## eLMSG 数据提交

微生物系统和基因组学图书馆(An eLibrary of Microbial Systematics and Genomics, eLMSG, https://www.biosino.org/elmsg/index)是一个提供微生物典型物种(Type Species) 多相分类鉴定基础数据、基因组序列及高质量注释结果、种群核心与泛基因组信息的数据库 平台,该平台同时提供分离菌株和宏基因组拼接菌株的基因组数据提交。

目前平台提交的流程如下:



一、提交原始数据

目前 eLMSG 数据上传采用线下上传的方式进行,您可以把您要提交的基因组序列直接邮件给我们 eLMSG 工作组(邮箱:<u>elmsg-support@picb.ac.cn</u>, 电话:17710710719)。

二、创建元数据

1、分离菌株的元数据创建

	A	В	С	D	E	Н	I	J	K	L	P	S	Т
1	This is a submis	sion template for	depositing mic	roorganis	m dati	a to eLMSG (h	https://www.biosino.c	org/elmsg/index).					
2	Red fields	Mandatory	It is necessar	ry, if it is e	empty	your submissio	on will fail. If informa	ition is unavailable, p	please enter 'not coll	ected'.			
3	Blue fields	Recommeded ma	These fields a	are import	ant foi	r microorganisi	m, please fill in as m	nuch as possible. If r	no information is ava	ilable, keep it empty	1		
4	Yellow fields	Conditional manu	If no informat	tion is avai	ilable,	keep it empty.							
5		Auto-generated	These fields v	will be auto	omatic	ally generated	by eLMSG system.						
6	CV	controlled vocabl	You can cho	ose the va	lue fro	om the controlle	ed vocabullary. If yo	ur value is not in the	list of controlled voc	abullary, please cor	ntact us through elmsg	-support@pi	icb.ac.cn
7	You can add the	se fields, includir	g Ecology, M	orphology,	, Multi	cellular morpho	ology, Colony morph	ology, C/N source,	Metabolite, Toleranc	e, Antibiotic, GC, Er	nzymology, Genomics		
8	All punctuation r	narks are in Engli	sh, and after o	comma ha	s a bla	ank.							
9				2. 植今印			FLOOPADES			10-11-ch		洗埴字	
10			×	の項ナね		ANI				推存子段		X2094 J	
11							Ger	neral Information	1			1	
12						· /		16S	rRNA				1
13						1			Ц /			/	
14	Item	Scientific Name	Strain Name	Type	ANI	F The value of fa	stAMI identification for	the strain with the	opology	Create date	Genbank 16SrRNA	Etymology	Sample T
15	if CV			CV		existing specie	s name, for more detail	please check				- <u></u>	
16	Species 1	Allenteromonas intesti	Z73T	negative	98	Thermaceae	agagiiigaiccigacic	single	Inear	31-Dec-03	LC066536	en.o.shi.mer	seawater s
17													
18													

元数据的创建目前采用 excel 方式提交,具体的填写表格参照 "eLMSG\_v5.0\_purified\_strain.xlsx"。表格第14行是我们 eLMSG 收集的字段,包括物种 基本信息(Taxonomy information)、采样环境信息(Ecology)、形态学和理化表型信息 (Morphology、Physiology、Biochemistry、Enzymology)和基因组信息(Genomics) 这几大块信息。如果没有测微生物形态学和理化表型信息,填写物种基本信息、采样环 境信息和基因组信息即可。

其中, <mark>红色</mark>标记的字段 (R, mandatory fields) 为必填字段, <mark>蓝色</mark>标记的字段 (RM, recommended mandatory fields) 表示推荐必填字段, <mark>黄色</mark>标记的字段 (C, conditional mandatory fields) 为选填字段。把鼠标放在对应的字段上, 会有该字段的解释和示例数 据供大家参考。一个菌株填一行。



有些字段有控制词表,我们在第 15 行增加了红色高亮加粗的 CV (controlled vocabulary),控制词表里边的词少,直接点击填写的空格,可以从下拉菜单中选择(如 "控制词表一"所示),当控制词表里的词比较多,下拉菜单无法解决的时候,我们的控制词表会单独列一个清单(如"控制词表二"所示),大家可以去清单找自己需要填写的值。 大家在填写的时候可以直接从我们控制词表中选择对应的值,如果您的值不在我们的控制词表里边,请先空着并告知 eLMSG 工作组(邮箱 :elmsg-support@picb.ac.cn,电话 : 17710710719)进行控制词表修改。

2、MAG (metagenome-assembled genome)的元数据创建



元数据的创建目前采用 excel 方式提交, 具体的填写表格参照 "eLMSG\_v5.0\_MAG.xlsx" 和 样 本 息 信 表 "eLMSG\_v5.0\_MAG\_Environment\_Metagenomic.xlsx"。 "eLMSG v5.0 MAG.xlsx"表格第 14 行是我们 eLMSG 收集的字段,包括物种基本信息 (Taxonomy information)、基因组 信息 (Genomics)、样本名称 (Sample name) 和参考文献 (References) 这几大块信息。 茸 中 "eLMSG v5.0 MAG.xlsx" 表 的 " Sample name " 与 "eLMSG\_v5.0\_MAG\_Environment\_Metagenomic.xlsx"表的"Sample name"对应。

