[별첨 2] 다부처유전체 연계 및 등록을 위한 2차 SOP

(국가생명연구자원정보센터, '18. 10. 2)

□ 1) BioProject 구조

필드	M/O	설명
Submitter		
name (영문)	M	Last name, First name (예, Hong, Gil Dong)
name (한글)	M	성 이름(예, 홍길동) - 만약 외국인일 경우 한글과 동일하게 "성 이름" 순으로 작성한다.
Email (1st)	M	1차 Email
Email (2st)	O	2차 Email
Organization (영문)	M	영문 기관명 (Organization)
Organization (국문)	M	한글 기관명 (Organization)
Department (영문)	M	영문 부서명 (Department)
Department (국문)	M	한글 부서명 (Department)
Phone	O	Phone (+82-2-1111-2222)
Fax	O	Fax (+82-2-1111-2222)
Address(영문)	M	예) 125 Gwahak-ro, Yeseong-gu, Daejeon 34141, Korea
Address(국문)	M	예) 34141 대전광역시 유성구 과학로 125
Country	M	국가명 선택
General Information		
Submission date	M	제출 날짜 (양식: YYYY-MM-DD, 예, 2018-12-22)
Release date selection	M	"Release immediately following curation (recommended)"와 "Release on specified date" 중 선택 (즉시 공개 또는 공개일자 지정 중에서 택1)
Release date	M	"Release date selection"에서 "Release on specified date"를 선택한 경우, 반드시 공개날짜를 입력하여야함. (양식: YYYY-MM-DD, 예, 2018-12-22)
Umbrella project	O	If your project is belonged to some umbrella project, choose correct one
Project		
Government department (국문)	M	7개 부처/청 중 선택하기 (과기정통부, 해양수산부, 보건복지부, 농림축산부, 산업부, 농진청, 산림청)
Project title (영문)	M	영문 project title
Project title (국문)	M	국문 project title
Project number (과제번호)	M	부처 전문기관에서 부여하는 과제 고유번호 (예, 한국연구재단에서 부여하는 과제번호 '2017M3C9A6048610')
Project NTIS number (과제 NTIS 번호)	M	과제의 NTIS 번호
Project type	M	과제가 총괄, 세부, 위탁, 단위, 용역, 기타_중에서 선택
Main project title (상위과제명)	M	총괄 과제명 (세부과제 등이 포함된 최상위 총괄과제 명) - 총괄과제 책임자가 입력 과제 책임자와 같은 경우도 반드시 입력하여야 함
Main project PI (상위과제 책임자)	M	총괄 과제 책임자
Main project number (상위과제 번호)	M	부처 전문기관에서 부여한 총괄과제 고유번호
Main project NTIS number (상위과제 NTIS 번호)	M	총괄 과제번호 (NTIS 번호)
Relevance	M	Select or provide the primary general relevance of the project: (Agricultural, Medical, Industrial, Environmental, Evolution, Model Organism)
Description (영문)	M	영문 과제 설명
Description (국문)	M	국문 과제 설명
Project data type	M	Indicate the general label of the primary study goal. (Table 1-1) 택1 또는 직접 입력
Sample scope	M	The scope and purity of the biological sample used for the study. (Table 1-2) 택1 또는 직접 입력
Consortium		
Consortium name	O	If study is carried out as part of a consortium, provide the consortium name
Consortium URL	O	If the consortium maintains a web site, provide the URL
Others		
Data provider	O	Indicate the data provider (data submitter) if it is someone other than the submitting organization or consortium
Biomaterial provider	O	Provide the information of the center and lab, or a culture collection identifier

Publications		
PubMed ID	O	Provide a PubMed ID

DOI	O	Provide a DOI if a PubMed ID is not available
Journal name	O	journal name
Article title	O	article title
Year	O	year
Author list	O	author list

