st evidence human infection HKU10 & SARSr-CoV	(16)
SADS-CoV	China 2017	>25,000 pig deaths seronegative people	Novel emergence of bat-origin CoV	(17)
Nipah virus	Kerala 2018, 2019	Killed 17/19 people 2018, 1 infected 2019	1st outbreak of NiV outside Bangladesh, W. Bengal focus	(18, 19)

that appears capable of infecting human cells (*47*); novel PMVs in bats and rodents consumed as bushmeat in Vietnam (*48*); a lineage C β-CoV in bats in China using the same cell receptor as MERS-CoV to infect human cells *in vitro* (*49*); MERSr-CoVs in bat guano harvested as fertilizer in Thailand (*50*), and directly in bats (Wacharapluesadee *et al.*, in prep.); 172

Table 1: Recent emergence events in SE Asia indicating potential for novel pathways of emergence, or unusual presentations for known or related viruses.

1

novel β-CoVs (52 novel SARSr-CoVs) and a new β-CoV clade ("lineage E") in bats in S. China that occur throughout the region (24, 49, 51, 52); 9 novel wildlife-origin CoVs and 27 PMVs in Thailand, and 9 novel CoVs in Malaysia reported by us from USAID-PREDICT funding (51, 53). These discoveries underscore the clear and present danger of zoonotic events in the region. The wide diversity of potentially pathogenic viral strains also has significant potential use in broadly active vaccine, immunotherapeutic and drug development (49).

Most high-impact zoonotic viruses originate in wildlife reservoirs, sometimes spilling over first into livestock 'amplifier' hosts, or directly into localized human populations with high levels of animal contact (Fig. 2). Efforts to prevent emerging zoonoses have targeted these high-risk populations in regions prone to disease emergence (31, 54). However, these regions tend to be in the developing world, where nations often lack the resources to investigate novel zoonotic events, or wildlife reservoirs of zoonoses, and prioritize more immediate public health concerns such as Dengue virus or NTDs (20). Zoonotic disease surveillance and control is also hampered by inadequate information on basic disease ecology (e.g. range of wildlife reservoirs, risk factors for spillover) and pathogen biology (viral diversity, pathogenesis), and by logistical challenges sharing novel diagnostics, animal models, samples and reagents at sites where outbreaks begin (55). Countermeasure and vaccine development is challenged by the small number of isolates available, and the presence of closely-related strains, some with evidence of transmission to people but with unknown pathogenicity. For example, using NIH and other funding, we found serological evidence of exposure to African henipaviruses in 3-4% of a sample of people in Cameroon, with a significant risk factor being butchering of bats; and serological responses to bat SARSr-CoVs in 6/209 (2.87%) and bat HKU10-CoV in 2/412 (0.5%) people living close to a bat cave in Yunnan, all novel viruses with unknown clinical impact (16, 56). We discovered two novel bat-borne reoviruses, Melaka and Pulau viruses, and showed that 14/109 (13%) of people living in close proximity to bat roosts on Tioman Island, Malaysia were exposed to these viruses (7) as well as 12/856 (1.4%) of a random sample screened in Singapore (8). Previous studies in the Philippines found famers and butchers who had contact with sick pigs tested positive for Reston ebolavirus antibodies but had not experienced illness (11), and it was our team that discovered Philippine bats as natural reservoir hosts for Reston virus (57). In preliminary screening in Malaysia with funding from DTRA we found serological evidence of exposure to henipaviruses (Hendra, Nipah and Cedar) in 14 bats and 6 human samples, and Ebola (Zaire and Sudan)-related viruses in 11 bats, 2 non-human primates (NHP) and 3 human samples (Hughes, in prep.).

PANDEMIC EMERGENCE
International Travel and Trade

LOCALIZED EMERGENCE
Expansion of the Wildlife-Human Interface

WILDLIFE RESERVOIR
Land use change

Encroachment into Wildlife Habitat

Even when spillover leads to outbreaks of illness and mortality, they are often unreported, undiagnosed or misdiagnosed. We initiated targeted syndromic surveillance of encephalitis patients in Bangladesh clinics and showed that Nipah virus (NiV) causes repeated annual outbreaks with an overall mortality rate of ~70% (58). NiV has been reported in people in West Bengal, India, which borders Bangladesh (28), in bats in Central North India (59), and now in people in Kerala, South India in 2018 and 2019, raising the specter of future spillover at other sites in the region (18, 60).

Fig. 2: Our Early Warning Approach: Most zoonotic viruses emerge from wildlife hosts, sometimes via livestock (lower panel), to cause small scale outbreaks (green) or larger chains (red) of human-to-human transmission (middle). In some cases, these spread more widely via air travel (upper). Our EIDRC proposal targets each stage: SA1 examines the diversity and spillover potential of viral threats in targeted SE Asian wildlife; SA2 seeks evidence of their spillover into focused high-risk human populations; SA3 identifies their capacity to cause illness and to spread more widely. Figure from (54).

Under a current NIAID R01, we have shown that we can use analytical approaches to target surveillance of wildlife and people and identify likely sites of spillover. We used phylogenetic analysis of bat hosts and their CoVs to identify a key cave complex in SW Yunnan Province that harbors all the genetic elements of SARS-CoV (61), as well as a series of SARSr-CoVs that can infect human cells in vitro, cause SARS-like clinical signs in humanized mice, and evade vaccine treatment and therapeutics (24, 49, 51, 52, 62-64). We conducted risk factor surveys and biological sampling of human populations living close to this cave and found serological evidence of spillover in people highly exposed to wildlife (16). This work provides proof-of-

2

concept for an early warning approach that we will expand in this EIDRC to target other viral groups in one of the world's most high-risk EID hotspots.

The overall premise of this EIDRC proposal is that there is substantial evidence that: 1) diverse CoVs, henipaviruses and FVs related to known human pathogens are circulating in wildlife reservoirs in Southeast Asia; **2)** these viruses likely spillover regularly to people, are often unreported or misdiagnosed, and their clinical manifestations and potential to cause pandemics are unknown and underestimated; and **3)** our strategy for targeted surveillance and detection of spillover and illness in at-risk human populations can be used as an 'early warning system' to conduct public health interventions and disrupt disease emergence. To scale-up this approach, we propose to launch **EID-SEARCH** (the **E**merging **I**nfectious **D**iseases - **S**outh **E**ast **A**sia **R**esearch **C**ollaboration **H**ub), to better understand, and respond to, the risk of zoonotic viral emergence in Southeast Asia. We will initially focus on the diversity of known and novel CoVs, henipaviruses, and FVs in wildlife reservoirs. We will use this work to build our center's interactive network throughout the region, bridge barriers, provide early focal point for basic and translational studies, develop sample handling and storage routines, communication networks, cohort management, infrastructure development, and cross-training in surveillance and novel diagnostic platforms. We will develop and transfer technology to test these viruses' capacity to infect human cells and mouse models, develop new specific and sensitive serological and molecular diagnostic tools, conduct surveillance of human communities with high risk of exposure to wildlife, and clinical cohorts to identify evidence of spillover of viruses causing previously 'cryptic' clinical syndromes in people. We will test the capacity of novel viruses to infect human cells and/or mouse models, and thus asses potential for spillover and spread which will feed into our risk models. Our group also has extensive experience in other viruses, including vector-borne flaviviruses, and alphaviruses, and will use this experience to provide robust long-term surveillance, diagnostics, and basic science opportunities. While the current proposal focuses its basic and applied research agendas on this group of viral EID, we note that the EID-SEARCH is poised to rapidly respond to any outbreak. **Thus, FVs, henipaviruses and CoVs are examples of sentinel surveillance systems in place that illustrate EID-SEARCH's capacity to identify and initiate a rapid response targeting any novel microbial pathogen that emerges in this region.**

2. Innovation: Previous work by our consortium has developed strategies to target surveillance in wildlife and people to better anticipate and identify early spillover events and pre-empt outbreaks of emerging viruses. In the EID-SEARCH we scale up this approach in a central research "hub" made up of leading laboratories, clinics and surveillance sites in three critically high-risk EID hotspot countries in SE Asia. EID-SEARCH's reach is expanded through its regional network of collaborators that covers nearly all of the greater SE Asia region. The innovation of the EID-SEARCH is in: **1)** Its multidisciplinary approach that combines modeling to target geographic and taxonomic sampling targets with on-the-ground zoonotic disease surveillance in human and wildlife populations; **2)** the biological characterization approach that identifies how likely viruses are to be able to infect people, and enables evaluation of existing countermeasure technologies; an approach that is grounded in our successful work on SARSr-CoVs; and **3)** the strategic sampling of both high-risk communities with extensive animal contact at the front-lines of viral spillover *and* syndromic clinical cohorts that present with signs of novel viral emergence, to identify risk of viral spillover in an early warning system approach **(Fig. 2)**. **In Aim 1, we will target viruses in archived and newly collected wildlife samples, and characterize their risk of spillover to people**. We will use spatial and phylogeographic analyses to identify the geography and species of wildlife to test samples from, and to further characterize viruses we have recently discovered. We will use *in silico* methods (novel ecological-phylogenetic risk ranking and mapping of receptor binding domains), cell culture, and animal models to assess their potential for spillover into high-risk human populations. **In Aim 2, we will conduct targeted cross-sectional serological surveys of human communities with high levels of animal contact to find evidence of viral spillover, and identify occupational and other risk factors for zoonotic virus transmission.** We will use serological tests targeting viruses identified in Aim 1 to identify the baseline spillover of known or novel CoVs, henipaviruses and FVs in these populations. Serological findings will be analyzed together with subject metadata (including a simple animal-contact survey, location data, viral sequence/isolates etc.) to identify zoonotic risk factors and routes of exposure from specific wildlife or vector species. **In Aim 3, we will use syndromic surveillance of clinical human cohorts to identify evidence of viral etiology for otherwise 'cryptic' outbreaks.** We will enroll and

3

collect biological samples from patients in local clinics and hospitals who present with syndromes previously linked to known viral agents and live within communities at high risk for disease emergence. We will conduct molecular and follow-up serological diagnostic assays to identify likely viral etiology of syndromes that may represent otherwise cryptic events. Where viruses are identified, we will attempt to isolate or molecularly re-derive them, and use survey data, ecological and phylogenetic analysis to identify likely reservoir hosts and inform intervention programs. Our US partner BSL-4 laboratory, NEIDL, will attempt isolation and characterization of viruses requiring BSL-4 of containment (e.g. novel filoviruses, close relatives of Hev/NiV).

3. Approach: 3.1. Research team: The EID-SEARCH builds on long-term collaboration among world leaders in emerging virus research with proven experience collaborating internationally on field, lab and human surveillance research **(Fig. 3)**. Over the past two decades, our consortium partners have conducted high profile research on EIDs within the region and globally, including identifying the wildlife reservoirs for NiV and Hendra virus (HeV), MERS- and SARS-CoVs (*24, 27, 59, 61, 65-67*), discovering SADS-CoV (*17*), and developing an array of serological, molecular, *in vitro* and *in vivo* approaches to characterizing high-risk CoVs, henipaviruses, and FVs (*62-64, 68-81*). Our team has substantial experience conducting human surveillance in the community and during outbreaks **(Sections 2.2.a., 2.2.b., 2.2.c, 3.2.a, 3.2.b)**, including: USAID-PREDICT, and DTRA-funded viral surveillance of indigenous communities in Peninsular Malaysia and syndromic surveillance of hospital patients in Sabah (Co-I Hughes); collection of >2,000 serum samples from healthy blood donors and sanitation workers, febrile and influenza like illness (ILI), and diarrheal stool samples from children under 5 years of age across Borneo Malaysia (Co-I Kamruddin); collaboration on the MONKEYBAR project with the

London School of Hygiene and Tropical Medicine to analyze risk and prevalence of the zoonotic malaria *Plasmodium knowlesi* through case control and cross-sectional studies in Sabah villagers (n=~800, 10,000 resp.) and clinical observational studies to determine the etiology of central nervous system infections and acute undifferentiated febrile illness in Sabah. (Co-Is William, Rajahram, Yeo); and community-based surveillance of hundreds of bat-exposed villagers in Thailand (Co-Is Wacharapluesedee, Hemachudha). As a proof-of-concept of how our team can rapidly scale up resources for diagnostics and surveillance in the face of a new EID, in 2016, we discovered a novel bat-origin HKU-2 CoV (*82, 83*) killing >20,000 pigs in S. China, designed PCR and serological assays, surveyed pig farmers for evidence of exposure, and conducted experimental infections, pathogenesis and molecular studies, **all in the span of 3 months** (*17, 84*).

Fig. 3: EID-SEARCH scope, core institutions, and roles.

