0331

A New Public Database for Near Full-Length HIV Single-Genome Sequences (https://pvsdb.ncifcrf.gov)

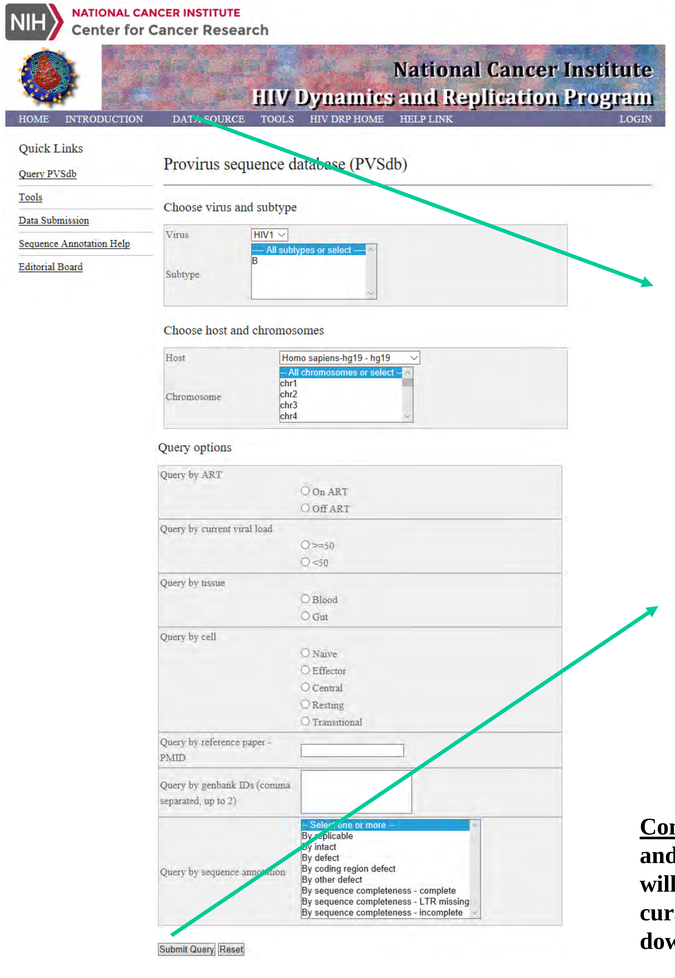
Wei Shao[1], Jigui Shan[1] ,Wei-Shau Hu[2], Elias K. Havas[3], Brian Luke[1], John W. Mellors[3], John M. Coffin[4], Mary F. Kearney[2]

[1]Frederick National Laboratory for Cancer Research, MD, USA; [2]National Cancer Institute, NIH., MD, USA; [3]University of Pittsburgh, PA, USA; [4]Tufts University, MA, USA

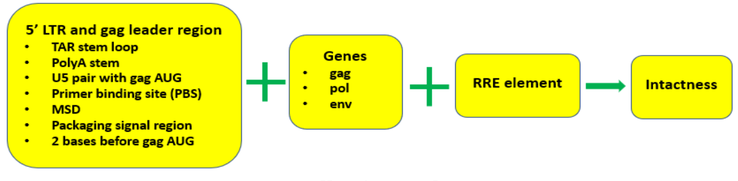
HIV Dynamics and Replication Program

Introduction: Despite the success of ART, HIV persists in reservoirs and viremia rebounds if treatment is interrupted. To facilitate understanding of the genetic structure and dynamics of the HIV reservoir, we developed a public database, PVSdb, for the storage and meta-analyses of near full-length (NFL) HIV genomic RNA and proviral sequences that persist in donors on ART or that rebound after ART is interrupted. PVSdb also has tools to annotate NFL sequences.