External Links		
Link Description	O	Link Description
URL	O	URL

□ 2) BioSample 구조

필드명	(M)/(O)	설명
Submitter		
name (영문)	M	Last name, First name (예, Hong, Gil Dong)
name (한글)	M	성 이름(예, 홍길동) - 만약 외국인일 경우 한글과 동일하게 “성 이름” 순으로 작성한다.
Email (1st)	M	1차 Email
Email (2st)	O	2차 Email
Organization (영문)	M	영문 기관명 (Organization)
Organization (국문)	M	한글 기관명 (Organization)
Department (영문)	M	영문 부서명 (Department)
Department (국문)	M	한글 부서명 (Department)
Phone	O	Phone (+82-2-1111-2222)
Fax	O	Fax (+82-2-1111-2222)
Address(영문)	M	예) 125 Gwahak-ro, Yeseong-gu, Daejeon 34141, Korea
Address(국문)	M	예) 34141 대전광역시 유성구 과학로 125
Country	M	국가명 선택
General Information		
Submission date	M	샘플정보가 제출된 제출일 (YYYY-MM-DD)
Release date selection	M	“Release immediately following curation (recommended)”와 “Release on specified date” 중 선택 (즉시 공개 또는 공개일자 지정 중에서 택1)
Release date	M	“Release date selection”에서 “Release on specified date”를 선택한 경우, 반드시 공개날짜를 입력하여야함. (양식: YYYY-MM-DD, 예, 2018-12-22)
Project accession	M	샘플과 관련된 project ID를 선택
Sample title (영문)	M	Provide a brief title, as a phrase of short sentence for public display
Sample title (국문)	M	국문 샘플제목 (1개 문장으로 간략하게 표현)
Description (영문)	M	Provide a description of the study goals and relevance.
Description (국문)	M	샘플의 설명을 국문으로 표현
Sample type	M	Select the package that best describes yours amples: **1. Clinical or host-associated, pathogen, 2. Environmental, food or other pathegen, 3. Microbe. 4. Model organism or animal sample, 5. Human sample, 6. Plant sample, 7. Virus sample**

□ 3) Experiment 구조

필드	M/O	설명
Submission date	M	submission date (양식: YYYY-MM-DD, 예, 2018-12-22)
Release date selection	M	“Release immediately following curation (recommended)”와 “Release on specified date” 중 선택 (즉시 공개 또는 공개일자 지정 중에서 택1)
Release date	M	“Release date selection”에서 “Release on specified date”를 선택한 경우, 반드시 공개날짜를 입력하여야함. (양식: YYYY-MM-DD, 예, 2018-12-22)
Experiment title (영문)	M	Experiment title
Experiment title (한글)	M	한글로 된 Experiment title
Description (영문)	M	Description
Description (한글)	M	한글로 된 Description
Platform	M	Sequencing platform. 아래 리스트에서 선택
Library Construction/Experiment Desing	M	Enter the details about your experimental design and molecular strategies including hybrid selection and affinity capture reagents; any detail that distinguishes your experiment from other similar experiments
Library	O	The library descriptor specifies the origin of the material being sequenced and any treatments that the material might have undergone that affect the sequencing result
Library name	O	The submitter's name for this library
Strategy	M	Sequencing technique intended for this library. 아래 전략 리스트 중 선택하거나 입력
Source	M	The library source specifies the type of source material that is being sequenced. 아래 소스 리스트 중 선택하거나 입력
Selection	M	Whether any method wasused to select and/or enrich the material being sequenced. 아래 선택리스트 중 선택하거나 입력
Layout		
Fragemnt/ Paired read	M	Library Layout specifies whether to expect single, Pair-end, or other configuration of reads. In the case of paired reads, information about the relative distance and orientation is specified.
Insert size	M(O)	Fragment sizefor Paired reads(Paired read만 해당)
Nominal size	O	Size of the insert for Paired reads(Paired read만 해당)
Nominal standard deviation	O	Standard deviation of insert size (typically ~10% of Nominal Size)(Paired read만 해당)

