HTG data

Input: Annotation file

Entry	Feature	Location	Qualifier	Value
COMMON	DATE		hold_date	20191130
	SUBMITTER		contact	Hanako Mishima
			ab_name	Mishima,H.
			ab_name	Yamada,T.
			ab_name	Park,C.S.
			ab_name	Liu,G.Q.
			email	mishima@ddbj.nig.ac.jp
			phone	81-55-981-6853
			fax	81-55-981-6849
			institute	National Institute of Genetics
			department	DNA Data Bank of Japan
			country	Japan
			state	Shizuoka
			city	Mishima
			street	Yata 1111
			zip	411-8540
	REFERENCE		ab_name	Mishima,H.
			ab_name	Yamada,T.
			ab_name	Park,C.S.
			ab_name	Liu,G.Q.
			title	Arabidopsis thaliana DNA
			year	2017
			status	Unpublished
			email	mishima@ddbj.nig.ac.jp
			phone	81-55-981-6853
			fax	81-55-981-6849
			phext	3207
	DIVISION		division	HTG
	KEYWORD		keyword	HTG
			keyword	HTGS_PHASE1
	COMMENT		line	Please visit our web site
			line	URL: http://ddbj.nig.ac.jp
ENT001	source	1..E	clone	CIC5D1
			clone_lib	AT01 BAC
			organism	Arabidopsis thaliana
			mol_type	genomic DNA
			chromosome	1
			ecotype	columbia
			ff_definition	@@[organism]@@ DNA, BAC clone: @@[clone]@@, chromosome @@[chromosome]@@, *** SEQUENCING IN PROGRESS ***
	assembly_gap	5001..5100	estimated_length	unknown
			gap_type	within scaffold
			linkage_evidence	within clone
	assembly_gap	24001..24100	estimated_length	unknown
			gap_type	within scaffold
			linkage_evidence	within clone
	assembly_gap	62001..62100	estimated_length	unknown
			gap_type	within scaffold
			linkage_evidence	within clone
	assembly_gap	105001..105100	estimated_length	unknown
			gap_type	within scaffold
			linkage_evidence	within clone

Output in DDBJ flat file

```
LOCUS       ############          123456 bp    DNA     linear   HTG 30-SEP-2017
DEFINITION  Arabidopsis thaliana DNA, BAC clone: CIC5D1, chromosome 1, ***
            SEQUENCING IN PROGRESS ***
ACCESSION   ############
VERSION     ############.1
DBLINK
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; malvids; Brassicales; Brassicaceae; Camelineae;
            Arabidopsis.
REFERENCE   1  (bases 1 to 123456)
  AUTHORS   Mishima,H., ada,T., Park,C.S. and Liu,G.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (dd-mmm-yyyy) to the DDBJ/EMBL/GenBank databases.
            Contact:Hanako Mishima
            National Institute of Genetics, DNA Data Bank of Japan; Yata 1111,
            Mishima, Shizuoka 411-8540, Japan
REFERENCE   2
  AUTHORS   Mishima,H., ada,T., Park,C.S. and Liu,G.Q.
  TITLE     Arabidopsis thaliana Sequencing
  JOURNAL   Unpublished (2017)
COMMENT     Please visit our web site
            URL:http://www.ddbj.nig.ac.jp/
FEATURES             Location/Qualifiers
     source          1..123456
                     /chromosome="1"
                     /clone="CIC6D5"
                     /clone_lib="AT01 BAC"
                     /ecotype="columbia"
                     /mol_type="genomic DNA"
                     /organism="Arabidopsis thaliana"
     assembly_gap    5001..5100
                     /estimated_length="unknown"
                     /gap_type="within scaffold"
                     /linkage_evidence="within clone"
     assembly_gap    24001..24100
                     /estimated_length="unknown"
                     /gap_type="within scaffold"
                     /linkage_evidence="within clone"
     assembly_gap    62001..62100
                     /estimated_length="unknown"
                     /gap_type="within scaffold"
                     /linkage_evidence="within clone"
     assembly_gap    105001..105100
                     /estimated_length="unknown"
                     /gap_type="within scaffold"
                     /linkage_evidence="within clone"
<The lines for BASE COUNT, ORIGIN, and nucleotide sequence are omitted>
//
```

DDBJ
DNA Data Bank of Japan