Table 1: Species found PCR-positive for SARSr-CoVs in our R01, with sample sizes and prevalence estimates.

Bat Species	Individuals tested	# positive	SARSr-CoV mean prev.	SARSr-CoV prev. range
Rhinolophus sinicus	1,328	113	8.5%	7.1 – 10.1%
R. macrotis	70	3	4.3%	0.9 - 12%
R. ferrumequinum	406	12	3.0%	1.5 – 5.1%
R. spp.	331	10	3.0%	1.5 – 5.5%
R. affinis	792	7	0.9%	0.4 – 1.8%
R. pusillus	1,023	8	0.8%	0.3 – 1.5%
Aselliscus stoliczkanus	269	2	0.7%	0.1 – 2.7%
Hipposideros pratti	323	2	0.6%	0.1 – 2.2%
H. armiger	1,188	1	0.1%	0.0 – 0.5%

We identified one cave system (the "Jinning Cave") in Yunnan Province that harbors *Rhinolophus* spp. bats with diverse SARSr-CoVs, including some with S proteins able to use human ACE2 as entry receptors. Bats in this cave carried SARSr-CoVs with **all unique genetic elements of the SARS-CoV outbreak virus**, suggesting that this site may be a potential public health risk (*29*).

We used a novel phylogeographic analysis, Maximum Clade Credibility (MCC) tree, to reconstruct the geographic areas of evolutionary origin for β-CoVs that we sequenced. Results suggest that: 1) Guangdong Province is the ancestral center of diversity of β-CoVs (data not shown); 2) Guizhou is the likely origin of the HKU3-related clade (lineage 1); and 3) Guangdong and Guangxi are the likely ancestral origins of the SARS-CoV outbreak sequences (lineage 2) **(Fig. 3)**. Despite our intensive sampling at some sites, around half of the 20 *Rhinolophus* spp. we identified were captured at sample sizes below the minimum required to detect SARSr-CoVs at prevalences we found (n=110, power 80%), and 5 others were SARSr-CoV negative in our study. To estimate sampling gaps, we used a viral 'mark-recapture' approach we previously published (*36, 37*). Results suggest we are approaching saturation of CoV strain discovery at some sites, whereas other sites contain rich pools of SARSr-CoVs that remain undiscovered **(Fig. 4)**. **In the current proposal,** we have used these analyses to estimate geographic and species-specific sampling targets **to more effectively identify new strains and CoV lineages needed to support experimental infection studies and risk assessment.**

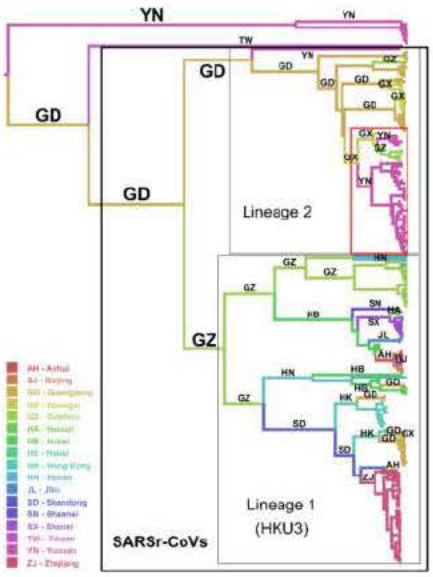

Fig.3 (left): MCC phylogeny of lineage B β-CoVs, including SARSr-CoVs (black box). Lineage 1 includes HKU3-related CoVs, lineage 2 includes SARS-CoV outbreak strains and close relatives (red box). Branches colored according to province of inferred ancestral origin (Guangdong GD, Yunnan YN, Guizhou, GZ).

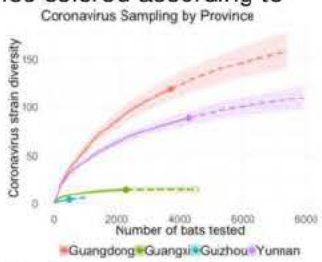

Fig. 4 (right): Estimates of SARSr-CoV strain diversity in the bats we sampled (strain defined as >10% sequence divergence in RdRp gene). GD and YN harbor highest CoV diversity, but discovery has not yet saturated. We estimate proposed additional sampling of 5,000 bats will identify >80% of remaining β-CoV strains in bat hosts from these regions.

In vitro & in vivo *characterization of SARSr-CoV potential for human infection*

We conducted *in vitro* and *in vivo* experiments to characterize the pathogenic potential of novel SARSr-CoVs. We isolated three SARSr-CoVs from bat feces: WIV1, WIV16 and Rs4874, with S protein sequences that diverged from SARS-CoV by 3% to 7% (*17, 22, 29*). We conducted full-length genome sequencing of 12 other novel SARSr-CoVs from the Jinning Cave, some highly similar to outbreak SARS-CoV in the most variable genes: N-terminal domain and receptor binding domain (RBD) of the S gene, ORF8 and ORF3 (*29*). Using our reverse genetics system, we constructed chimeric viruses with SARSr-CoV WIV1 backbone and the S gene of different variants, including WIV1-Rs4231S and WIV1-Rs7327S. All 3 SARSr-CoV isolates and the two chimeric viruses replicated efficiently in Vero E6 cells and in HeLa cells expressing hACE2, but not in HeLa cells that don't express ACE2 (*17, 22, 29*) **(Fig. 5a)**. In collaboration with Ralph Baric (UNC), we used the SARS-CoV reverse genetics system (*38*) to generate a chimeric virus with a mouse-adapted SARS-CoV