□ 4) RUN 구조

필드		M/O	설명
Accession number		M	각 부처정보센터에서 부여한 고유번호
Submission date		M	submission date
Release date selection		M	“Release immediately following curation (recommended)”와 “Release on specified date” 중 선택 (즉시 공개 또는 공개일자 지정 중에서 택1)
Release date		M	“Release date selection”에서 “Release on specified date”를 선택한 경우, 반드시 공개날짜를 입력하여야함. (양식: YYYY-MM-DD, 예, 2018-12-22)
File type		M	The information about supported formats of the submitted sequence data. Fastq, BAM, VCF, SFF, Other format 가능
FastQ format	**File name**	M	We only accept GZIP and BZIP2 compression formats. Especially we don't accept 7-ZIP or TAR compressed files.
	MD5 for file	M	MD5 checksums are a 32-character alphanumeric string. For Mac and Linux system users, the native command line tools "md5sum"(Linux) and "md5"(Mac OX) can be used to generate MD5 checksums. Windows users must need to download a third-party utility
BAM format	**Reference assembly Name**	M	The Reference's assembly name
	File name	M	Submitted BAM files must be readable with SAMtools. BAM file names are required to end up with the .bam suffix (e.g. 'a.bam')
	MD5 for file	M	MD5 for bam file
	reference file name	M	The Reference's file name
	MD5 for reference file	M	MD5 for reference file
VCF format	**Reference assembly Name**	M	The Reference's assembly name
	File name	M	VCFfile names are required to end up with the .vcfsuffix (e.g. 'a.vcf')
	MD5 for file	M	MD5 for vcf file
	reference file name	M	The Reference's file name
	MD5 for reference file	M	MD5 for reference file
SFF format	**File name**	M	SFF file names are required to end up with the .sffsuffix (e.g. 'a.sff')
	MD5 for file	M	MD5 for SFF file
Other format (한국전용)	**File name**	M	기타 파일명
	MD5 for file	M	MD5 for SFF file
	description	M	기타 파일에 대한 양식 또는 간단한 설명(영문 또는 한글 가능)

[관련 테이블]

1-1) proejct data type 테이블

type	*Description*
Whole Genome sequencing	whole, or partial, genome sequencing project (with or without a genome assembly)
Clone ends	clone-end sequencing project
Epigenomics	DNA methylation, histone modification, chromatin accessibility datasets
Exome	exomeresequencing project
Map	project that results in non-sequence map data such as genetic map, radiation hybrid map, cytogenetic map, optical map, and etc.
Metagenome	sequence analysis of environmental samples
Phenotype/Genotype	project correlating phenotype and genotype
Random Survey	Sequence generated from a random sampling of the collected sample; not intended to be comprehensive sampling of the material.
Targeted Locus (Loci)	project to sequence specific loci, such as a 16S rRNA sequencing
Transcriptome or Gene Expression	Large scale RNA sequencing or expression analysis. Includes cDNA, EST, RNA_seq, and microarray.
Variation	Project with a primary goal of identifying large or small sequence variation across populations.
Other	a free text description is provided to indicate Other data type

1-2) sample scope 테이블

Sample scope	*Description*
Monoisolate	A single animal, cultured cell-line, inbred populationor possibly a heterogeneous population (a single genome assembly is generated from the pooled sample; not preferred).
Multiisolate	Multiple individuals, a population (representation of a species).
Multi-species	Sample represents multiple species.
Environment	Species content of the sample is not known. Nucleic acid is directly isolated from an environmental sample for analysis. This is used for metagenome studies.
Synthetic	Sample is synthetically created in a laboratory.
Single cell	Single cell sequencing examines the sequence information from individual cells
Other	Specify the sample scope that was used.

2-1) Sample type : 1. Clinical or host-associated 구조

필드	M/O	설명
Sample name	M	sample name. it concise, unique and consistent within your lab, and as informative as possible. Every Sample Name from a single Submitter must be unique
Organism	M	The most descriptive organism name for this sample
Isolate	O	Identification or description of the specific individual from which this sample was obtained
Strain	O	microbial or eukaryotic strain name
Collected by	M	Name of persons or institute who collected the sample
Collection date	M	Date of sampling (YYYY-MM-DD)
Geographic location	M	Geographical origin of the sample; Use a colon to separate the country or ocean from more detailed information about the location
Host	M	The natural (as opposed to laboratory) host to the organism fromwhich the sample was obtained
Host disease	M	Name of relevant disease, e.g. Salmonella gastroenteritis
Isolation source	M	Describes the physical, environmental and/or local geographical source of the biological sample from which the sample was derived
Latitude and longitude	M	The geographical coordinates of the location where the sample was collected. Specify as degrees latitude and longitude in format "d[d.dddd] N|S d[dd.dddd] W|E", e.g., 38.98 N 77.11 W
Culture collection	O	Name of source institute and unique culture identifier. See the description for the proper format and list of allowed institutes
Genotype	O	observed genotype
Host age	O	Age of host at the time of sampling
Host description	O	Additional information not included in other defined vocabulary fields
Host disease outcome	O	Final outcome of disease, e.g., death, chronic disease, recovery
Host disease stage	O	Stage of disease at the time ofsampling
Host health state	O	Information regarding health state of the individual sampled at the time of sampling
Host sex	O	Gender or physical sex of the host
Host subject id	O	a unique identifier by which each subject can be referred to, de-identified, e.g. #131
Host tissue sampled	O	Type of tissue the initial sample was taken from
Passage history	O	Number of passages and passage method
Pathotype	O	Some bacterial specific pathotypes (example Eschericia coli -STEC, UPEC)
Serotype	O	Taxonomy below subspecies; a variety (in bacteria, fungi or virus) usually based on its antigenic properties. e.g. serotype="H1N1" in Influenza A virus CY098518
Serovar	O	Taxonomy below subspecies; a variety (in bacteria, fungi or virus) usually based on its antigenic properties. Same as serovar and serotype. Sometimes used as species identifier in bacteria with shaky taxonomy, e.g. Leptospira, serovar saopaolo S76607 (65357 in Entrez)
Specimen voucher	O	Identifier for the physical specimen
Subgroup	O	Taxonomy below subspecies; sometimes used in viruses to denote subgroups taken from a single isolate
Subtype	O	Used as classifier in viruses (e.g. HIV type 1, Group M, Subtype A)