The administration of this center **(Section 4.1.)** will be centralized at EcoHealth Alliance (EHA) led by PI Daszak, who has >20 years' experience managing international research collaborations on emerging zoonoses, and has collaborated with all partners in the EIDRC consortium for 5-20 yrs on NIAID- and USAID-funded research. He will oversee all activities in this EIDRC, supported by Deputy lead, Dr. Olival – who has managed human and wildlife disease surveillance projects in SE Asia for >10 years, and a **Core Executive Committee (Section 4.1.a)**. Co-Is Olival and Zambrana are leaders infectious disease analytics, and head the modeling and analytics work for USAID-PREDICT. Co-Is Wang, Baric, Broder, Anderson, and Laing have developed a unique array of *in vitro* cell culture, serological and molecular reagents to identify and characterize emerging viruses in people and animals (henipaviruses and other PMVs, CoVs, FVs and many others). Co-Is Wacharapluesedee, Hemachudha, and Hughes and collaborators have coordinated clinical and community surveillance of people, and of wildlife and livestock within Thailand and Malaysia for the past 15 years, in collaboration with Ministries of Health, Agriculture and Environment. Finally, this consortium has extensive external funding from NIH, DoD, USAID, DHS and other USG agencies supporting laboratory, analytical and field studies directly related to the current proposal, and will bring substantial leverage to our EIDRC.

4

3.2. Geographical focus: The three core (hub) countries for this proposed EIDRC are Thailand, Malaysia and Singapore, three contiguous EID hotspot countries that stretch through one of the most significant foci of EID origins globally. Our proposed work includes sampling wildlife and people in Thailand, Peninsular Malaysia, Sarawak and Sabah, and thus contains overlapping wildlife biogeography with India, Bangladesh, Myanmar, SW China, Laos, Cambodia, Vietnam, Philippines, and Indonesia. It encompasses diverse cultural and socioeconomic backgrounds which have correspondingly diverse and intensive interactions with wildlife reservoirs and the global travel network. We have already begun sampling of wildlife hunters, livestock workers, wildlife farmers and market workers, rural villagers living close to high risk wildlife species, people otherwise highly exposed to animals in these countries, and clinical cohorts, as part of current/prior NIH, DTRA, and USAID funding. Our core member's extensive network of collaborators in clinics, research institutes and public health laboratories in almost every other Southeast Asian country positions the EID-SEARCH within a much larger catchment area for emerging zoonoses and will allow us to rapidly scale up to include additional sites in response to an outbreak or shift in research priorities. We will conduct regular meetings with key research, public health and community leaders, including >50 leading laboratories/institutions in 10 countries **(Section 4.2)** to maintain these collaborative relationships with the core members of our consortium (**Fig. 4**).

EID-SEARCH Core Countries
Collaborating Partners

We will use these, and our annual meetings, to share information on novel research and diagnostic approaches, pathogens that are of key pandemic potential, regions or populations at high risk of spillover, and information from the greater network on likely outbreaks of novel disease. This platform will coordinate sample sharing and diagnostic platforms and help build a rapid response to outbreaks across the region.

Fig.4 : Map of Southeast Asia indicating three hub countries for this proposed EIDRC (**Red**: Thailand, Singapore, Malaysia) and countries in which Key Personnel have existing collaborating partners via other funded work (**Green**), indicating that EID-SEARCH provides access to key collaborators and sites throughout the region.

3.3. Capacity for linkage with other NIAID studies, and scaling up to respond to outbreaks: EID-SEARCH US-based partners include senior scientists with significant experience on NIAID-funded projects, extensive collaborations with other NIAID-funded scientists, and previous experience working with NIAID leadership through workshops, review of programs (e.g. PI Daszak's role in NIAID CEIRS external review), and work on study sections. These contacts ensure that the EID-SEARCH senior leadership will be ready and able to rapidly set up data, sample and reagent/assay sharing protocols with NIAID researchers, the other EIDRCs, the EIDRC CC and other NIAID research centers. EID-SEARCH in-country partners include senior medical doctors and infectious disease specialists at national and regional hospitals in each country and administrative state in this proposal. On news of potential outbreaks, **EID-SEARCH will rapidly mobilize its core staff and their extensive collaborative network across almost all SE Asian countries (Fig. 4) (Sections 4.1.a, 4.2. 4.3).** The core analytical approach can be used to rapidly identify sampling sites and targets to harvest vectors, vector borne arboviruses, birds and mammalian species that harbor most other viral threats. Through existing projects such as the USAID PREDICT Project and DTRA-funded biosurveillance projects, EHA is currently partnered with hospitals and researchers across South and Western Asia, West and Central Africa, and Latin America. Our partners already have capacity to conduct human and wildlife sampling and testing for novel viral agents, so expanding in geography or scale is readily achievable given sufficient resources.

Aim 1: Identify, characterize and rank spillover risk of high zoonotic potential viruses from wildlife.

1.1 Rationale/Innovation: The vast majority of emerging zoonotic pathogens are linked directly or indirectly to wildlife host, particularly from important mammalian groups like bats, rodents and NHPs (3). Our previous work on wildlife origin viruses demonstrates that zoonotic viral emergence tends to occur in specific regions with high diversity of wildlife that harbor viruses of zoonotic potential, and human populations that have close contact to these wildlife **(Fig. 1)** (2, 3). In Aim 1 **(see Fig. 9 for overview)**, we will strategically conduct EID

5

surveillance by applying risk analytics to prioritize the regions, wildlife species, and human populations that have a particularly high risk of viral spillover. We will target high-risk zoonotic host reservoirs (bats, rodents and NHPs) in these regions and apply novel predictive models to estimate viral diversity and host range to identify wildlife species that have been ignored or under-sampled by previous zoonotic disease surveillance efforts. We will strategically collect specimens from these species and screen them, together with some of the tens of thousands of recently collected samples we have archived, for known and novel agents. We will begin screening using conserved PCR assays to identify known and novel viruses in high risk RNA viral families, *Coronaviridae*, *Paramyxoviridae*, and *Filoviridae*. We will expand to other groups of pathogens, including arboviruses, pending available resources and priorities established with the EIDRC coordinating center (EIDRC-CC) and other regional EIDRCs. Using risk ranking algorithms and a series of ecological and virological risk factors, we will triage the wildlife viruses we discover for further genomic characterization and viral isolation. On this subset of viruses, we will then use *in vitro* and *in vivo* approaches (viral rescue, cell lines, gene edited humanized transgenic mice and the collaborative cross genetic reference mouse population) previously developed and widely used by our team to predict capacity of novel viruses to infect people and spillover (*63, 85-88*). These high-risk viruses and their close relatives will be targets for human community serosurveys and clinical sampling in Aims 2 and 3, respectively.

1.2 Preliminary data: 1.2.a. Geographic targeting: EHA has led the field in analyzing where the highest risk of zoonotic spillover is across the globe, identifying SE Asia as a major EID hotspot (*1, 2*). These hotspots are sites where high wildlife biodiversity, human population density and rapid ecologic change collide **(Fig 1)**. Using host and ICTV viral data for all known mammalian viruses, and generalized additive models to correct for sampling bias, we are able to predict the relative number of yet-to-be-described viruses that a species harbors (*3*). This demonstrates substantial undiscovered viral diversity in mammals across Southeast Asia, and particularly in regions of Thailand, Singapore and Malaysia **(Fig. 1)**. Separately, we found that factors driving spillover (host diversity, climate) differ from those driving spread in human populations (population density, bushmeat hunting, livestock production) (*89*). We therefore analyze capacity for viral spread as a separate risk factor, using our unique demographic and air travel and flight data modeling software (*90, 91*). In the current proposed work, we will use all the above approaches to target wildlife sampling (archival and new field collections) in Aim 1, to inform sites for human sampling in Aim 2, and to assess risk of spread in Aim 3.

1.2.b. Reservoir host species targeting: Our analysis of all known mammalian host-viral associations demonstrates that bats, rodents and NHPs represent the three most important reservoirs for zoonoses, responsible for the highest proportion of known and predicted viral diversity of zoonotic potential (*3*). Thus, the bulk of our proposed wildlife sampling and testing will consist of these three taxa with a focus on bats, because this group is the known or hypothesized reservoir for the majority of high-impact zoonotic CoVs, PMVs and FVs (*50, 92-94*). Bats have the highest proportion of predicted zoonotic viruses in any mammalian group (*2*), and are exceedingly diverse in Southeast Asia (~500 species). To further prioritize taxonomic sampling targets within these diverse geographic hotspots, we used a novel Bayesian phylogeographic algorithm to identify host genera and species that are the most important centers of evolutionary diversity for a given viral group, e.g. bat genera for β-CoVs **(Fig. 5)** (*51*). This approach will be extended to other mammal taxa and viral families of interest, allowing us to identify the reservoir hosts of specific zoonotic viral groups, where viral diversity is likely highest.

Fig 5: Preliminary analysis of sequence data collected under prior NIH funding, identifying SE Asian bat genera with the highest β-CoV evolutionary diversity. We will apply this Bayesian discrete phylogeographic model to identify taxonomic and geographic centers of viral diversification and sharing for relevant PMV, CoV, FV and other virus clades. Line thickness proportional to probability of virus sharing between two genera. Inter-family switches in red.

Finally, we have designed a strategy to estimate sampling gaps, using a viral 'mark-recapture' approach we previously published (*95, 96*) **(Fig. 6)**. Using prior data for bat, rodent and NHP viral discovery vs. sampling rates, we can identify the likely total viral diversity for specific host groups in a region, and use this to estimate sample sizes to discovering the majority of unknown viruses within a specific family or genus. **In the current**

6

proposal, we will use these analyses to estimate geographic and species-specific sampling targets **to more effectively identify new strains to support experimental infection studies and risk assessment.**
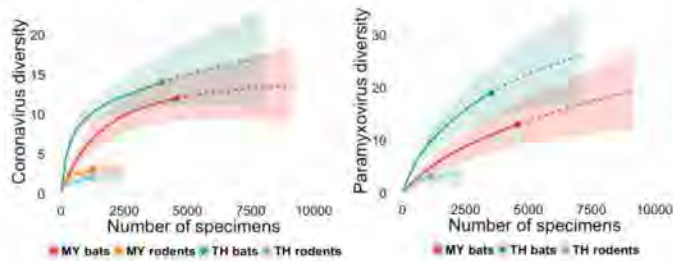

Fig. 6: Estimated CoV **(left)** and PMV **(right)** diversity in bats and rodents from Thailand and Malaysia, using data from PCR screening and RdRp sequences from >10,000 specimens in bats and 4,500 in rodents. Bats have 4X more viral species than rodents, controlling for sampling effort. We estimate that additional collection of 5k-9k bat specimens and testing of our archived bat and rodent specimens alone will identify >80% of remaining CoV and PMV viral species in these key reservoirs, yielding >800 unique viral strains.

1.2.c. Sample testing to identify known and novel viruses: We have conducted extensive wildlife surveillance for novel viruses across Southeast Asia in prior NIAID funded work and as part of the USAID PREDICT project. Under the 10-year PREDICT program globally, we identified 912 novel and 210 known viruses, more than the total number of viruses previously recognized in mammals by the ICTV (97). This work included collecting nearly 300,000 individual mammal samples from 14 EHA-led PREDICT countries, and PCR-screening of more than 60,000 individual specimens (53). In southern China alone, we identified 178 β-CoVs, of which 172 were novel, discovered a new β-CoV clade, "lineage E" (41), diverse HKU3r-CoVs (179 sequences) within a 'sister' clade to the SARS-CoV lineage, and a new bat-origin SADS-CoV, responsible for killing >25,000 pigs in Guangdong Province (17). We have collected 28,957 samples from bats, rodents and NHPs in Thailand and 47,178 in Malaysia, **but have only tested a minority of these using PCR.** We have identified 100 novel viruses in Thailand and 77 in Malaysia. **Furthermore, we have only sequenced a small segment of one gene for all novel viruses found. We have archived duplicate samples which are now available for use in this project,** including fecal, oral, urogenital, serum samples, and biopsies from bats, rodents, and NHPs. In the current proposed work, we will attempt to sequence, isolate and characterize those viruses that our analyses below suggest are most likely to be able to infect humans.