1. Provirus Sequence Database Web Page



2. NFL Sequence Annotation and Intactness Determination



3. Data Collection and Storage

s - Show sample information of the paper

Pubmed_ID	Title	Author	Journal	Year	Comments
27500724	Defective proviruses rapidly accumulate during acute HIV-1 infection	Bruner, KM et al.	Nat Med	2017	
29045846	Identification of Genetically Intact HIV-1 Proviruses in Specific CD4+ T Cells from Effectively Treated Participants	Hiene, B et al.	Cell Reports	2017	
24243014	Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure.	Ho, YC et al.	Cell	2013	

4. Basic Query: All Stored Sequences and their annotation

Total sequences found 1056, Download sequence file , Download annotation file

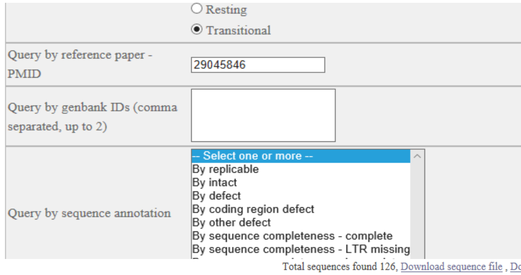
1	2	3	4	5	6	7	8	9
pubmed_id	genbank_id	seq_length	5LTR_length	TAR_stem_loop_insertion	TAR_stem_loop_deletion	polyA_insertion	polyA_deletion	u5_gag_pair_R1_insertion
27500724	KX505733.1	638	638					
27500724	KX505627.1	2748	5		61		6	

10	11	12	13	14	15	16	17
u5_gag_pair_R1_deletion	first_half_pbs_insertion	first_half_pbs_deletion	MSD_status	packing_insertion	packing_deletion	gag_AUG_pair_U5_insertion	gag_AUG_pair_U5_deletion
		9	missing		116		11
11		9	missing		116		11

18	19	20	21	22	23	24	25	26	27
2base_before_gag_deletion	gag_start_ATG	gag_insertion	gag_deletion	gag_premature_stop	pol_insertion	pol_deletion	pol_premature_stop	env_insertion	env_deletion
deletion	deletion		1503						2571
deletion	deletion		1503					6	65

28	29	30	31	32	33	34	35
env_premature_stop	rre_status	3LTR_length	two_LTR_similarity	intactness	sequence_completeness	replicable	annot_comment
	bad		0.00	defective	LTR incomplete		NA
70	good		0.00	defective	LTR incomplete		NA

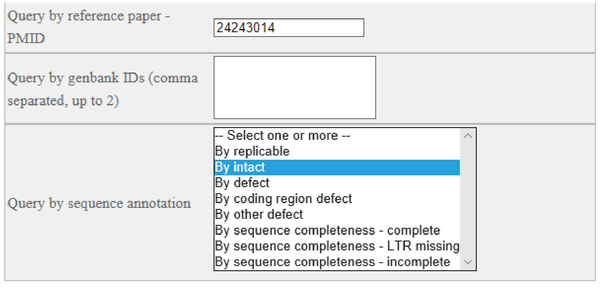
5. Advanced Query: Sequences from a Cell type Reported in a Paper



Total sequences found 126, Download sequence file , Download annotation file

1	2	3	4	5	6	7	8	9
pubmed_id	genbank_id	seq_length	5LTR_length	TAR_stem_loop_insertion	TAR_stem_loop_deletion	polyA_insertion	polyA_deletion	u5_gag_pair_R1_insertion
29045846	KY778360.1	4437			61		6	
29045846	KY778424.1	8991			61		6	
29045846	KY778427.1	1262			61		6	

6. Advanced Query: Intact Sequences Reported in a Paper



29	30	31	32	33	34	35
rre_status	3LTR_length	two_LTR_similarity	intactness	sequence_completeness	replicable	annot_comment
good	565	0.00	intact	LTR incomplete		NA
good	565	0.00	intact	LTR incomplete		NA
good	519	0.00	intact	LTR incomplete		NA
good	505	0.00	intact	LTR incomplete		NA
good	519	0.00	intact	LTR incomplete		NA
good	565	0.00	intact	LTR incomplete		NA
good	578	0.00	intact	LTR incomplete		NA
good	565	0.00	intact	LTR incomplete	replicable	NA
good	565	0.00	intact	LTR incomplete		NA
good	519	0.00	intact	LTR incomplete		NA
good	565	0.00	intact	LTR incomplete	replicable	NA

Conclusions: This public HIV NFL sequence database and accompanying tools provide easy querying and analyses of HIV DNA and RNA genomes that persist on ART. Meta-analyses of HIV NFL genetics will contribute to our understanding of HIV persistence and may reveal targets for potential future curative interventions. The query results including NFL sequences and their annotation files can be downloaded.

Frederick National Laboratory for Cancer Research
sponsored by the National Cancer Institute