

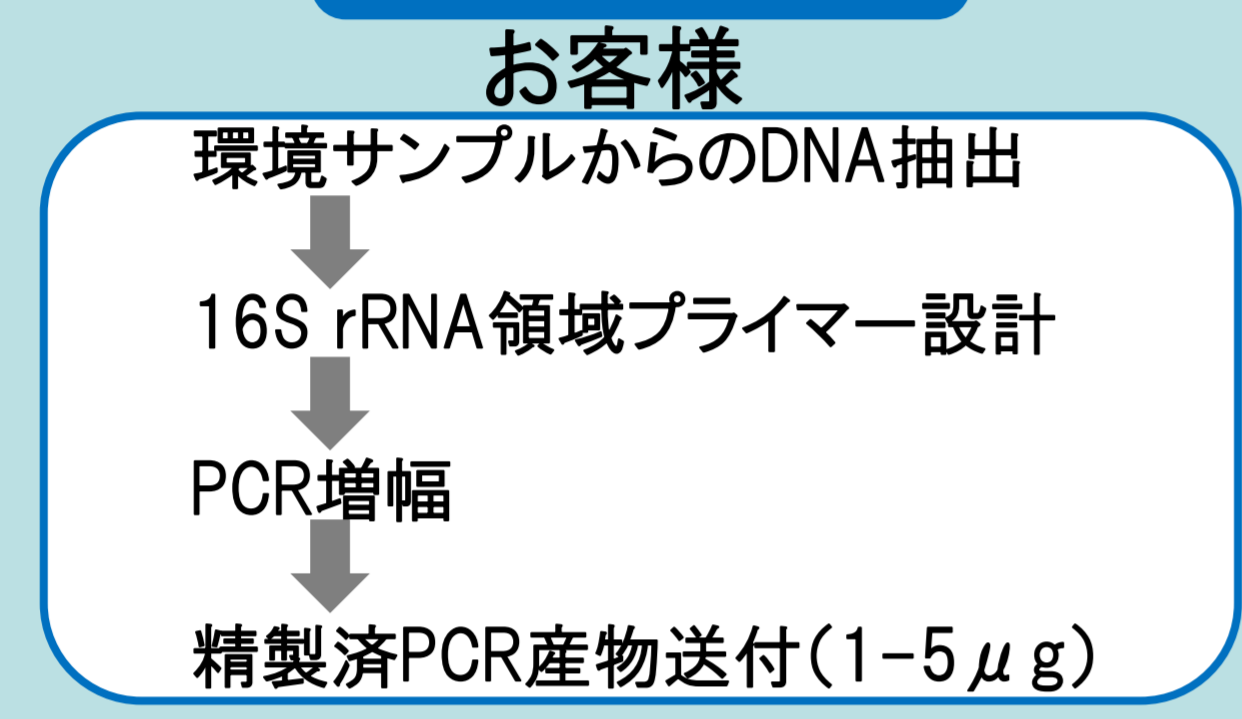
次世代シーケンサーを用いた解析のご紹介<16S rRNAシーケンス解析①>

タカラバイオ株式会社 ドラゴンジェノミクスセンター

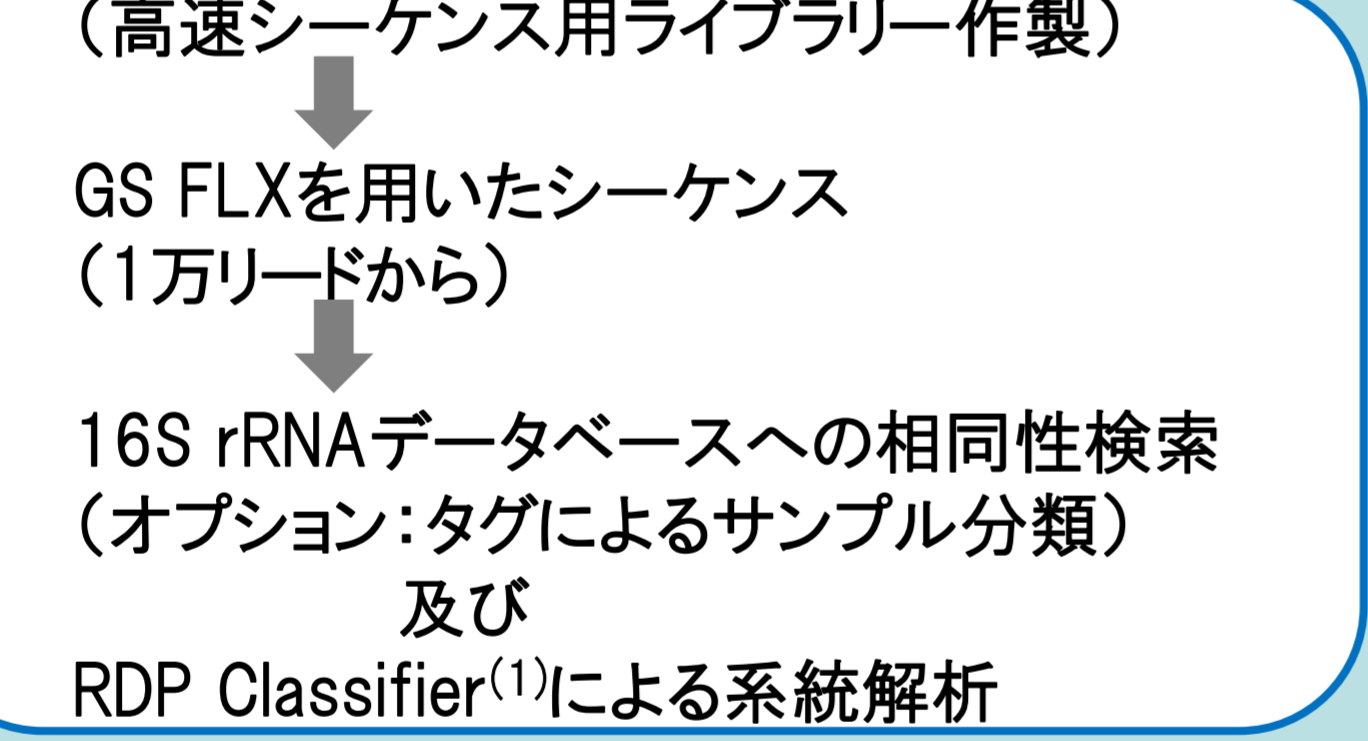
次世代シーケンサーを用いるメリット

1. クローニングバイアスを考慮する必要がない
2. プライマーにタグ配列を付加することで、複数検体を同時に解析可能
3. サンガー法と比較して大量の配列を取得可能