1.2.d. In vitro & in vivo *characterization of viral potential for human infection:* Using Coronaviridae as an example, we have conducted *in vitro* and *in vivo* experiments to characterize the pathogenic potential of novel SARSr-CoVs from bats. We isolated three SARSr-CoVs from bat feces: WIV1, WIV16 and Rs4874, with Spike (S) protein sequences that diverged from SARS-CoV by 3-7% (24, 61, 98). We conducted full-length genome sequencing of 12 other novel SARSr-CoVs, some highly similar to outbreak SARS-CoV in the most variable genes (61). Using our reverse genetics system we constructed chimeric viruses and rederived full length recombinant SARSr-CoV from in silco sequence. All 3 SARSr-CoV full length isolates and the two chimeric viruses replicated efficiently in Vero E6 cells and in HeLa cells expressing hACE2, civet and bat ACE2, but not in those without ACE2 (24, 61, 98). We used the SARS-CoV reverse genetics system (72) to generate a chimeric virus with a mouse-adapted SARS-CoV backbone expressing SHC014 S protein with 10% sequence divergence from SARS-CoV S. This and the other full length and chimera viruses replicated in primary human airway epithelium, using the human ACE2 receptor to enter into cells (62). Thus, **SARSr-CoVs with diverse variants of SL-CoV S protein without deletions in their RBD can use human ACE2 as receptor for cell entry.** The Baric lab has a well-established hACE2 and DPP4 transgenic mouse model that **we used to assess the capacity of novel SARSr-CoVs and MERS-CoV to infect humans and cause disease** (85, 99). Infection of bat SARSr-SHC014 or WIV1 caused SARS-like clinical signs in the transgenic hACE2 mouse model **that weren't reduced by immune-therapeutic monoclonals that attenuate SARS-CoV pathogenicity, nor after challenge following vaccination against SARS-CoV** (62) **(Fig. 7)**. We repeated
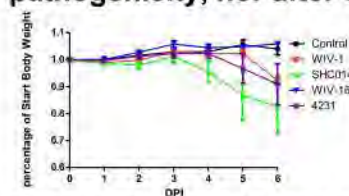
this virus characterization approach with chimeras using HKU3r-CoV S proteins that are ~25% divergent from SARS-CoV S, **and found that they are unable to use the ACE2 receptor.** Additionally, we were unable to culture HKU3r-CoVs in Vero E6, or human cell lines.

Fig. 7: Body weight change in hACE2 transgenic mice after SARSr-CoV infection, demonstrating that bat-origin SARSr-CoVs cause similar clinical signs to outbreak SARS-CoV strain (4231) (62, 63).

7

A similar approach to the above pipeline for CoVs will be applied to novel henipaviruses we discover during our research. The broad mammalian tropism of HeV and NiV (68) is likely mediated by their usage of highly conserved ephrin ligands for cell entry (19, 100). A third isolated henipavirus, Cedar virus does not cause pathogenesis in animal models. **Co-Is Broder and Laing** have developed a reverse genetics system and rescued a recombinant CedV to test ephrin receptor usage/tropism and pathogenicity during henipavirus infections (77). CedV is unable to use the Ephrin-B3 receptor (77, 101) which is found in spinal cord and may underlie NiV encephalitis (102) and that pathogenesis also involves virulence factors V and W proteins. The putative henipaviruses, Ghana virus and Mòjiāng virus, predict V and W protein expression, with GhV able to bind to ephrin-B2, but not -B2 (103), and the receptor for MoJV remains unknown (104) but is likely ephrin-B2 (105). Our group has also used similar methods to test the hypothesis that novel Filoviruses discovered in wildlife are able to infect people. NiV, EBOV and MERS-CoV all infect primary microvascular endothelial cells, which are available at Co-I Baric's lab for EID-SEARCH collaborators (71, 106). Primary human lung endothelial cells are highly susceptible to Ebolavirus infection (107), and Huh7 cells can be used to isolate virus from clinical samples (108) offering advantages for reverse genetic recovery of novel FVs (109). The three filovirus genera, *Ebolavirus*, *Marburgvirus* and *Cuevavirus*, use Niemann–Pick type C1 (NPC1) protein as cell entry receptor (110). For novel filoviruses, cell culture would be carried out under BSL-4, however they can readily be characterized following identification of their encoded GP sequences which can be pseudotyped into recombinant GP-bearing vesicular stomatitis virus (VSV) pseudovirions (111). These pseudovirions can be safely used to characterize the tropism and entry filoviruses mediated by GP without a need for BSL4 containment. **Co-Is Wang and Anderson** used cells originating from various mammalian species to determine spillover and/or zoonotic potential of MLAV, a newly discovered Asian filovirus (47). Despite the low amino acid sequence identity (22–39%) of the glycoprotein with other filoviruses, MLAV is capable of using NPC1 as entry receptor. Cells from humans, monkeys, dogs, hamsters and bats were able to be infected with a MLAV pseudovirus, implying a broad species cell tropism with a high risk of interspecies spillover transmission.

Mouse models. For novel viruses across all families, we will use the UNC collaborative cross mouse that captures 90% of the diversity within all mouse groups to assess the most suitable model for further experiments (112). We have used this model for CoV, FV (Ebola), Flaviviruses, and alphaviruses infections. Clinical signs such as weight loss, hemorrhage, encephalitis, and, acute or chronic arthritis were recapitulated in these mice following SARS, EBOV, West Nile virus and Chikungunya infections, respectively (86-88, 113-116). Note that CC OR3015 shows hemorrhagic disease phenotypes after EBOV infection **(Fig. 8)**. *Bat Models.* Duke-NUS has developed two models for bat *in vivo* studies: a colony of *Eonycteris spelaea* (cave

dwelling small fruit bat) and a bat mouse model, which contains the bat immune system in an immunodeficient mouse (117). In a proof of concept study, **Co-Is Wang and Anderson** infected *E. spelaea* bats with MERS-CoV to demonstrate their susceptibility and suitability for infection with a bat borne virus under BSL3 containment. These bats and the bat mouse model will serve to complement and expand studies from the UNC collaborative cross mouse, pending further funding from EIDRC CC. Recombinant viruses, transgenic mouse models and experimental recombinant protein constructs described above will be made available to the EID-SEARCH consortium and other EIDRCs following standard procedures **(see Resource Sharing Plan)**.
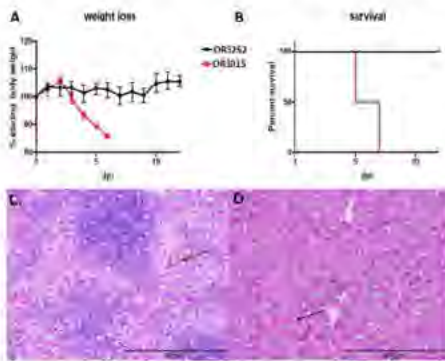
Fig. 8: EBOV Infection in Collaborative Cross Mouse. **Panel A/B:** Wt EBOV in two different CC lines demonstrate different disease patterns. **Panel C/D:** Hemorrhagic phenotypes on d. 6 (arrow) in spleen and liver, resp. From (86).

1.3. General Approach: We will conduct analyses to geographically- and taxonomically-target testing of samples from wild mammals that are most likely to harbor known high-profile zoonotic pathogens, or close relatives with potential to infect people. This includes selecting from tens of thousands of samples collected on recent projects and stored in freezers in our laboratories, and new sampling of wildlife in high risk locales. We will use viral family level PCR assays to identify and partially characterize viruses, then follow up sequencing of relevant receptor binding proteins to assess homology and predict binding affinity, then attempt to isolate and

8

biologically characterize viruses that phylogenetic and other data suggest have high spillover and pandemic potential. We will follow this with *in vitro* binding assays and cell culture experiments, and *in vivo* animal model infections to assess their potential to infect people and cause disease **(Fig. 9)**. EHA will lead the study design, targeted sampling, and data analysis; Co-Is Hughes and Wacharapluesadee will lead field sampling and in-country testing; Duke-NUS, UNC, USUS will lead the serological and PCR diagnostic development and the *in vitro* and *in vivo* analyses, as well as cross-training and tech transfer to the hub labs in Thailand and Malaysia. This will include via support of one Malaysian Ph.D student, Mei-Ho Lee (CM Ltd.) who has worked with the EID-SEARCH collaborators for 9 years. The focus of her study will be infection potential of novel CoVs and PMVs already found in bats in Malaysia under preliminary data for this proposal.
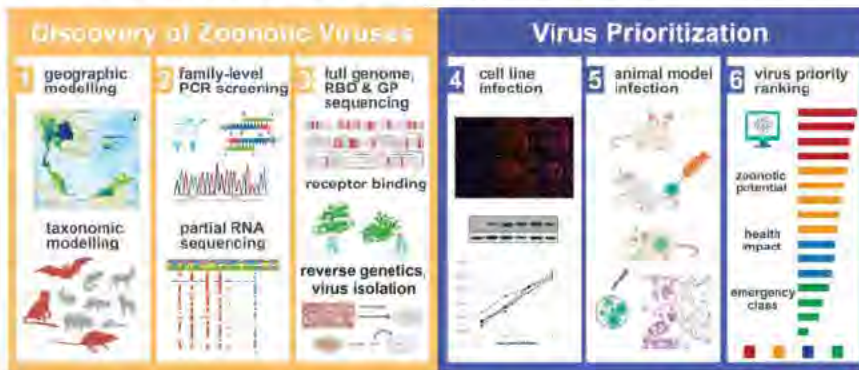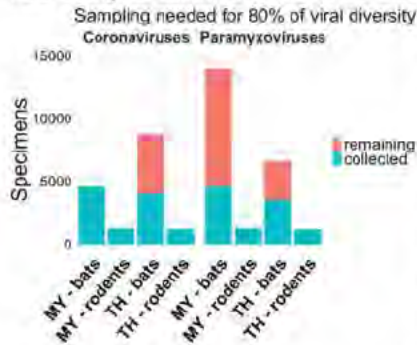

Fig. 9: Aim 1 will use analytical tools to target selection of archived and newly collected samples, conduct PCR and sequencing of novel viruses, recover by reverse genetics, and assess zoonotic potential using *in vitro* and *in vitro* models and analyses.

1.4 Wildlife samples: 1.4.a. Geographic and taxonomic targeting for newly collected wildlife samples: We will prioritize sites within each country that rank highest for both the risk of zoonotic disease emergence (*2*) and the predicted number of 'missing' zoonotic viruses (*3*). Our preliminary analysis **(Fig. 1)** suggests priority areas include: the Kra Isthmus (S. Thailand) and adjacent forests in N. Peninsular Malaysia, NW Thailand (near the Myanmar-Laos border), and lowland rainforests of Sarawak and Sabah. In each of these 'hottest of the hotspots' regions, we will identify 2-4 distinct sampling sites based on the geographic distribution of key hosts we prioritize for additional wildlife sampling. We will priority rank bat, rodent, and NHP species using host trait-based models (*3*). We will also identify species predicted to be centers of evolutionary diversity for CoVs, PMVs, and FVs using ancestral host reconstruction of available sequence data and Bayesian discrete phylogeographic models (*118-120*). We will leverage our strong collaborations with wildlife experts in Malaysia and Thailand, including many of whom are based at our core partner institutions, to ground-truth geographic and taxonomic model outputs with local expert knowledge to ensure the sampling plan is realistic based on current species distributions, population size estimates, seasonality, and accessibility to sites. Most zoonotic viruses are naturally carried by multiple wildlife host species, and there are strong, predictable relations between host phylogeny and viral taxonomy (*3*). We will use a combined network analysis and phylogeographic model to rapidly predict and prioritize new (unsampled) hosts for important, novel viruses we discover during our research. We have shown this relatively simple model (using just wildlife species range overlap and phylogenetic similarity between host species) is both generalizable to estimate host range for all mammalian and avian viruses *and* robust – accurately predicting hosts 98% of the time when cross-validated using data from known viruses (*121*).

1.4.b. Sample size justifications for testing new and archived wildlife specimens: We calculate initial sample sizes targets using our preliminary data on average PCR prevalence from screening tens of thousands of bat, rodent and NHP specimens for CoV and PMVs based on previous studies in Malaysia and Thailand. We will refine these sample size targets, in real-time, over the course of our project using viral discovery curve analyses **(Fig. 6)** to either scale-up or scale-back new sampling and testing of archived specimens for each species depending on viral capture rates. We will only collect additional specimens projected from each prioritized taxa to achieve statistically robust sample sizes and to complement our archived specimens. For example, *Rousettus* spp. bats, known FV reservoirs, were not adequately sampled under PREDICT research in Thailand. Preliminary analysis indicates that, for all viral target families, we will require ~14,000 new samples to identify 80% of remaining viral species diversity in our study region, ~5,000 samples from Thailand and ~9,000 samples from Malaysia **(Fig. 10)**. Viral strain diversity from a sampling effort this size is expected to number in the hundreds of strains. For example, given ~6% prevalence of CoVs in bat species previously sampled under the PREDICT project, a target of 14,000 individuals would yield ~800 PCR positive individuals,

9

representing, an estimated ~600 novel sequences/strains. Similarly, for PMVs, which have an average PCR detection prevalence of ~1.5%, sampling of 14,000 individuals would yield ~200 positives and an estimated 150 unique strains. Filovirus estimates are not feasible yet due to the small number of positive samples in prior studies, but will be estimated as the project begins. From each mammal, we will collect serum, whole blood, throat swabs, urine and fecal samples or rectal swab using our previously-validated nondestructive methods. We will sample sites at different time-points throughout each year, and sample an equal number of males and females to account for seasonal or sex-specific differences viral shedding (See Vertebrate Animals) (122, 123). Dead or animals euthanized due to injury will be necropsied.

Fig. 10: Stacked bar plot showing number of currently-archived specimens (blue) and new specimens required (orange) to discover 80% of predicted number of CoVs and PMV species from high-risk bat and rodent taxa in Malaysia and Thailand.

