

## Sample Annotation File: WGS (Whole Genome Shotgun) is used as piece of Scaffold CON entry

This line is not required for annotation file.

Entry	Feature	Location	Qualifier	Value
COMMON	DATATYPE		type	WGS
	KEYWORD		keyword	WGS
			keyword	NON_CONTIGUOUS_FINISHED
	DBLINK		project	PRJDB12345
			biosample	SAMD90000001
			sequence read archive	DRR990000
	SUBMITTER		contact	Hanako Mishima
			consrtm	Zimmermannella faecalis genome sequencing consortium
			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 Gene
			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	Kim,K.-H.
			ab_name	Park,C.S.
			ab_name	Liu,G.Q.
			title	Mouse whole genome shotgun seq
			status	Unpublished
			year	2017
			hold_date	20191130
	DATE		line	Please visit our web page
	COMMENT		line	URL: http://www.ddbj.nig.ac.jp/
	ST_COMMENT		tagset_id	Genome-Assembly-Data
			Assembly Method	Velev v. 2.0
			Assembly Name	Mmus_assembly01
			Genome Coverage	60x
			Sequencing Technology	454 GS FLX; ABI 3730
	source	1..E	organism	Mus musculus
			mol_type	genomic DNA
			strain	15706
			ff_definition	@@[organism]@@ @@[strain]@@ DNA, @@[submitter_seqid]@@
			submitter_seqid	@@[entry]@@

Enter "WGS" on DATATYPE and KEYWORD lines.

Use consrtm qualifier for group name.

Enter two qualifiers, "project" for BioProject ID and "biosample" for BioSample ID, on DBLINK line, respectively. When the sequence is assembled by using short reads which were output from NGS, accession number of the Run Data issued by SRA database should be described in "sequence read archive".

Enter the full name for contact qualifier.

Enter the date to withhold your data. If the data should be publicized immediately, delete the

In case of submission, entry name is not necessary. Header description for each entry in sequence file is reflected as contig name (value of /submitter\_seq\_id) on a publicized flat file.

Enter the ST\_COMMENT by which depends on the submitted sequences. For eukaryotic genome entry(ies), "assembly name" qualifier is mandatory. Please see the Mass Submission System Making Files (COMMENT/ST\_COMMENT) for details.

Enter the contig name using submitter\_seqid qualifier. If an entry name represents value of submitter\_id qualifier, you can use metadescription, @@[entry]@@, to show sequence (contig) name, instead of concrete sequence name, respectively.

Meta-base position, "E", can be used to describe location of a feature instead of numerical number of sequence end point, if a given feature, including source, spans full length of a submitted sequence. Please visit a following site for detail, [http://www.ddbj.nig.ac.jp/sub/mss/file\\_format](http://www.ddbj.nig.ac.jp/sub/mss/file_format)

Meta description can be used to represent value for some qualifiers, such as organism, submitter\_seqid, in ff\_definition. Please visit the following site for detail. [http://www.ddbj.nig.ac.jp/sub/mss/file\\_format-e.html#annotation](http://www.ddbj.nig.ac.jp/sub/mss/file_format-e.html#annotation)

Source feature can be described on a submission which all information, feature, qualifier and its value are same for all