Nanopore vs Illumina metadata

The metadata model is conserved.

In general in Nanopore metadata files all the same attributes as in Illumina are filled.

Experiment design>
Experiment design use the same metadata tags and value formats (see experiment.grid_nanopore.pdf for the distribution of values of experiment design from sequencing exp in the 329 files)

Sample attributes>
Nano pore metadata files have 13 sample attributes (vs 31 in Illumina files) but most of the useful ones (those in proposed schema) are common between them.
Here are the list of them: (and their mapping in schema)
10 are in common with illumina xmls: (schema ones are marked)

- "strain"   (➜ VIRUS strain)

   "isolate"

   "collected_by"

- "collection_date"   (➜ BIOSAMPLE collection_date)
- "geo_loc_name"   (➜ HOST geo_origin)
- "host"   (➜ HOST taxon_id)
- "host_disease"   (➜ HOST disease)
- "isolation_source"   (➜ BIOSAMPLE sample_type)

   "BioSampleModel"

   3 are new ones: (idk what these are)

   "barcode_identifiers"

   "barcode identifiers"

   "ARTIC barcode identifiers"

WHAT IS LACKING (from attributes in schema that do are in Illumina files)>
- [ ] age (➜ HOST age)
- [ ] sex (➜ HOST sex)
- [ ] tissue (but there is isolation_source, and also tissue is not very used in Illumina either (2/50) although non/redundant) (➜ BIOSAMPLE sample_type)
- [ ] laboratory_host (cell line when cell culture is source), anyway could be too much detail and thus not needed in schema (➜ BIOSAMPLE   culture_cell)
- [ ] passage_history (cell line passage details when cell culture is source), anyway could be too

much detail and thus not needed in schema (➜ BIOSAMPLE   culture_passage_history)

☐ link_addit_analys (but there is a new attribute, study_links, *within STUDY with the same info) ➜ (VARIANT IN SAMPLE study_info:study_ref or HANDOVER) e.g "https://submit.ncbi.nlm.nih.gov/subs/sra/SUB7181455/bp_info"

*study links

xml$EXPERIMENT_PACKAGE_SET$EXPERIMENT_PACKAGE$STUDY$STUDY_LINKS $STUDY_LINK$URL_LINK$URL1