**Working report 1: stress sub-optimal sequence coverage**

Run simulation with 5000 individuals and 8000 individuals, per each run consider different cases defined below:

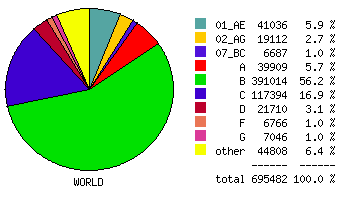
# **Scenario 1**

* one subtype of the virus (HIV-1-B) for all seeds
* complete sampling for a transmission network of one seed
* same sampling time interval (e.g.: five or three years) for a transmission network of one seed

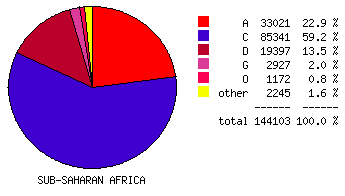
# **Scenario 2**

* one subtype of the virus (HIV-1-B) for all seeds
* complete sampling for a transmission network of all seeds
* same sampling time interval (e.g.: five or three years) for a transmission network of all seeds

Current HIV-1 subtypes variants found in the World and in Sub-Saharan Africa: <https://www.hiv.lanl.gov/components/sequence/HIV/geo/geo.comp>



*Figure 1: Pie slice of frequency of subtypes of HIV-1 in the World*



*Figure 2: Pie slice of frequency of subtypes of HIV-1 in Sub-Saharan Africa*

We choose: subtype **A**, **C**, **D**, and **G**.

* **C**: full length 279-294 (<https://www.hiv.lanl.gov/components/sequence/HIV/search/search.comp>) 27/10/2017 [remove 282 & 283 double 281 / remove 285 double 284 / remove 288-281, double 292] >> **9 sequences retained**
* **A**: full length 4-11 & 15 ([https://www.hiv.lanl.gov/components/sequence/HIV/search/d\_search.comp?ssam\_subtype=A%20OR%20A1%20OR%20A2&ssam\_organism=HIV-1&ssam\_sample\_georegion=ssa&ssam\_sample\_country=[A-Z](https://www.hiv.lanl.gov/components/sequence/HIV/search/d_search.comp?ssam_subtype=A OR A1 OR A2&ssam_organism=HIV-1&ssam_sample_georegion=ssa&ssam_sample_country=[A-Z)])
* **D**: not full length (+8k), 671-677 & 6,9 (<https://www.hiv.lanl.gov/components/sequence/HIV/search/search.comp>)
* **G**: full length (+8k), 2, 6, 13, 583, 780, 894-897 (<https://www.hiv.lanl.gov/components/sequence/HIV/search/search.comp>)

# **Scenario 3**

* different subtypes of the virus (**HIV-1-B,C,D,A**) for all seeds
* complete sampling for a transmission network of one seed
* same sampling time interval (e.g.: five or three years) for a transmission network of one seed

# **Scenario 4**

* different subtypes of the virus (**HIV-1-B,C,D,A**) for all seeds
* complete sampling for a transmission network of all seeds
* same sampling time interval (e.g.: five or three years) for a transmission network of all seeds