1.4.c. Sample collection, testing, viral isolation: Wildlife will be captured and sampled using previously published techniques, by personnel wearing appropriate personal protective equipment (Tyvek suits or dedicated long external clothing, double nitrile gloves, leather gloves, an N95 or p100 respirator, and safety glasses) (53, 93, 95, 96, 122, 124). All samples will be placed in a vapor phase liquid nitrogen dry shipper immediately upon collection, and then transferred to a -80C freezer once back in the lab, until testing. Viral RNA will be extracted from bat fecal pellets/anal swabs. RNA will be aliquoted, and stored at -80C. PCR screening will be performed with pan-coronavirus (125, 126), filovirus (127) and paramyxovirus primers (128), in addition to virus specific primers where additional sensitivity is warranted in key viral reservoirs, i.e. Nipah virus or MERS virus specific primers (129, 130). PCR products will be gel purified and sequenced. We will attempt isolation on samples that contain viruses of interest (determined in **Aim 1.5 below**), using Vero cells and primary cell culture from the wildlife taxonomic order (bat, rodent and NHP cell lines). Over 70 bat cell lines are maintained at Duke-NUS from five different bat species. These include primary lung, brain, bone marrow, heart, kidney, intestine, liver, lymph node, muscle, spleen, PBMC and ovary/testes from *Eonycteris spelaea*, *Cynopterus bracyhotis*, *Rhinolophus Lepidus*, *Myotis muricola* and *Pteropus alecto* bats. In addition we have 8 immortalized cell lines, including those derived from lungs, therefore suitable for the isolation of respiratory viruses.

1.4.d. Moving beyond RNA viruses: To detect pathogens from other families, such as non-RNA viruses, we will combine conserved PCR assays with NGS. We will perform either unbiased NGS, or use a pan-viral enrichment library to sequence full-length genomes (131). We will conduct this work via EID-CC research projects or as part of outbreak detection/investigation in the region from DNA viruses or other agents.

1.4.e. Sequencing Spike Glycoproteins: Based on earlier studies, our working hypothesis is that zoonotic CoVs, FVs and henipaviruses encode receptor binding glycoproteins that elicit efficient infection of primary or continuous human cell lines (e.g., lung, liver, etc.) (16, 61, 62). Characterization these for SARSr-CoVs suggest that viruses up to 10%, but not 25%, different in the spike glycoprotein use human and bat ACE2 receptors for docking and entry. Uneven sampling gaps (e.g., no strains were found with 10-25% spike variation) prevent a thorough understanding of the transition point where the most divergent strains lose human ACE2 receptor usage. Our viral discovery curves **(Fig. 6)** suggest further sampling will reveal a rich diversity of as-yet-undiscovered SARSr-CoVs and strains of other viruses that fill in the phylogenetic tree between currently-described viruses. In this aim we will use phylogenetic and viral discovery analysis to specifically target bat species and regions that are under-sampled to allow sufficient power to identify and characterize missing strain diversity for potential zoonoses. We will sequence the S proteins of novel SARSr-CoVs to prioritize viruses for experimental work in Aim 3 to test this hypothesis. Likewise, we will test the hypothesis that phylogenetic related zoonotic receptor binding protein genes of MERS-related, Ebola-related and NiV/HeV-related viruses will also program efficient entry via human and ortholog receptors. For all novel virus strains, we will sequence the complete entry spike genes by amplifying overlapping fragments using degenerate primers as shown previously (24, 61). **Reverse Genetics:** Full-length genomes of selected CoVs, FVs or henipaviruses (representative across subclades) will be sequenced using NEBNext Ultra II DNA Library

10

Prep Kit for Illumina and sequenced on a MiSeq sequencer. PCR or MinIon sequencing will be performed to fill gaps in the genome. The full length spike gene sequences, including the amount of variation in the receptor binding residues that bind the appropriate human receptor will be used to select strains for Aim 3 experiments. We will sequence full length genomes of high risk strains that are antigenically distinct and escape SARS or MERS-CoV cross neutralization, synthetically reconstruct a small subset (1-2) and evaluate growth in Vero, BHK and Huh7 cells (63, 132). From full length sequences, synthetic clones will be ordered to recover recombinant viruses using reverse genetics as described by our groups (62, 63, 77, 109, 133, 134). Recombinant virus growth will be accessed in Vero, BHK, Huh7 and select bat cells and recovered viruses used for downstream studies. Virus strain prioritization is described below.

1.5. Assessing risk for spillover. 1.5.a. Capacity to infect human cells: We will use a series of cell cultures to assess the capacity of target viruses to enter human cells. For CoVs, FVs and henipaviruses, we will use primary human ciliated airway epithelial cells (HAE) and lung endothelial cell cultures derived from the lungs of de-identified transplant recipients. HAE are highly differentiated and grown on an air-liquid interface for several weeks prior to use (62, 63, 132). We will prepare HAE and lung endothelial cultures from three different patient codes in triplicate in collaboration with the tissue procurement facility at the Cystic Fibrosis Center at UNC. Cultures will be inoculated with chimeric bat SARSr-CoVs or novel FVs or henipaviruses to assess efficient replication. At 72 hpi, cultures will be fixed for immunofluorescent staining using antisera to the SARS-CoV conserved nucleocapsid protein (N) or to structural genes of the other viruses (64, 71). SARSr-CoVs that differ significantly in S protein sequence (11-24%) from epidemic SARS-CoV yet replicate *in vitro*, will also be evaluated for sensitivity to neutralization in Vero cells by PRNT50 assays using broadly SARS-CoV cross reactive human mAbs S227.14, S230.15, S215.17, and S109.8 (63, 135) or monoclonal antibodies targeting the Ebola or Marburg E glycoprotein (Available through BEI), or NiV/HeV glycoproteins (136). As controls or if antisera is not available, the S genes of novel SARSr-CoV or spike entry genes of other novel viruses under investigation will be inserted into VEE 3526 replicon vectors (VRP3526-S), packaged and used to vaccinate mice (137). Polyclonal sera against test bat-SARS virus or other spike genes will be harvested and tested for ability to cross neutralize SARS-CoV, GD03, WIV-1, SHC014, WIV-16, other novel SARSr-CoV and HKU3-SRBD by PRNT50 assay (63, 138, 139). Using PRNT50 titers with sera (n=4 each) among these viruses, antigenic cartography (140) will allow comparison of antigenic vs. phylogenetic distance, identifying the transition at which SARSr-CoV strains escape SARS-CoV based vaccines, informing immunotherapeutic and vaccine design strategies (141-143). Similar approaches will be applied to novel MERS-related viruses, other CoV, FVs or henipaviruses. While most wildlife-origin viruses with unknown transmission potential or pathogenicity in people uncategorized, we use BSL-3 for isolation as standard, and cell culture for potentially BSL-4 pathogens will be undertaken at collaborating institute NEIDL. To facilitate this, duplicate samples in culture medium from animals PCR +ve for FVs and very close relatives of NiV or HeV will be shipped to NEIDL **(See letter of support)** for culture, characterization and sharing with other approved agencies including NIH Rocky Mountain Laboratories where PI Daszak and Co-Is Baric have ongoing collaborations. Similarly, some duplicate samples of animals PCR +ve for CoVs will be shipped to UNC for further characterization by Co-I Baric. Where feasible and appropriate, training opportunities at NEIDL and UNC for in-country staff will be given, thereby enriching hands-on collaboration across sites.

1.5.b. Host ACE2, DPPR, Niemann-Pick C1 (NPC1) and EphrinB2/EphrinB3 receptors: We will sequence relevant host receptors for select mammalian wildlife species we find positive for strains of high-priority CoV clades, and all new henipaviruses and FVs. Of particular interest is homology across 18 bat and human orthologue ACE2 contact interface residues that engage the SARS RBD as a potential indicator of SARSr-CoV cross species transmission potential and growth in human cells (70). Likewise, variation in the NPC1 receptor has been shown to alter EBOV ability to infect in humans and bats, so this can be structurally mapped for prediction and downselection of strains (144). Signatures of positive selection in bat NPC1 were concentrated at the virus-receptor interface, with the strongest signal at the same residue that controls EBOV infection in *Eidolon helvum* cells. Thus NPC1 is a genetic determinant of FV susceptibility in bats, and some NPC1 variations likely reflect host adaptations to reduce FV replication and virulence (145). NiV and HeV use ephrin B2 or ephrin B3 as a receptor, and variation affects its ability to use human vs bat receptor molecules (103, 146).

11

1.5.c. Host-virus evolution and predicting receptor binding: We will use Bayesian and Maximum Likelihood phylogenetic analyses of RNA-dependent RNA polymerase (RdRp) sequences from PCR screening, receptor binding glycoproteins, and/or full genome sequence data (when available) to reconstruct the evolutionary history of the novel bat SARSr-CoVs, bat MERSr-CoV, FVs and henipaviruses that we identify. We will run co-phylogenetic analyses to map out instances of host-switching (*147, 148*) and ancestral state reconstruction analyses **(Fig. 4)** to identify the most important species for shaping evolutionary diversity for these viruses (*51, 149*), allowing for more targeted future surveillance and spillover mitigation interventions. For receptor binding glycoprotein sequence data from characterized viruses, we will apply codon usage analyses and codon adaptation indices, as used recently on henipaviruses (*150*), to assess the relative contributions of natural selection and mutation pressure in shaping host adaptation and host range.

1.5.d. Viral strain prioritization: Of the expected 100-200 novel SARSr-CoV strains, and approximately 600 total CoV strains, we will down-select to prioritize for further characterization based on S genes that are: i) different from SHC014, WIV1, SARS-CoV with diversity ranges of 10-25%; ii) have virus S RBD that could use human/bat receptors; iii) have recombinant chimeric spikes indicative of gene flow between clade I and II strains; iv) have bat ACE2 or DPP4 receptors that might select for spike RBDs that can use human receptors for entry (15/18 conserved residues in human/bat ACE2 molecules that bind SARSs-CoV S RBD domains are likely more efficient receptors than 3/18 conserved sites). Using structural models based on the SARS S glycoprotein, the extent and location of antigenic variation will be annotated onto the structure, further resolving the locations of highly evolving and conserved structural motifs/epitopes that function in protective immunity (*135, 151-153*). For novel henipaviruses, we will examine the receptor-binding proteins (e.g. envelope attachment (G) proteins) for sequence and structural homology with particular attention to conversation of amino acids that mediate ephrin receptor binding. We will prioritize strains that are predicted to interact with ephrin-B2 and expression V/W proteins to determine whether these viruses represent zoonotic threats. Using the recombinant Cedar virus (rCedV) platform, we will be able to rescue rCedV that express heterotypic G proteins to explore receptor utilization and tissue dissemination in animal models. We will assess ability to infect human cells using a continuous lung epithelial cell line, Calu3, which our group has previously used to generate data on NiV and HeV growth curves and a variety of host expression changes after infection (*154*). We will also use primary human airway cultures and microvascular endothelial cells to assess human infection for FVs, henipaviruses and MERSr- and SARSr-CoVs (*155*) (*156*) . There is some evidence for NiV and HeV replication in mouse lung, low level titers in young animals, 10 fold higher titers in aged mice (*157, 158*). Thus it is likely these models could be improved in the collaborative cross mouse resource at UNC. We have identified CC lines that are highly vulnerable to wildtype SARS and others that develop hemorrhagic disease following mouse adapted Ebola infections (*86-88*).

1.5.e. Animal models: We will use <u>transgenic mouse models</u> and the <u>UNC Collaborative Cross model</u> to assess spillover potential of viruses. For work on CoVs, we will follow the protocols in the <u>transgenic mouse model</u> used by the Baric lab previously. Briefly, in the BSL3 lab, 10- to 20-week old hACE2 or hDPP4 transgenic mice will be intranasally inoculated with 5 x 10^4 PFU of full length wildtype rbat CoV, or chimeric SARSr-CoV or MERSr-CoV encoding different bat CoV spike proteins respectively, then monitored daily for weight loss, morbidity, and clinical signs of disease. Mice will be euthanized at 2, 4, and 6 dpi (endpoint of 14 dpi), organs harvested and virus quantified by plaque assay or genome (mRNA1) RT-PCR. After 7 days in formalin, tissues for histopathology will be removed from the BSL3 and stained with H&E, and for immunohistochemistry using polyclonal antibodies against the N protein. We will conduct limited evaluation of existing countermeasures using therapeutic monoclonal antibodies *in vitro*. Existing SARSr-CoV or MERS-CoV mAbs (*159, 160*) will be diluted in DMEM starting at 1:20, and serial dilutions mixed with an equal volume of 50 PFU of chimeric bat SARSr-CoVs encoding different spike proteins and RFP indicator genes, then incubated on Vero E6 cells at 37°C for 1 h. Antibody neutralization titers will be evaluated by plaque quantification at 4-5dpi. In select instances, hACE2 transgenic mice will be injected with SARS-CoV mAbs, and infected with chimeric bat SARSr-CoVs. Clinical signs and morbidity will be assessed and tissue pathology examined and compared with mice without treatment of mAbs to determine the therapeutic effect on SARSr-CoV infection, and protection of SARSr-CoV by wildtype SARS-S based vaccines assessed as described (*132, 161*). We will sequence full length genomes of high risk strains that are antigenically distinct and escape SARS or MERS-

12

CoV cross neutralization, synthetically reconstruct a small subset (1-2), recover full length viruses and evaluate the ability of nucleoside analogues to inhibit growth in HAE cultures and/or *in vivo* (*63, 132*). For the Collaborative Cross model, we will infect six to eight mice from 6-8 Collaborative Cross mouse lines (10 weeks of age) with six test viruses/yr, intransasally or by subcutaneous infection with 1 x 10e4 virus. Animals will be followed for weight loss, morbidity, mortality and other clinical disease symptoms (e.g., hunching, hindleg paralysis). One half the animals will be sacrificed at day 3 and day 7 pi. to evaluate virus growth in various organs, and organ pathology. In highly vulnerable lines, additional experiments will be performed using more animals, evaluating virus growth, clinical disease symptoms, respiratory pathology, tropism and organ pathology through day 28 pi. We anticipate that some CC lines will prove highly vulnerable to select wildtype viruses, because of host susceptibility combinations that promote severe disease outcomes.