解析の流れ

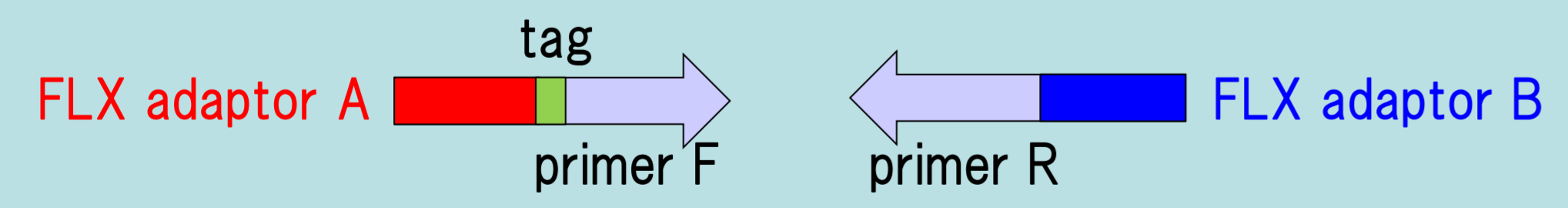


タカラバイオ

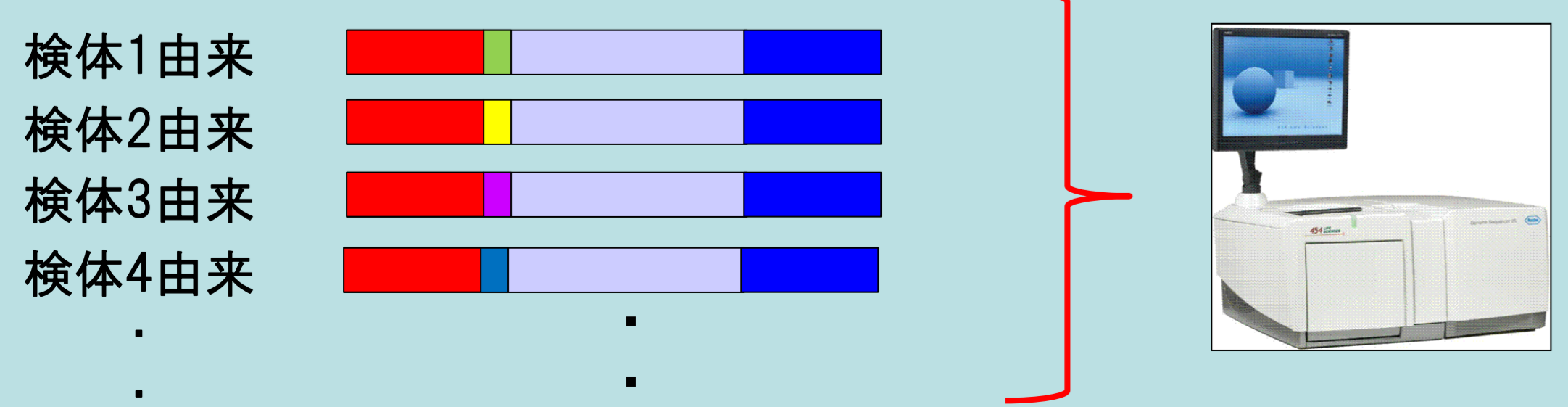


複数検体を混合する場合の解析例

1. 検体を識別するタグ配列を付加したプライマーで鋳型DNAをPCR



2. PCR産物を混合し、タグ付きライブラリーを作製 ⇒ シーケンス



3. タグ配列を利用して検体ごとに取得リードを分別して情報解析

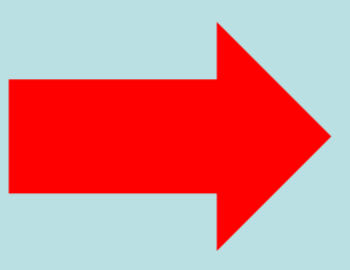
各解析手法の特徴

相同性検索

・データベース内の配列と高い相同性を示す配列が存在する場合は、Species、Strainレベルで細菌種の特特定が可能

系統分類解析

・Domain~Genus各レベルにおいて、分類の信頼性(ブートストラップ確率)を表示 ⇒ 例えば、“*Escherichia*属”とは言い切れないが、“*Enterobacteriaceae*科”であることは確からしい、といった解釈が可能