2-2) Sample type : 2. Environmental, food or other pathogen 구조

필드	M/O	설명
Sample name	M	sample name. it concise, unique and consistent within your lab, and as informative as possible. Every Sample Name from a single Submitter must be unique
Organism	M	The most descriptive organism name for this sample
Isolate	O	Identification or description of the specific individual from which this sample was obtained
Strain	O	microbial or eukaryotic strain name
Collected by	M	Name of persons or institute who collected the sample
Collection date	M	Date of sampling
Geographic location	M	Geographical origin of the sample; Use a colon to separate the country or ocean from more detailed information about the location
Isolation source	M	Describes the physical, environmental and/or local geographical source of the biological sample from which the sample was derived
Latitude and longitude	M	The geographical coordinates of the location where the sample was collected. Specify as degrees latitude and longitude in format "d[d.dddd] N|S d[dd.dddd] W|E", e.g., 38.98 N 77.11 W
Culture collection	O	Name of source institute and unique culture identifier. See the description for the proper format and list of allowed institutes
Genotype	O	observed genotype
Passage history	O	Number of passages and passage method
Pathotype	O	Some bacterial specific pathotypes (example Eschericia coli -STEC, UPEC)
Serotype	O	Taxonomy below subspecies; a variety (in bacteria, fungi or virus) usually based on its antigenic properties. e.g. serotype="H1N1" in Influenza A virus CY098518
Serovar	O	Taxonomy below subspecies; a variety (in bacteria, fungi or virus) usually based on its antigenic properties. Same as serovar and serotype. Sometimes used as species identifier in bacteria with shaky taxonomy, e.g. Leptospira, serovar saopaolo S76607 (65357 in Entrez)
Specimen voucher	O	Identifier for the physical specimen
Subgroup	O	Taxonomy below subspecies; sometimes used in viruses to denote subgroups taken from a single isolate
Subtype	O	Used as classifier in viruses (e.g. HIV type 1, Group M, Subtype A)

2-3) Sample type : 3. Microbe 구조

필드	M/O	설명
Sample name	M	sample name. it concise, unique and consistent within your lab, and as informative as possible. Every Sample Name from a single Submitter must be unique
Organism	M	The most descriptive organism name for this sample
Strain	M	microbial or eukaryotic strain name
Collected by	O	Name of persons or institute who collected the sample
Collection date	M	Date of sampling
Geographic location	M	Geographical origin of the sample; Use a colon to separate the country or ocean from more detailed information about the location
Altitude	O	The altitude of the sample is the vertical distance between Earth'ssurface above Sea Level and the sampled position in the air
Biomaterial provider	O	The name and address of the lab or PI, or a culture collection identifier
Depth	O	Depth is defined as the vertical distance below surface, e.g. for sediment or soil samples depth is measured from sediment or soil surface, respectively. Depth can be reported as an interval for subsurface samples
Environmental biome	O	descriptor of the broad ecological context of a sample. Examples include: desert, taiga, deciduous woodland, or coral reef
Host	O	The natural (as opposed to laboratory) host to the organism fromwhich the sample was obtained
Isolation source	M	Describes the physical, environmental and/or local geographical source of the biological sample from which the sample was derived
Latitude and longitude	O	The geographical coordinates of the location where the sample was collected. Specify as degrees latitude and longitude in format "d[d.dddd] N|S d[dd.dddd] W|E", e.g., 38.98 N 77.11 W
Culture collection	O	Name of source institute and unique culture identifier. See the description for the proper format and list of allowed institutes
Genotype	O	Observed genotype
Host tissue sampled	O	Type of tissue the initial sample was taken from
Identified by	O	The name of the taxonomist who identified the specimen.
Lab host	O	Scientific name and description of the laboratory host used to propagate the source organism or material from which the sample was obtained, e.g., Escherichia coli DH5a, or Homo sapiens HeLa cells
Passage history	O	Number of passages and passage method
Sample size	O	Amount or size of sample (volume, mass or area) that was collected
Serotype	O	Taxonomy below subspecies; a variety (in bacteria, fungi or virus) usually based on its antigenic properties. e.g. serotype="H1N1" in Influenza A virus CY098518
Serovar	O	Taxonomy below subspecies; a variety (in bacteria, fungi or virus) usually based on its antigenic properties. Same as serovar and serotype. Sometimes used as species identifier in bacteria with shaky taxonomy, e.g. Leptospira, serovar saopaolo S76607 (65357 in Entrez)
Specimen voucher	O	Identifier for the physical specimen
Temperature	O	temperature of the sample at time of sampling