1.6. Potential problems/alternative approaches: Challenges with logistics or obtaining permission for wildlife sampling in sites we select. We have a >20-year track record of successful field work in Malaysia, and >10 years in Thailand, and have strong established partnerships with local wildlife authorities and communities to ensure access. We have existing permissions in place which will be renewed at the start of this project. Due to the abundance of wildlife habitats in the region, if access to a site is withdrawn, we will rapidly identify a suitable, complementary site with similar bat species composition and abundance. Finally, we have already collected tens of thousands of wildlife samples from the region and will focus on these while waiting for permission. **We may not identify novel viruses in our sample wildlife species due to seasonality of viral shedding or other factors.** We have calculated sample sizes based on realistic detection rates from our extensive preliminary data, and are confident that we will detect and identify large numbers of potentially zoonotic pathogens, particularly for CoVs and PMVs. We acknowledge that FV strain diversity may be harder to capture, given relatively low seroprevalence in bat populations – suggesting lower levels of viral circulation (*42, 43*). However, our team was the first and only group to sequence FV RNA from Asian wildlife species (*46, 47, 57*), and we are confident that with our targeted sampling and testing strategy offers the best chance globally to identify additional strains of Ebola and related viruses. The sampling regions we have selected are subtropical, and our previous data, and even published studies in temperate regions (*162*), do not suggest a strong pattern of seasonality in CoV shedding, and where seasonal patterns do occur, i.e. for henipaviruses in Thailand (*163*), or can be predicted from wildlife serology data, we will be sure to conduct sampling at different timepoints throughout the year to account for this.

<u>Aim 2: Identify evidence and analyze risk factors for viral spillover in high-risk communities using novel serological assays.</u>

2.1 Rationale/Innovation: Rapid response requires early detection. Assessing the spillover of rare or novel zoonotic agents will require targeted surveillance of high-risk communities that interface frequently with wildlife that harbor high risk emerging viral strains. To enhance the low statistical probability of identifying these rare events, In Aim 2 we will target rural human communities inhabiting regions identified in Aim 1 as having high viral diversity in wildlife, that also engage in practices that increase the risk of spillover (e.g. hunting and butchering wildlife). We will initially identify 4 subsites in each of Thailand, Peninsular Malaysia, Sarawak, and Sabah for sampling. We will design and deploy human risk factor questionnaires, and specific and sensitive serological assays to identify the baseline frequency, and causes of, spillover of known and related novel viral pathogens in these populations. **We will use the results to test a key hypothesis: that hunting wildlife increases the risk of exposure to novel viruses, compared to the rest of the community.** The alternative hypothesis is that seroprevalence <u>in the general community is not statistically different to those who hunt and butcher wildlife, because exposure to wildlife pathogens is largely through inhabiting a biodiverse landscape where wildlife make indirect contact with people</u> (e.g. via fomites). For participants who report symptoms related to 1) severe/acute respiratory illness (SARI/ARI); 2) Influenza-like illness (ILI); 3) fever of unknown origin (FUO); 4) encephalitis; 5) hemorrhagic fever; or 6) unusual presentation of high fever with severe diarrhea; all within the last 10 days, an additional sample type will be collected, two nasal or oropharyngeal swabs. Samples will be marked for additional PCR-based assays to identify presence of known and novel CoVs, henipaviruses and FVs, and for isolation and biological characterization of potential pathogens, using the approach laid out in Aim 1.

13

2.2 Preliminary data on high-zoonotic risk human community surveillance: 2.2.a Biological sample collection in SE Asia: Our long-term collaboration in the region has included qualitative and quantitative survey data, and collecting large archives of banked human sera and other biological samples from high zoonotic-risk populations. For example, under the USAID-PREDICT project, EHA has collected 1,400 and 678 specimens in Malaysia and Thailand respectively from people living in zoonotic disease interfaces, all of which will be available for retrospective serological testing by EID-SEARCH which we have ensured is compliant with our existing IRB. Our core group has collected and tested many thousand additional specimens from other important community cohorts **(Table 2)**, and some of these will continue under EID-SEARCH **(Section 2.4)**.

Country, site	Lead	# enrolled, focus	Findings
Peninsular Malaysia	Hughes	9,800+ samples, Orang Asli indigenous pop., for PCR/serol.	25+ novel CoVs, influenza, Nipah ab+ve, FV ab+ve
Malaysia Sabah	Kamruddin	1,283 for serology	40% JE, 5% ZIKV ab+ve
Malaysia Sabah	William	10,800 for zoonotic malaria study	High burden of macaque-origin malaria
Malaysia Sabah	Hughes	150 bat cave workers	Ongoing
Malaysia Sarawak	Siang	500 Bidayuh and Iban people	8% PCR prevalence HPV in women
Malaysia PREDICT	Hughes	1,400 high zoonotic-risk communities for viral PCR, serology	Ongoing. Multiple known and novel CoVs, PMVs, others.
Thailand	Wacharaplu-esadee	100s of bat guano harvesters/villagers	Novel HKU1-CoV assoc. clinical findings
Thailand PREDICT	Wacharaplu-esadee	678 high zoonotic-risk communities for viral PCR, serology	Ongoing. Multiple known and novel CoVs, PMVs, others.
Singapore	Wang	856, for Melaka virus	7-11% MELV ab+ve

Table 2: Biological sample collection from healthy populations conducted <u>by members of **EID-SEARCH**</u> in our hub countries.

Of particular interest is the access to indigenous communities in Peninsular Malaysia (Orang Asli) and Sarawak (Bidayuh, Iban people), and the bat guano harvesters in Thailand. These communities hunt, butcher and eat wildlife daily, and work in supremely high-risk zoonotic interfaces (e.g. bat caves). <u>One of our indigenous community sites is experiencing an ongoing outbreak of unknown etiology that we are investigating currently</u> **(Section 3.2.a)**.

2.2.b Human Contact Risk Factors: EHA is the lead organization across all 27 USAID-PREDICT project countries for collecting data on, and analyzing the human aspects of, zoonotic spillover risk (*164, 165*). We conducted exploratory studies using standardized one-on-one semi-structured interviews and site observation data with people living in hotspot areas or engaged in high-zoonotic spillover risk activities, using these data to assess local social and cultural norms underlying wildlife contact (*165*). We then developed a risk factor survey to accompany human biological sampling to evaluate the type and frequency of animal contact, wildlife exposure in daily life, occupational risk in markets and long the wildlife value chain, and unusual illnesses reported over the past 12 months. We conducted cross-sectional studies in high risk populations in 27 countries, and study participants provided biological samples (oral/oropharyngeal swab, nasal/nasopharyngeal swab, urine sample/urogenital swab, fecal sample/rectal swab, whole blood, serum, plasma). For example, survey and biological samples were collected from: 1,585 participants from 7 sites in Yunnan, Guangxi, and Guangdong provinces, China; 1,400 participants from 4 sites in Malaysia; and 678 participants from 4 sites in Thailand. Samples were tested with novel serological assays and results analyzed against risk factors from survey data (*16*). In China, CoV-seropositive individuals were mostly farmers/peasants (8/9), living in Yunnan province (7/9), 41-60 yrs old (7/9), with domestic contact with rodents (6/9), that are male (6/9). Although these results are preliminary and don't provide detailed information on routes of exposure, <u>they identify key subpopulations to target that would likely increase the opportunity for identifying evidence of spillover for other CoVs, PMVs and FVs.</u> Data from PREDICT surveys have been used to help identify target populations in Aim 2, identify strategies to better target at-risk people, and conduct focused **human risk factor surveys and serosurveys to produce statistically significant findings for henipaviruses, CoVs and FVs.** In Aim 2, we will use combined biological sampling and risk factor surveillance in targeted populations within the community to 1) test the hypothesis that occupational (i.e. wildlife hunting) risk factors correlate with seropositivity to henipaviruses, FVs and CoVs; 2) assess possible health effects of infection in people; and 3) where recent

14

illness is reported, obtain biological samples for PCR testing and potential viral isolation and characterization. Obtaining this information could be a significant step in understanding the likelihood of recent 'hidden' spillover events and their public health impacts, as well as the risk of future emergence of viral disease.

2.2.c. Risk factors for illness: Our preliminary data in Thailand, Malaysia and other SE Asian countries revealed that frequent contact with wild (e.g. bats, rodents, NHPs) and domestic animals (e.g. poultry, swine) among local communities was associated with a high percentage of self-reported severe/acute respiratory illness (SARI/ARI) and Influenza-like illness (ILI) symptoms **(Section 3.2.b)**. In addition, people involved in crop production, hunting and butchering of wildlife reported statistically significantly higher SARI and ILI symptoms (Li *et al.*, in prep.). **These preliminary data suggest a target population and sites for current proposal with the correct targets for serological surveillance and risk factor surveys in Aim 2 to produce statistically significant findings.** Coupled with better serological tools from our team **(Section 2.2.d)**, we will be able to identify the likely routes of exposure to known or novel coronaviruses, henipaviruses, and FVs in the study region, as well as build data on the illnesses they cause in people.

2.2.d. Serological platform development: Most emerging viruses produced a short-lived viremia in people so that large samples sizes are required to provide enough PCR-positive individuals to analyze infection patterns. For Aim 2, we will focus on serological sampling in the community, because antibodies are typically long-lived and can be used to assess prior virus exposure or infection with much smaller samples sizes (*123*). Most serological assays target a single protein, and for emerging viruses, it's often unclear which viral antigen will be immunodominant during human infection, and therefore which to target in developing serological assays. Co-Is Wang, Anderson (Duke-NUS) and Broder, Laing (USUS) have developed a series of approaches to deal with this problem, and which will be used in testing human biological samples in Aims 2 and 3. Henipaviruses, FV and CoVs express envelope receptor-binding proteins that are the target of protective antibody responses. Co-I Broder has led efforts to express and purify henipavirus native-like virus surface proteins for immunoassay application (*166, 167*), developing monoclonal antibodies (*168, 169*) and as subunit vaccines (*170, 171*), has developed a novel, multiplex microsphere immunoassay (MMIA) allowing serum samples to be simultaneously tested for antibodies reactive with henipaviruses, FVs and CoVs. This MMIA uses a Luminex-based multiplex platform and custom designed virus receptor-binding proteins expressed in a mammalian cell-culture system and purified as native-like oligomeric proteins. These oligomeric antigens retain post-translational modifications and quaternary structures, allowing capture of conformational-dependent antibodies raised during a natural infection, which represent the majority of the protective humoral response. The known diversity of henipaviruses includes five presently described viruses, NiV, HeV, CedV, GhV and MojV and this assay has been expanded to include all described species **(Fig. 11)**. In the past, our collaborators have used NiV and HeV G protein antigens in a MMIA assay to identify a serological footprint of NiV-related African henipaviruses and exposure to virus(es) antigenically-related to NiV in livestock sampled in Bangladesh (*78, 79*). To strengthen serological data interpretation, the G protein from each virus is integrated
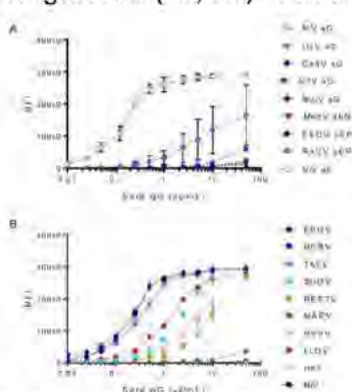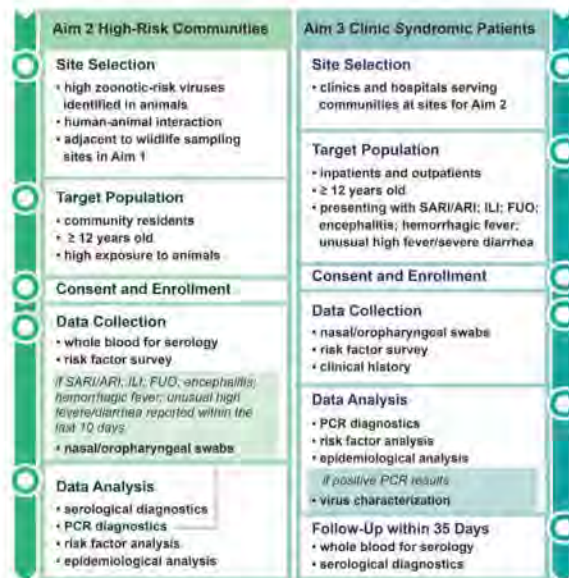

into the MMIA so that virus-specific vs. henipavirus cross-reactive antibodies can be detected and address specificity of past virus exposure. Extensive cross-reactivity exists among ebolaviruses and between ebolaviruses and marburgviruses (*172-174*). We are validating the MMIA pan-FV assay for specificity, and as part of this, in collaboration with Duke-NUS, we have found serological evidence of past exposure to a filovirus(es) most antigenically-related to ebolaviruses (e.g. Ebola virus, Bundibugyo virus), but unreactive with Reston virus GP in three under-sampled fruit bat species, suggesting that novel FVs circulate within bat hosts in SE Asia (*43*). **This platform has now been set up for use in Malaysia and Thailand and will be available for use in this proposed work.** Co-Is Hughes, Broder, Laing have conducted initial screening of human, wildlife and livestock samples at high-risk interfaces in Malaysia, finding serological evidence of past exposure to viruses antigenically-related to NiV and EBOV in humans, bats and non-human primates (NHPs).