其中, <mark>红色</mark>标记的字段 (R, mandatory fields) 为必填字段, <mark>蓝色</mark>标记的字段 (RM, recommended mandatory fields) 表示推荐必填字段, <mark>黄色</mark>标记的字段 (C, conditional mandatory fields) 为选填字段。把鼠标放在对应的字段上, 会有该字段的解释和示例数 据供大家参考。一个菌株填一行。



有些字段有控制词表,我们在第 15 行增加了红色高亮加粗的 CV (controlled vocabulary),控制词表里边的词少,直接点击填写的空格,可以从下拉菜单中选择(如 "控制词表一"所示),当控制词表里的词比较多,下拉菜单无法解决的时候,我们的控制词表会单独列一个清单(如"控制词表二"所示),大家可以去清单找自己需要填写的值。 大家在填写的时候可以直接从我们控制词表中选择对应的值,如果您的值不在我们的控制词表里边,请先空着并告知 eLMSG 工作组(邮箱 :elmsg-support@picb.ac.cn,电话 : 17710710719)进行控制词表修改。

三、数据发布

在保证元数据都准确填写的情况下,我们会给出您提交的数据对应的基因组 ID 和数据链接。私有数据和公开数据会有两种不同的 ID 给出方式:

1、公开数据

直接给您连接, 点击对应的 "Download" 连接查看基因组数据。例如 https://www.biosino.org/elmsg/record/MSG050834

😍 e LMSG				Search
Antifaeciroseburia inte	estinalis str. D21-19T			
General information				
Name D21-19T	Rank	Strain		
Lineage Root. Bacteria.	Firmicutes. Bacilli. Lactobacillales. L	actobacillaceae. 'A	Antifaeciroseburi	'. 'Antifaeciroseburia intestinalis'. D21-19T
Nomenclature type pos	itive			
Ecology				
Isolation source	isolated from the cecal contents of	an 8-week-old male	Lepob/ob C57B	L/6J mouse.
Taxonomy marker				•
MK287676 Marvinbrya	ntia sp.			Download
🛗 Create date: 17-DEC-2018				
Length	1413	Topology	linear	Strandedness double
Genomics				
LMSG_G00000060.1	'Antifaeciroseburia intestinalis'			Download
🖺 Submitter date: 2019-12-30 1	4:40:35 Update date: 2019-12-30 14:40:35			
Source	Chang Liu, Institute of Microbiology	, Chinese Academy	of Science	点击Download按钮可以查看基因组数据

2、私有数据

私有数据我们会给您两种链接模式,一种链接是正常的网页连接,您可以通过 该连接查看您该基因组的元数据信息,但是基因组序列信息无法查看,例如 https://www.biosino.org/elmsg/record/MSG051763;另一种连接是审稿人的专用链 接,通过该链接,审稿人可以对私有数据进行查看,点击对应的"Download"连接查 看基因组数据。

General information					
Name YSL	Rank Strain				
Lineage Root. Archaea	. Euryarchaeota. Methanobacteria. Methanobacteriales. Methanobacteriaceae. Methanobacterium. 'Methanobacterium sp.'. YSL				
Nomenclature type positive					
Ecology					
lealation source	andimant				
Environment	seumen				
Coordination	loka Diras Vantai				
Geographic location	Jene Kwei, Tankai				
Conection date	2013/07				
Environment feature	harbor				
Environment reature	eadimont				
Latitude and	27 00222 27 CE0008 NI 400 42222 400 622228 E				
longitude	31.00333-31.3000 N IZU.13333-IZU.03333 E				
Genomics					
LMSG_G000000181.1	'Methanobacterium sp.'				
🗂 Submitter date: 2020/01/02					
Source	Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences 私有数据基因组数据无法查看				
Assembly status	complete genome Assembly class haploid 需诵讨特殊链接查看				

## 四、文章发表

在文章发表的时候, 审稿人可以通过您文章里边的链接或者基因组 ID 号对您的数 据进行审核。

文章中可选择的引用模式如下:

1、公开数据:

\*\*\*\*\* are available in eLMSG (an eLibrary of Microbial Systematics and Genomics, https://www.biosino.org/elmsg/index) under accession numbers \*\*\*\*\*\*.

2、 私有数据:

1) 逐个株菌描述:

Strain \*\*\*\*\*\* (eLMSG genome ID: LMSG\_G\*\*\*\*) can be viewed in eLMSG (https://www.biosino.org/elmsg/index) through the URL: https://www.biosino.org/elmsg/record/\*\*\*\*\*

2) 多个株菌一起描述:

All data can be viewed in eLMSG (<u>https://www.biosino.org/elmsg/index</u>) through the URL in document "genome\_link.xlsx". Column A refers to strain name. Column B indicates eLMSG genome ID. Column C refers to URL of genome in eLMSG.