2-4) Sample type : 4. Model organism or animal sample 구조

필드	M/O	설명
Sample name	M	sample name. it concise, unique and consistent within your lab, and as informative as possible. Every Sample Name from a single Submitter must be unique
Organism	M	The most descriptive organism name for this sample
Breed	O	breed name -chiefly used in domesticated animals or plants
Strain	O	microbial or eukaryotic strain name
Age	O	age at the time of sampling; relevant scale depends on species and study
Biomaterial provider	M	The name and address of the lab or PI, or a culture collection identifier
Sex	M	physical sex of sampled organism
Tissue	M	Type of tissue the sample was taken from
Birth date	O	The date of birth
Birth location	O	The location of birth
Breeding history	O	The history of breeding
Breeding method	O	The method of breeding
Cell line	O	Name of the cell line
Cell subtype	O	The subtype of cell
Cell type	O	Type of cell of the sample or from which the sample was obtained
Collected by	O	Name of persons or institute who collected the sample
Culture collection	O	Name of source institute and unique culture identifier. See the description for the proper format and list of allowed institutes
Death date	O	The date of death
Development stage	O	Developmental stage at the time of sampling
Disease	O	List of diseases diagnosed; can include multiple diagnoses
Disease stage	O	Stage of disease at the time of sampling
Genotype	O	Observed genotype
Geographic location	O	Geographical origin of the sample; Use a colon to separate the country or ocean from more detailed information about the location
Growth protocol	O	The protocol of growth.
Health state	O	Health or disease status of sample at time of collection
Isolation source	O	Describes the physical, environmental and/or local geographical source of the biological sample from which the sample was derived
Latitude and longitude	O	The geographical coordinates of the location where the sample was collected. Specify as degrees latitude and longitude in format "d[d.dddd] N|S d[dd.dddd] W|E", e.g. 38.98 N 77.11 W
Phenotype	O	Phenotype ofsampled organism
Specimen voucher	O	Identifier for the physical specimen
Storage conditions	O	Explain how and for how long the sample was stored before DNA extraction
Study book number	O	Study book number of sample
Treatment	O	Treatment of sample

2-5) Sample type : 5. Human sample 구조

필드	M/O	설명
Sample name	M	sample name. it concise, unique and consistent within your lab, and as informative as possible. Every Sample Name from a single Submitter must be unique
Organism	M	The most descriptive organism name for this sample
Isolate	M	List of diseases diagnosed: can include multiple diagnoses
Age	O	age at the time of sampling; relevant scale depends on species and study
Biomaterial provider	M	The name and address of the lab or PI, or a culture collection identifier
Sex	M	physical sex of sampled organism
Tissue	M	Type of tissue the sample was taken from
Disease	O	
Cell line	O	Name of the cell line
Cell subtype	O	The subtype of cell
Cell type	O	Type of cell of the sample or from which the sample was obtained
Culture collection	O	Name of source institute and unique culture identifier. See the description for the proper format and list of allowed institutes
Development stage	O	Developmental stage at the time of sampling
Disease stage	O	Stage of disease at the time of sampling
Ethnicity	O	Ethnicity of the subject
Health state	O	Health or disease status of sample at time of collection
Karyotype	O	Karyotype of sampled organism
Phenotype	O	Phenotype of sampled organism
Population	O	for human: ; for plants: filial generation, number of progeny, genetic structure
Race	O	Race of sample
Type	O	Sample type, such as cell culture, mixed culture, tissue sample, whole organism, single cell, metagenomic assembly
Treatment	O	Treatment of sample