Fig. 11: Validation of multiplex microsphere immunoassay (MMIA) specificity and identification of immunologically cross-reactive viruses for A) NiV B) EBOV.

15

Co-Is Wang and Anderson (Duke-NUS) have developed a phage-display method to screen antibodies for all known and related bat-borne viruses, including henipaviruses, FVs and CoVs and have set up a similar Luminex-based assay. Additionally the Duke-NUS team has demonstrated capacity for rapid and cost-effective development of LIPS serological assays against novel viral agents **(Sections 2.6.a, 3.2.a)**. These will be used for verification and expanded surveillance. For molecular investigation (Aim 2 & 3), we will combine targeted PCR with a discovery-driven enrichment NGS strategy to not only survey currently known viruses, but also to discover potential novel viruses. Both the phage-display serological and enrichment platforms can be rapidly updated to include new viruses identified.

2.3 General Approach/Innovation: Our human sampling approach uses two approaches in Aims 1 & 2 (Fig. 12). In Aim 2 we will conduct community-based surveillance using a brief, focused questionnaire to identify the prevalence and frequency of risk-factors for viral spillover in tandem with collection of high quality biological samples to determine the seroprevalence of spillover viruses in at-risk human populations. We will use generalized linear models to analyze correlation between biological results and risk-factors for viral spillover. **In Aim 3**, we will conduct clinic-based syndromic surveillance and biological sampling at clinics/hospitals that are the catchment healthcare facilities for the communities sampled in Aim 2. Both community-based and clinic-based syndromic surveillance programs are cross sectional case-control studies designed with the sample sizes necessary to statistically quantify (with a power of 80%) risk factors and health impacts for viral spillover, linked to serological status and PCR prevalence in symptomatic patients. All human sampling conducted under Aims 2 and 3 will adhere to strict criteria for inclusion and recruitment and protection of human subjects from research risk **(see Human Subjects and Clinical Trials Information)**.

Fig. 12: Human sampling approach. Aim 2 **(left)** focuses on serology in high-risk communities to identify evidence of population exposure to novel zoonotic viruses. Aim 3 **(right)** sets up clinical cohorts at hospitals serving these high-risk communities to conduct syndromic surveillance and identify the etiology of new viral syndromes

2.4 Target population & sample sizes: We will target populations in the same geographic regions as those identified in Aim 1 that harbor wildlife with high potential zoonotic viral diversity, that are EID hotspots, and that are well-connected through regional travel and trade hubs to the global travel network **(Fig 13)**. Within each geographic region we will identify populations at four sub-sites, building on our preliminary data **(Table 2)**. We will target specific communities based on our analysis of previously collected behavioral questionnaire data, and other publicly available data (e.g. wildlife hunting licenses sold in a region, number of guano caves accessed for fertilizer etc.) to identify key at-risk populations and occupations with high exposure to wildlife. We will conduct concurrent sampling trips at visits to sub-sites during which members of the community will be enrolled and wildlife surveillance activities conducted. Individuals living or working around wildlife habitats (e.g. bat caves/roosts), who hunt wildlife, work with wildlife or livestock, or slaughtering wildlife will be considered for enrollment. Participants will complete a consent form and questionnaire in addition to having biological samples collected. During these community trips after data collection the field research staff will conduct community meetings to help inform the public health risk of zoonoses. **Target populations:** *Thailand (Co-I Wacharapluesadee):* 100 subjects of healthy high-risk community from 2 study sites, Chonburi (NiV, other PMVs and CoVs found by PCR in bats in our previous work), and Ratchaburi (bat guano harvester villages where we found a MERSr-CoV in bat guano (*50, 175*) and serological evidence of α-CoVs, HKU1-CoV in people (*175*). *Peninsular Malaysia (Co-I Hughes, CM Ltd.):* We will expand our sampling of the Orang Asli indigenous communities to include sub-sites in communities in Districts we have already sampled and

16

additional Districts in the States of Kelantan Perak, Pahang, and Kelantan all close to bat caves. Samples will be stored and tested at NPHL. *Sarawak (Co-I Tan, Fac. Med. Hlth. Sci. UNIMAS):* We will expand our sampling of the Bidayuh and Iban indigenous people to include the Dayak and Orang Ulu ("people of the interior") because this group has a more intimate connection to remote wildlife-rich areas than the other two, albeit all are hunters. Human sampling will be led by UniMAS in collaboration with CM Ltd / EHA. *Sabah:* We will conduct sampling at the Madai bat and bird cave that has 200 permanent residents who harvest bird's nests for soup, and >1,000 during the harvesting season and at Gomantong cave – one of the largest bat and cave swiftlet nesting sites in the world. *Sabah: (Co-I Hughes):* We will conduct sampling at the Madai bat and bird cave that has 200 permanent residents who harvest bird's nests for soup, and >1,000 during the harvesting season and at Gomantong cave – one of the largest bat and cave swiftlet nesting sites in the world. Protocols follow those for Peninsular Malaysia *Singapore: (Co-I Wang, Duke-NUS):* Active human surveillance will be not be conducted in Singapore, but archived human specimens e.g. from our Melaka virus serosurvey (8), will be made available to conduct retrospective serosurveys for specific pathogens that are identified as high interest to the EIDRC, NIH, and EID-CC.
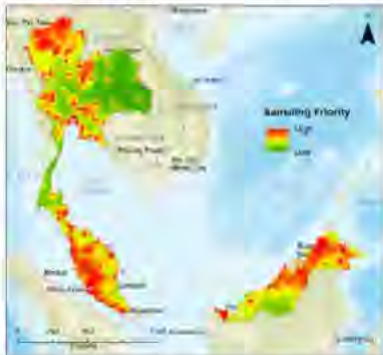

Fig. 13: Targeting high-risk sites for human zoonotic disease surveillance in Thailand and Malaysia. This map plots areas with high numbers of expected viruses in wildlife (3), greater spillover probability based on EID risk factors (2), and higher levels of connectivity (measured by road density data) that may facilitate person-person spread.

Sample sizes: From our previous work we conservatively anticipate that 10-30% of the community population will have had exposure to wildlife allowing us to capture highly exposed and non-exposure individuals at each site. Individuals living or working around wildlife habitats (e.g. bat caves/roosts), who hunt wildlife, work with wildlife or livestock farming, transportation, selling, or slaughtering wildlife in the surveyed areas will be targeted so that they make up ≥30% of the sampled population in each community. We will stratify sampling to ensure appropriate representation of sex, demographic, and socio-economic factors at each community site. In each of our regions, we will conduct sampling at four sub-sites in populations with high prevalence of high occupational and environmental risk of exposure to wildlife populations, sampling 175 persons at each sub-site. We will target populations with, on average, 30% or higher prevalence of occupational or environmental exposures to wildlife based on data from previous studies. This design will give us >80% power to detect 5 percentage point differences in seroprevalence against any of our viral groups between baseline and high-risk subgroups. We conservatively assume 5% base seroprevalence, as well as assuming high spatio-temporal variance amongst sub-sites, with baseline varying 0-25% amongst sub-sties and the fraction of the population at risk varying from 20-40%. (Power calculation conducted by 500 simulations of a generalized linear mixed model (GLMM)). For community-based surveillance, we will enroll 700 individuals per study region, allowing us to make county/province-level comparisons of differing effects; the total sample will be 3,500 participants over 5 years.

2.5 Data & sample collection: Following enrollment, biological specimens (5mL will be drawn for whole blood and serum collection, samples will be aliquoted, at least one max 500 µL of whole blood and two 500 µL serum samples) will be collected by locally certified healthcare professional proficient in phlebotomy techniques and a short questionnaire will be administered by trained project staff that speak appropriate local language to all participants. Through the serological findings and the survey we will investigate the details of five risk factors related to high risk wildlife exposure, to maximize the power of the analyses, based on continuing analysis of our previous work, and will focus on: 1) occupation and occupational exposures; 2) observed or reported interactions with priority wildlife, especially bats, in/around house; 3) proximity of residence or work place to environments of increased risks, e.g. nearby bat roosts; 4) working or regular visitor to animal markets; 5) self-reported ILI/SARI clinical diagnosis or symptoms in the past 12mo and lifetime. For participants who report the symptoms relating to 1) severe/acute respiratory illness (SARI/ARI); 2) Influenza-like illness (ILI); 3) fever of unknown origin (FUO); 4) encephalitis; or 5) hemorrhagic fever within the last 10 days, an additional sample type will be collected, nasal or oropharyngeal swab (2x). Blood of participants whose serum tests positive for emerging viral pathogens will be taken and Peripheral Blood Mononuclear Cells (PBMCs) frozen for future use.

17

These will be used for future harvesting of polyclonal and monoclonal antibodies as potential therapeutics/ diagnostics **(see Letter of Support NEIDL)**.

2.6: Laboratory analysis: 2.6.a Serological testing: We will screen all human specimens for henipaviruses, FVs and CoVs using our novel, multiplex microsphere immunoassay (MMIA) **(Section 2.2.d)**. We will use ELISA and serum neutralization tests (SNT) as confirmatory, where feasible and appropriate for the biocontainment level given sensitivity and specificity variation, and the need for live virus for SNTs **(See Select Agent Research)**. Serologic evidence of local spillover will be used to prioritize zoonotic strains for recombinant virus recovery from full length sequences. Preliminary data suggest ELISA will be effective for some viral groups. We previously developed a SARSr-CoV specific ELISA for serosurveillance using the purified NP of a bat SARSr-CoV (Rp3). The specificity of this assay was evaluated using polyclonal antibodies against HKU1, OC43, 229E, NL63, MERS-CoV and EBOV and no significant cross-activity detected (*16*). **This suggests that if we can expand our NP serology tests to cover other bat CoVs, and that doing so may identify many more seropositive individuals.** For CoVs, we will therefore use two serological testing approaches: 1) testing human sera collected from both community- and clinic- based sampling for a panel of bat CoVs that will include SARS-CoV, MERS-CoV, SADS-CoV and a range of bat SARSr-CoVs (*16*); 2) testing for antibodies to common human CoVs (HCoV NL63, OC43 – **Section 2.8**). We will test panels of NP and G recombinant proteins to assess if this approach will act as a comparison to MMIA for FVs and PMVs.

2.6.b RT-PCR testing. Specimens from individuals in the community who reported being symptomatic within the last 10 days **(Section 2.5)** will be screened using RT-PCR initially for CoVs, PMVs, and FVs following published protocols **(Section 1.4.c)**. Positive samples will be subjected to full genome sequencing and RT-PCR amplification of the glycoprotein genes. Samples from the clinic-based syndromic surveillance will also be tested using RT-PCR for Influenza A & B, HCoV NL63, OC43, HKU1, SARS-CoV & 229E, NiV, HeV, Ebola (WHO PCR protocol) based on the symptom as rule-outs.

2.6.c Biological characterization of viruses identified: Novel viruses identified in humans will be recovered by cultivation or viruses re-derived from full length genome sequences using reverse genetic approaches. Viruses will be characterized for growth in primary human cell lines and in mouse models, including the Collaborative Cross Mice or humanized mice expressing human receptors from related strains. The approaches are discussed in more detail under Aim 1.