研究目的に応じた使い分けが可能

納品物イメージ

○相同性検索結果:最新の16S rRNAデータベースに対して相同性検索を行い、上位5件をリスト化

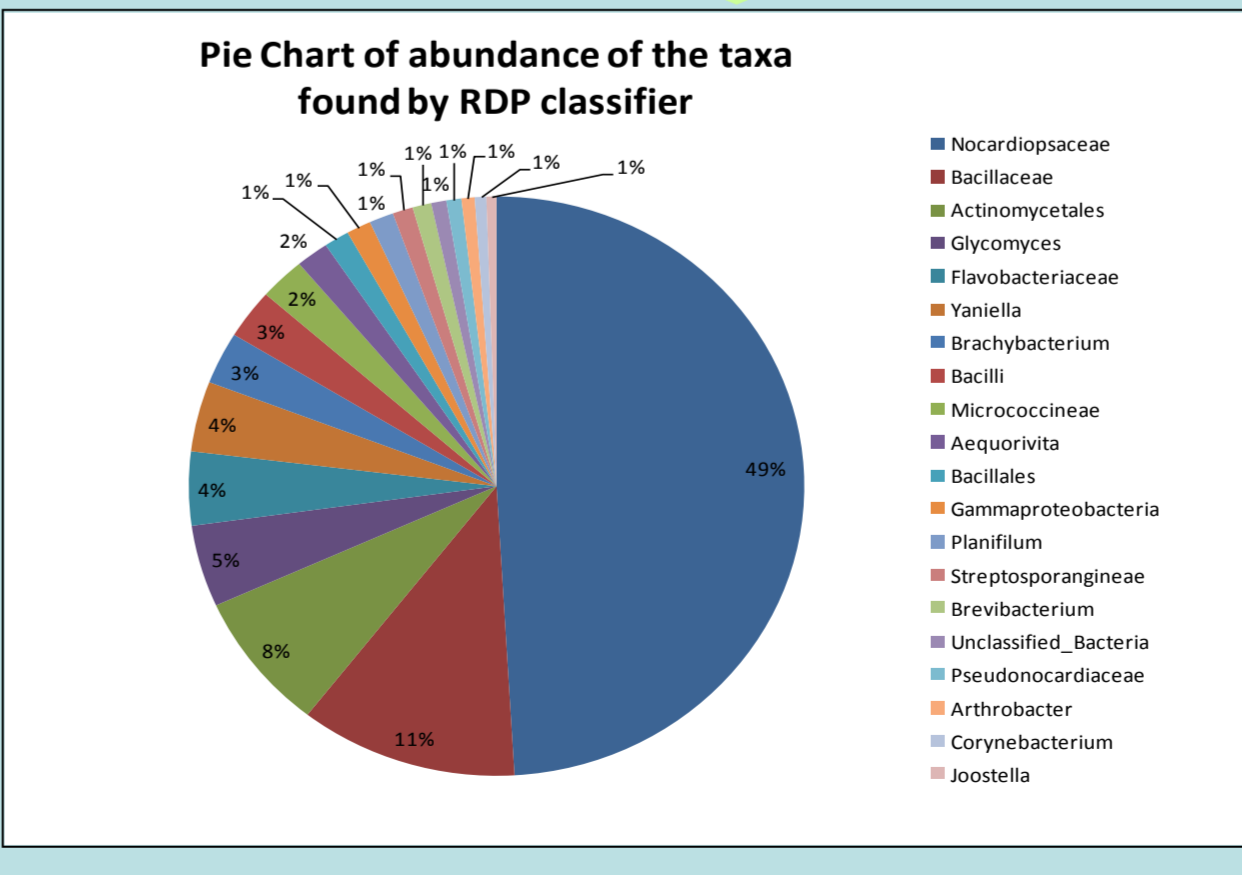
Query	QUERY SIZE	QUERY HIT	Comment	ORG	Comment2	Length	HitPosNo	Score	Expect	Identities_1	Identities_2	Gaps	Strand	QUERY START	QUERY END	SUBJECT START	SUBJECT END	QUERY STRING	SUBJECT STRING
FG7DF1X1	120	1	EF010836	Virgibacillus sp. 171-4.2-CV-H26(h)-06	EF010836.1	438	1	174	4.00E-43	112	116	3	No	5	120	337	225	actcctgcttactcctgctc	actcctgcttactcctgctc
FG7DF1X1	120	2	EU037266	Bacillus sp. G1DM-84	EU037266.1	1492	1	167	9.00E-41	110	116	2	No	5	120	357	244	actcctgcttactcctgctc	actcctgcttactcctgctc
FG7DF1X1	120	3	EF023842	Bacillus cereus	EF023842.1	1429	1	167	9.00E-41	93	96	0	No	5	100	335	240	actcctgcttactcctgctc	actcctgcttactcctgctc
FG7DF1X1	120	4	EF113684	Bacillus thuringiensis	EF113684.1	699	1	167	9.00E-41	110	116	2	No	5	120	334	221	actcctgcttactcctgctc	actcctgcttactcctgctc
FG7DF1X1	120	5	EF101186	Virgibacillus sp. 30-7	EF101186.1	637	1	167	9.00E-41	111	116	3	No	5	120	345	233	ccccaactgctccccaactgctc	ccccaactgctccccaactgctc