2-6) Sample type : 6. Plant sample 구조

필드	M/O	설명
Sample name	M	sample name. it concise, unique and consistent within your lab, and as informative as possible. Every Sample Name from a single Submitter must be unique
Organism	M	The most descriptive organism name for this sample
Cultivar	M	Cultivarname: cultivated variety of plant
Age	O	age at the time of sampling; relevant scale depends on species and study
Biomaterial provider	M	The name and address of the lab or PI, or a culture collection identifier
Tissue	M	Type of tissue the sample was taken from
Cell line	O	Name of the cell line
Cell type	O	Type of cell of the sample or from which the sample was obtained
Collected by	O	Name of persons or institute who collected the sample
Collection date		
Culture collection	O	Name of source institute and unique culture identifier. See the description for the proper format and list of allowed institutes
Development stage	O	Developmental stage at the time of sampling
Disease	O	List of diseases diagnosed; can include multiple diagnoses
Disease stage	O	Stage of disease at the time of sampling
Genotype	O	Observed genotype
Growth protocol	O	The protocol of growth.
Height or length	O	Measurement of height or length
Isolation source	O	Describes the physical, environmental and/or local geographical source of the biological sample from which the sample was derived
Latitude and longitude	O	The geographical coordinates of the location where the sample was collected. Specify as degrees latitude and longitude in format "d[d.dddd] N|S d[dd.dddd] W|E", e.g. 38.98 N 77.11 W
Phenotype	O	Phenotype ofsampled organism
Population	O	for human: ; for plants: filial generation, number of progeny, genetic structure
Type	O	Sample type, such as cell culture, mixed culture, tissue sample, whole organism, single cell, and metagenomics assembly
Sex	M	Physical sex of sampled organism
Specimen voucher	O	Identifier for the physical specimen
Temperature	O	Temperature of the sample at time of sampling
Treatment	O	Treatment of sample

2-7) Sample type : 7. Virus sample 구조

필드	M/O	설명
Sample name	M	sample name. it concise, unique and consistent within your lab, and as informative as possible. Every Sample Name from a single Submitter must be unique
Organism	M	The most descriptive organism name for this sample
Isolate	M	identification or description of the specific individual from which this sample was obtained
Strain	O	microbial or eukaryotic strain name
Host	M	The natural (as opposed to laboratory) host to the organism from which the sample was obtained
Lab host	O	Scientific name and description of the laboratory host used to propagate the source organism or material from which the sample was obtained, e.g., Escherichia coli DH5a, or Homo sapiens HeLa cells
Collection date	M	Date of sampling
Geographic location	M	Geographical origin of the sample; Use a colon to separate the country or ocean from more detailed information about the location
Isolation source	M	Describes thephysical, environmental and/or local geographical source of the biological sample from which the sample was derived
Altitude	O	The altitude of the sample is the vertical distance between Earth's surface above Sea Level and the sampled position in the air
Biomaterial provider	O	name and address of the lab or PI, or a culture collection identifier
Collected by	O	Name of persons or institute who collected the sample
Culture collection	O	Name of source institute and unique culture identifier. See the description for the proper format and list of allowed institutes
Depth	O	Depth is defined as the vertical distance below surface, e.g. for sediment or soil samples depth is measured from sediment or soil surface, respectively. Depth can be reported as an interval for subsurface samples
Disease	O	List of diseases diagnosed; can include multiple diagnoses
Environment biom	O	descriptorof the broad ecological context of a sample. Examples include: desert, taiga, deciduous woodland, or coral reef
Genotype	O	Observed genotype
Host tissue sampled	O	Type of tissue the initial sample was taken from
Identified by	O	name of the taxonomist who identified the specimen
Latitude and longitude	O	The geographical coordinates of the location where the sample was collected. Specify as degrees latitude and longitude in format "d[d.dddd] N|S d[dd.dddd] W|E", e.g. 38.98 N 77.11 W
Passage history	O	Number of passages and passage method
Sample size	O	Amount or size of sample (volume, mass or area) that was collected
Serotype	O	Taxonomy below subspecies; a variety (in bacteria, fungi or virus) usually based on its antigenic properties. e.g. serotype="H1N1"in Influenza A virus CY098518
Specimen voucher	O	Identifier for the physical specimen
Temperature	O	temperatureof the sample at time of sampling