2.7 Epidemiological analysis: We will conduct a cross sectional case-control study to identify risk factors for Henipavirus, FVs, and CoVs spillover. "Cases" are defined as participants whose samples tested positive for Henipavirus, FVs, or CoVs by serological tests. "Controls" will be selected from the pool of participants admitted to the studies and whose samples tested negative for Henipavirus, FVs, or CoVs. We will use nearest neighbor matching to pair cases demographically (e.g. age, sex, location) with controls at a 1-to-3 ratio or greater. We will use generalized linear models and least absolute shrinkage and selection operator (LASSO) regression analyses (*176*) to analyze correlation between biological results, serological/PCR status, and risk factors related to high risk wildlife exposure including: 1) occupation and occupational exposures; 2) observed or reported interactions with priority wildlife, especially bats, in/around house; 3) home or work proximity to environments of increased risks, e.g. nearby bat roosts; 4) working or regular visitor to animal markets; 5) self-reported ILI/SARI clinical diagnosis or symptoms lifetime and past 12mo. Additionally, we will use this procedure to determine how clinical presentation differs between Henipavirus, FVs, or CoVs in exposed and unexposed participants, as well as in the time course of illness, severity of symptoms, and type of symptoms.

2.8 Potential problems/alternative approaches: Rarity of spillover events means it may be difficult to identify sufficient seropositives to statistically analyze human risk profiles. First, we will be targeting our community-based surveillance to subpopulations with high-levels of wildlife exposure, at sites selected for diverse and prevalent viruses to increase likelihood of finding positive individuals. Second, our serology testing will include a multi-plex panel of assays for a large diversity of viruses, increasing likelihood of positives. For example, in previous work our overall bat CoV PCR prevalence in the community was, 11.8%; β-CoV, 3.4%; α-CoV, 9.1%). Thus, using this broad serological panel to screen individuals prone to contact with species of higher zoonotic risk increases the potential for detecting spillover with enough power for statistical analyses,

and will shed light on behaviors that predispose to spillover. Third, we will include common human CoVs, PMVs, and influenza viruses in our panel, so that even if low prevalence of wildlife viruses is found, we will be able to conduct a valuable cross-sectional study of the seroprevalence of more common human viruses. Finally, we will be able to assess relative measures of human-wildlife contact from our survey work. We will analyze intensity of contact with high-risk taxa against other risk factors to provide useful proxy information for spillover risk. **Serological testing may be difficult to interpret in the presence of known and novel viruses.** We will use the three-tiered serological testing system outline in **2.6.a** to try to identify exposure to these 'novel' viruses. However, we know exposure to a range of related viruses generates antisera that are cross-reactive with known, related virus antigens - a serious challenge to serological surveillance and diagnostics. Our collaborations have already demonstrated that SE Asia bat sera samples screened with our pan-filovirus MMIA have been exposed to an Ebola-like virus, that is antigenically-distinct from known Asiatic filoviruses, Reston virus and Měnglà virus (*43*). Using these three serological platforms we will address inherent challenges of indirect virus surveillance through antibody capture by testing multiple binding assay platforms and antigens, using positive control samples to establish serological profiles against known viruses that will aid in our interpretations of cross-reactive antisera responses likely representative of 'novel' viruses.

Aim 3: Identify and characterize viral etiology of 'cryptic' outbreaks in clinical cohorts.

3.1 Rationale/Innovation: Our previous research **(Table 1)** and other data suggests that there are substantial numbers of outbreaks or clinical cases for which the etiologic agent is not diagnosed, e.g. (*177, 178*). To investigate this in SE. Asia, we have assembled a multi-disciplinary team that includes senior infectious disease doctors and clinicians and will work directly with local clinics and regional hospitals to identify the 'cryptic' outbreaks or cases that occur in the region. **In Aim 2, we will conduct serology in high zoonotic-risk communities to find evidence of prior exposure in people,** i.e. evidence of viral spillover into the community. In Aim 3 we identify members of these communities presenting at clinic with undiagnosed illness that may have high-impact zoonotic viral etiology. **Therefore, in Aim 3 we focus on PCR to identify the putative viruses that are replicating in these actively-infected patients (Fig. 12).** We will enroll, and collect high quality biological samples from patients who live in rural communities linked to Aim 2 and who present at regional clinics or hospitals with symptoms typical of high impact zoonotic viruses. We will collect detailed survey data to assess their contact type and frequency with wildlife and livestock and conduct molecular and serological diagnostic assays to test causal links between their syndromes and the known and novel viral agents identified in Aim 1. We will conduct molecular viral discovery assays on patients who present with symptoms indicative of a high-risk zoonotic virus. If novel pathogens are identified by PCR, we will attempt to isolate, obtain full genome sequence and/or biologically characterize the wildtype or molecularly re-derived pathogen, using the *in vitro* and *in vivo* strategy laid out in Aim 1.

3.2 Preliminary data clinical surveillance: 3.2.a. Clinical surveys and outbreak investigations: Our longterm collaboration in the region has included the following activities investigating clinical syndromes: *Peninsular Malaysia*: **At the time of writing, an outbreak of suspected undiagnosed viral illness in the indigenous Batek Orang Asli ("people of the forest") community living at Gua Musang ("civet cat cave") district, in Kelantan, Malaysia, close to Taman Negara – a National Park containing the oldest known rainforest.** This community is one where we have conducted prior routine surveillance and biological sampling, demonstrating proof-of-concept that our two-pronged high-risk community/clinic approach will work. Dozens of people are affected and the outbreak is currently being investigated by our group in collaboration with the Malaysian NPHL. The Orang Asli continue to practice traditional subsistence hunting of wild animals (bats, rodents, primates). These communities are remotely located, in heavily forested areas, with limited access to medical services. This outbreak, and others that will continue to emerge just like it, underscore the critical need for our EIDRC - to improve the technological tools, training, and human capacity for rapid EID surveillance, diagnosis, and response at the frontlines of disease emergence in SE Asia. **Investigating this outbreak is a key priority if EID-SEARCH is funded.** *Sabah*: Co-I Kamruddin (BMHRC) is leading an asymptomatic leptospirosis study on sanitation workers, an ILI study in a Kota Kinabalu clinic and diarrheal study in children under 5 yrs. Co-Is William, Tan and others have conducted clinical observational studies to determine the etiology of central nervous system infections and acute undifferentiated febrile illness in Sabah.

19

Co-Is William and consultant Yeo have collaborated with Co-I Wang to enroll 90 adult and over 100 pediatric participants with central nervous system infections, negative tests for malaria and dengue from 3 district hospitals in Sabah to determine the various etiologies. While a proportion tested positive for rickettsial agents, the majority had no clear diagnosis. Co-Is William, Rajahram, and Yeo have conducted clinical studies of over 200 inpatients with acute febrile illness with negative tests for malaria and dengue from 3 district hospitals in Sabah to determine the various etiologies. While a high proportion tested positive for rickettsial agents, the majority had no clear diagnosis. _Sarawak:_ Co-I Siang (UNIMAS) is conducting a survey of high-risk female Kelabit indigenous people and has found 8% with HPV52 as the dominant strain. He is working on clinical cohorts of hand, foot and mouth disease and will recruit patients from that study in the current proposed work. Finally, Co-Is Siang, Kamruddin are conducting a long-term study of rotavirus and other viral etiologies of acute gastroenteritis in Kuching, Sarawak that will include swine herders and farmers. The Baric lab is a global leader in norovirus VLP diagnostic reagents to detect serologic evidence of norovirus gastroenteritis, as well as RT-PCR detection of novel noroviruses in humans, and has identified novel bat noroviruses, suggesting further opportunities for expansion of this work (_142, 179-181_). This work is particularly relevant in Malaysia because pig farms are shunned by the government and local villagers to they are placed far away from town centers, deep in the forest, and it is this scenario that led to the emergence of Nipah virus in 1998-9 (_6, 22_). _Thailand:_ Co-Is Hemachudha, Wacharapluesedee (TRC-EID, Chulalongkorn Univ. Hosp.) work in coordination with the Thailand Ministry of Health to investigate outbreak etiology, including the first case of imported MERS-CoV in Thailand, **for which we produced sequence confirmation within 24 hours from acquiring the specimen**. Three imported MERS cases have been detected in Thailand since 2015, and all cases were confirmed at the TRC-EID. We worked during the Zika outbreak to identify > 500 PCR-positive Zika patients with sequence confirmation. We have also worked on returning and suspected Ebola travelers, syndromic surveillance for SARI, encephalitis, Hand Foot and Mouth disease in children and viral diarrhea. Consultant Hickey (US CDC, Thailand), screened specimens collected between 1998 and 2002 from a large cohort of children from Kamphaeng Phet Province to examine Dengue virus and JE epidemiology (_182, 183_). Among the 784/2,574 convalescent sera henipavirus-specific antibodies were detected in 13 serum samples (1.7% seroprevalence), suggesting prior exposure to endemic henipaviruses or a related PMV in rural Thailand. _Singapore:_ Duke-NUS has worked with the Ministry of Health to investigate Zika cases (_184_), and other EIDs. With EHA and others, Duke-NUS has developed novel CoV assays for clinical surveillance, and for expedited analysis during outbreaks. Our pipeline of novel virus detection, reservoir host identification, serological assay development, and serological testing is as rapid as several months, with the development of novel serological assays from sequence data in 7 days (_17_). **For example, with the discovery of SADS-CoV, we identified and fully sequenced a new virus, screened archived wildlife specimens to identify the natural reservoir host species (_Rhinolophus_ spp. bats), developed LIPS serological assays, and tested human and swine samples to assess seroprevelance in less than 3 months (_17_).**

3.2.b Analysis of self-reported illness: We have developed a standardized approach for analyzing data from study participants on self-reported symptoms related to target illnesses: fever with cough and shortness of breath or difficulty breathing (severe acute respiratory illness - SARI); fever with muscle aches (fevers of unknown origin (FUO); and, cough or sore throat (influenza-like illness - ILI). For all of our prior human clinical surveillance in 20+ countries, including in Malaysia and Thailand, we used a least absolute shrinkage and selection operator (LASSO) regression to identify associations between participants who reported ILI and/or
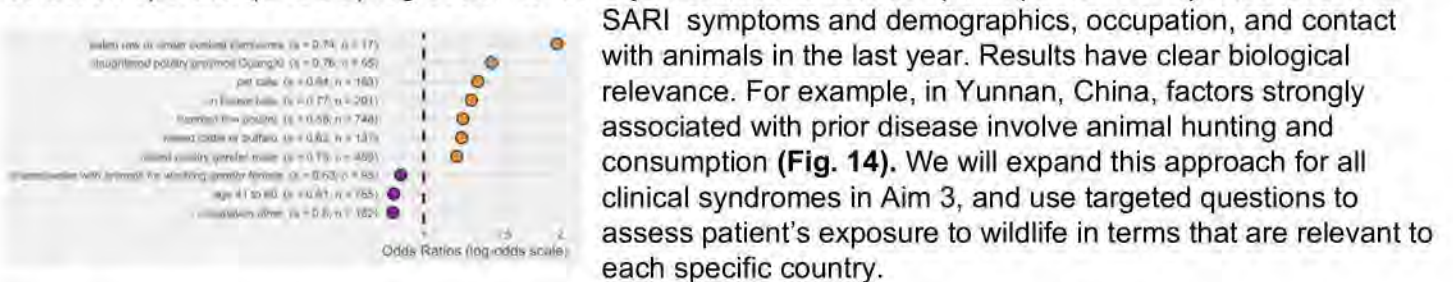

SARI symptoms and demographics, occupation, and contact with animals in the last year. Results have clear biological relevance. For example, in Yunnan, China, factors strongly associated with prior disease involve animal hunting and consumption **(Fig. 14)**. We will expand this approach for all clinical syndromes in Aim 3, and use targeted questions to assess patient's exposure to wildlife in terms that are relevant to each specific country.

Fig. 14: Factors associated with self-reported ILI & SARI in prior 12 months (s = bootstrap support; n = #+ve out of 1,585 respondents). **Orange circles** = odds ratios > 1 (+ve association); **purple** = odds ratios <1 (-ve association).

20

3.3 General Approach: Our previous research suggests that there are substantial numbers of outbreak cases for which the etiologic agent has not been identified. To capture these, we will expand on our current syndromic surveillance to enroll, and collect high quality biological samples from patients who present at regional clinics and live in rural communities linked to Aim 2. Case definitions and enrollment criteria include: 1) severe/acute respiratory illness (SARI/ARI); 2) Influenza-like illness (ILI); 3) fever of unknown origin (FUO); 4) encephalitis; 5) hemorrhagic fever; or 6) high fever/diarrhea with unusual presentation. We will collect survey data to assess contact with wildlife, and occupational and environmental exposures, conduct molecular and serological assays to test causal links between their syndromes and the known and novel viral agents identified in Aim 1. If novel pathogens are identified by PCR, we will attempt to isolate and biologically characterize the pathogen, using the collaborative cross to identify an appropriate animal model to conduct preliminary pathogenesis work. We will then use behavioral survey data, ecological and phylogeographic analyses to assess the likely reservoir hosts to inform intervention programs.