3-1) Sequencing platform

454 GS 20
454 GS FLX
454 GS FLX Titanium
.....

3-2) Strategy

type	*Description*
WGA	Random sequencing of the whole genome following non-PCR amplification
WGS	Random sequencing of the whole genome
WXS	Random sequencing of exonic regions selected from the genome
RNA-Seq	Random sequencing of whole transcriptome
miRNA-Seq	Random sequencing of small miRNAs
Tn-Seq	Sequencing from transposon insertion sites
WCS	Random sequencing of a whole chromosome or other replicon isolated from a genome
CLONE	Genomic clone based (hierarchical) sequencing
POOLCLONE	Shotgun of pooled clones (usually BACs and Fosmids)
AMPLICON	Sequencing of overlapping or distinct PCR or RT-PCR products
CLONEEND	Clone end (5′, 3′, or both) sequencing
FINISHING	Sequencing intended to finish (close) gaps in existing coverage
ChIP-Seq	Direct sequencing of chromatin immunoprecipitates
MNase-Seq	Direct sequencing following MNase digestion
DNase-Hypersensitivity	Sequencing of hypersensitive sites, or segments of open chromatin that are more readily cleaved by DNaseI
Bisulfite-Seq	Sequencing following treatment of DNA with bisulfite to convert cytosine residues to uracil depending on methylation status
EST	Single pass sequencing of cDNA templates
FL-cDNA	Full-length sequencing of cDNA templates
CTS	Concatenated Tag Sequencing
MRE-Seq	Methylation-Sensitive Restriction Enzyme Sequencing strategy
MeDIP-Seq	Methylated DNA Immunoprecipitation Sequencing strategy
MBD-Seq	Direct sequencing of methylated fractions sequencing strategy
OTHER	Library strategy not listed (please include additional info in the “design description”)

3-3) Source 선택리스트

type	*Description*
GENOMIC	Genomic DNA (includes PCR products from genomic DNA)
TRANSCRIPTOMIC	Transcription products or non-genomic DNA (EST, cDNA, RT-PCR, screened libraries)
METATRANSCRIPTOMIC	Transcription products from community targets
METAGENOMIC	Mixed material from metagenome
SYNTHETIC	Synthetic DNA
VIRAL RNA	Viral RNA
OTHER	Other, unspecified, or unknown library source material (please include additional info in the “design description”)

3-4) Selection 선택리스트

type	*Description*
unspecified	Library enrichment, screening, or selection is not specified (please include additional info in the “design description”)
RANDOM	Random selection by shearing or other method
PCR	Source material was selected by designed primers
RANDOM PCR	Source material was selected by randomly generated primers
RT-PCR	Source material was selected by reverse transcription PCR
HMPR	Hypo-methylated partial restriction digest
MF	Methyl Filtrated
CF-S	Cot-filtered single/low-copy genomic DNA
CF-M	Cot-filtered moderately repetitive genomic DNA
CF-H	Cot-filtered highly repetitive genomic DNA
CF-T	Cot-filtered theoretical single-copy genomic DNA
MDA	Multiple displacement amplification
MSLL	Methylation Spanning Linking Library
cDNA	complementary DNA
ChIP	Chromatin immunoprecipitation
MNase	Micrococcal Nuclease (MNase) digestion
DNAse	Deoxyribonuclease (MNase) digestion
Hybrid Selection	Selection by hybridization in array or solution
Reduced Representation	Reproducible genomic subsets, often generated by restriction fragment size selection, containing a manageable number of loci to facilitate re-sampling
Restriction Digest	DNA fractionation using restriction enzymes
5-methylcytidine antibody	Selection of methylated DNA fragments using an antibody raised against 5-methylcytosine or 5-methylcytidine (m5C)
MBD2 protein methyl-CpG binding domain	Enrichment by methyl-CpG binding domain
CAGE	Cap-analysis gene expression
RACE	Rapid Amplification of cDNA Ends
size fractionation	Physical selection of size appropriate targets

Padlock probes capture method	Circularized oligonucleotide probes
Poly-A	polyA enriched RNA-seq
other	Other library enrichment, screening, or selection process (please include additional info in the “design description”)