3.4 Clinical cohorts. 3.4.a Cohort selection: We will initiate active clinic-based syndromic surveillance at 2 town-level level clinics and 2 provincial-level hospitals in each Thailand, Peninsular Malaysia, Sabah and Sarawak. These serve as the catchment healthcare facilities for people in our community-based surveillance of Aim 2 and are inclusive of wildlife sampling sites in Aim 1. Patients $\geq$ 12 years old presenting at the health facility who meet the syndromic and clinical case definitions above will be recruited into the study after completing a consent form. We will aim to enroll 300 participants per hospital, with that we will have a 95% probability of detecting live virus in patients in each hospital assuming a population prevalence of 1%. This sampling effort will total of 4,800 individuals for clinical syndromic surveillance. We assume up to 40% loss from follow-up, therefore yielding 2880 participants that will be available for follow-up blood sampling **(Section 3.4.b)**. While enrollment is limited by the number of patients that come to the healthcare facility presenting with relevant clinical symptoms, in our work in the PREDICT project we enrolled an average of 100 participants per year, so we are confident this is an achievable enrollment target. Study data will be pooled across country regions, as clinical patients are limited by the number of individuals presenting at hospitals. "Cases" will be determined as participants whose samples tested positive for either CoVs, henipavirus, or FVs, PCR tests. "Controls" will be selected from the pool of participants admitted to the studies and whose samples tested negative. We will use nearest neighbor matching to pair cases demographically (e.g. age, sex, location) with controls at a 1-to-3 ratio or greater. Sites: _Thailand:_ We will work with two primary hospital, Phanat Nikhom Hosp., Chonburi district (Key Pers. Dr. P. Hemachudha) & Photharam Hosp., Ratchaburi Province, and a large tertiary hospital, King Chulalongkorn Memorial Hospital, Bangkok Thailand (Co-I Dr. T. Hemachudha). _Peninsular Malaysia:_ **Focusing immediately on the current outbreak in Orang Asli, we will work closely with the Malaysian Ministry of Health to support their investigation, and with 6 Orang Asli clinics which focus solely on this indigenous community.** Co-I Sellaran (Lintang Clinic, Kuala Kangsar) refers patients to Sungai Siput Hospital and Gen. Hospital, Ipoh. Key Pers. I Lotfi (Pos Betau clinic, Kuala Lipis) will continue collaboration with Co-I Hughes in the communities at that site. _Sarawak:_ Key Pers. Diyana (Director, Bario Clinic) refers patients to Miri Hospital (Co-I Utap) and serves people from the Kelabit, and nomadic tribes in the region, as well as populations within a number of large towns. _Sabah:_ We will continue our collaboration with Queen Elizabeth Hospital, QEH (Key Pers. Lee, Rajahram) and Hospital University Malaysia Sabah, HUMS (Co-I Lasimbang, Director). Both clinics include our targeted sites for **Aim 2** in their catchment, and both have ongoing collaboration with the Borneo Medical Health Ctr (Co-I Kamruddin, Director). _Singapore:_ Currently, there are no plans to conduct clinic surveillance based on budget constraints. However, in the event that that is an evolving need for clinic surveillance, enrollment, and sampling, we have close working relationships with all hospitals in the country, and the central infectious disease reference lab (key clinicians at which are aware of our proposal). This would allow us to rapidly on-board a hospital based site.

3.4.b Clinic enrollment and follow-up: We will recruit inpatients and outpatients to participate in the study after initial screening to determine if they meet our clinical case definitions. Once consented and enrolled, biological samples will be collected by trained and locally certified hospital staff and a questionnaire administered by trained hospital or research staff that speak appropriate local language. Every effort will be made to take samples concurrently when collecting samples for normative diagnostics. For both inpatients and outpatients, samples will be collected no more than 10 days of reported onset of illness to increase the chance

21

of PCR detection of viruses (*185*). Where possible, we will follow up 35 days after enrollment to collect an additional two 500 µL serum samples conduct a standardized questionnaire supplement to collect additional data on the course of symptoms in the interim period. This gives adequate time for development of IgG, which occurs <28 days after onset of symptoms for SARS patients (*186-188*). Serum samples will be used to screen panels of high risk human and local zoonotic Filovirus, Henipavirus and Coronavirus strains as described in Aim 2. These samples may prove valuable to collaborators in that the PBMC could be used in production of therapeutic antibodies.

3.4.c Sampling and clinical interview: Biological specimens whole blood (5mL will be drawn for whole blood and serum collection, samples will be aliquoted, at least one max. 500 µL whole blood; two 500 µL serum samples) and two nasal or oropharyngeal swabs will be collected from all participants, and a questionnaire will be administered. We will investigate five risk factors, to maximize the power of the analyses, all related to high risk wildlife exposure, based on continuing analysis of our previous work, and will include: 11) occupation and occupational exposures; 2) observed or reported interactions with priority wildlife, especially bats, in/around house; 3) proximity of residence or work place to environments of increased risks, e.g. nearby bat roosts; 4) working or regular visitor to animal markets; 5) self-reported ILI/SARI clinical diagnosis or symptoms in the past 12mo and lifetime. With additional consent considerations for participants enrolled from clinical settings, we will review clinical records to collect data on medical history, clinical syndromes, and patient etiology. Blood will be taken for future harvesting of Peripheral Blood Mononuclear Cells (PBMCs), as per **Section 2.5**.

3.5 Sample testing: The standard syndromic diagnostic PCR assays for common pathogens will be conducted and the results will be shared with the hospital with 48 hours. Specimens will be further tested for 3 viral families; CoV, PMV and filovirus by consensus PCR **(Section 1.4.c)**. Necropsy samples from deceased patients will be further characterized by NGS if previous PCRs are negative and unable to identify the cause of infection. Co-Is Wang, Anderson will conduct a VirScan serological assay using paired serum samples at Duke-NUS that identifies antibodies with a four-fold rise in titer in the convalescent sample compared with the acute sample, thereby indicating a possible etiological agent. A PCR test with the specific primer from each virus antibody positive case will be further tested from acute specimens to confirm infection.

3.6 Viral risk characterization and potential for pandemic spread: We will use statistical models built from collated biological, ecological, and genetic data to assess the likelihood of pathogenicity for the range of diverse viruses we will characterize. The aim is to quickly identify and assess which viruses pose the most significant risk to human health and thus prioritize additional resources to more fully characterize, better understand pathogenicity, and invest in our animal model pipeline and ultimately therapeutic development for those pathogens. We will combine two independent but related statistical models to predict pathogen spillover risk and pathogenicity. First, we will expand a simple generalized additive modeling approach (*3*) to assess spillover risk probability to a new virus that is currently not known to infect people, when we have only limited genetic data from initial PCR screening and sequencing. Additional data from receptor binding protein homology and cell line infection experiments will be incorporated into the model framework as available. Second, given high likelihood to spillover and infect humans, or direct evidence of human infection from our clinical surveillance, we will then assess the likelihood of pathogenicity based on traits for phylogenetically related nearest neighbor viruses as a first approximation and incorporating more detailed data from genomic characterization and animal model experiments (Aim 1). Epidemiological trait data for ~300 viral species known to infect people (e.g. case fatality rate, global estimates of number of human cases per year or decade, R_0, infectious period, and primary symptoms) have been collated by recent studies and will be used as

predictor variables (*189*). Thirdly, we will apply tools already developed, and being refined by EHA under DTRA and DHS supported research, to predict pandemic spread for viruses using global flight models, as well as additional datasets on human movement and connectivity across Southeast Asia (*90, 91*) **(Fig. 15)**.

Fig. 15: Probability feed from EHA's Flight Risk Tracker tool (FLIRT). This image is of the likely pathways of spread and establishment for a Nipah-like virus emerging from Sabah, based on human movement and airline flight data (*90, 91*).

22

3.7 Potential problems/alternative approaches: Patients visiting clinics may have cleared virus, but not yet developed IgG antibodies, reducing seropositive cases. Our 35 day follow-up sampling should account for this because the maximum lag time between SARS infection and IgG development is ~28 days and Ebola virus infection and IgG approximately 20 days (*185-188, 190*). We also expect that patients in rural communities will only visit clinics when symptoms have progressed, likely coinciding with late illness and onset of IgG. We will also have data from our community study that can be analyzed in conjunction with, or in the absence of, clinical data from the hospital data to identify PCR- or seropositives. Finally, the risk of limited detection is outweighed by the potential public health importance of discovering active spillover of new henipaviruses, FVs or CoVs infection.

4. Administrative Plan

4.1. Project management: 4.1.a. Administrative core: The EID-SEARCH organizational partner roles are laid out by Specific Aim above **(see Fig. 3)**. EcoHealth Alliance (EHA) will lead the administrative core of the EID-SEARCH, including: coordination and management of all human and animal research; communication with consortium partners, NIH, other EIDRCs, and the EIDRC-CC; data management and quality assurance; modeling and analysis; and compliance with all US, international, and institutional rules and regulations. The administration will be led by PI Daszak, who has >20 years' experience managing international research collaborations on emerging zoonoses, and has collaborated with all partners in the EIDRC consortium for 5-20 yrs on NIAID- and USAID-funded research **(Fig. 16)**. He also has direct experience as one of two external reviewers of one of the CEIRS partners for NIAID (Program officer Diane Post). PI Daszak will be assisted by Deputy Lead, co-I Olival, who has overseen zoonotic disease research in Southeast Asia for >10 years and working directly with our partners in Thailand and Malaysia for 15 years. Senior program manager (Chmura)

and full-time program assistant (Hamel) will provide administrative support. The EID-SEARCH data management, epidemiological analysis, and modeling activities will be led by co-Is Olival and Zambrana who have led all modeling and analytics on USAID USAID PREDICT for the last 10 years. They will overse a programmer/data scientist (Mendelsohn), evolutionary biologist/bioinformatician (Latinne), and human epidemiologists (Dawson, Li, Martinez). Our human epidemiology team will work directly with country teams in Malaysia and Thailand to manage human surveillance, IRB applications and approvals, and ensure compliance with all human subjects work. EHA's senior veterinarian (Field) and wildlife field biologist (Phelps) have extensive experience working that bats, rodents, and primates in SE Asia and will coordinate all training and field research in Malaysia and Thailand and ensure compliance with IACUCs.
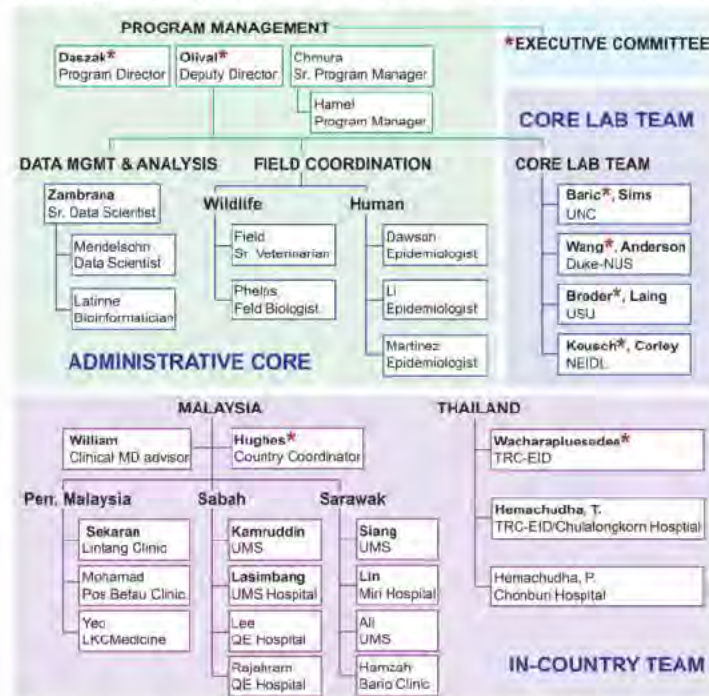
Fig. 16: Administration and program management for the **EID-SEARCH**

The EID-SEARCH core laboratory team includes global experts in virology, molecular pathogenesis, and the development of assays, reagents, and therapeutics for high consequence viral zoonoses (Co-Is Wang, Anderson at Duke-NUS; Baric, Sims at UNC; Broder, Laing at USU; Keusch, Corley at NEIDL). The core laboratory team will work with in-country diagnostic labs in Thailand (Wacharapluesedee, Hemachudha at the TRC-EID, Chulalongkorn University) and Malaysia (Hughes, Lee, CM Ltd.). Wacharapluesedee and Hughes will additionally serve as country coordinators for the EID-SEARCH, liaising directly with EHA on weekly conference calls, and working with in-country partners to ensure the smooth operation and coordination of clinical and community surveillance and wildlife research in Thailand and Malaysia, respectively. **A Core Executive Committee** comprising PI Daszak, Deputy Lead and Co-I Olival, and leads from each core partner and country: Co-Is Hughes, Wacharapleusadee, Baric, Wang, Broder, Keusch (or alternates), will conduct

23