FG7DF1X1	121	1	EU274333	Micromonospora sp. QAI13	EU274333.1	1372	1	186	1.00E-46	117	122	2	No	1	121	296	176	ccccactgctccccactgctc	ccccactgctccccactgctc
FG7DF1X1	121	2	EF216369	Knoellia sp. TFS 104	EF216369.1	1490	1	186	1.00E-46	117	122	2	No	1	121	356	236	ccccactgctccccactgctc	ccccactgctccccactgctc
FG7DF1X1	121	3	AF060684	Kineococcus-like bacterium AS3088	AF060684.1	1525	1	186	1.00E-46	117	122	2	No	1	121	347	227	ccccactgctccccactgctc	ccccactgctccccactgctc
FG7DF1X1	121	4	FJ482120	bacterium SRU 17	FJ482120.1	743	1	186	1.00E-46	117	122	2	No	1	121	264	144	ccccactgctccccactgctc	ccccactgctccccactgctc
FG7DF1X1	121	5	AF060675	Kineococcus-like bacterium AS2965	AF060675.1	1516	1	186	1.00E-46	117	122	2	No	1	121	347	227	ccccactgctccccactgctc	ccccactgctccccactgctc

○系統分類結果:DomainレベルからGenusレベルまで分類した結果をリストおよびグラフ化

Reed_name	Domain	Bootstrap value	Phylum	Bootstrap value	Class	Bootstrap value	Subclass	Bootstrap value	Order	Bootstrap value	Suborder	Bootstrap value	Family	Bootstrap value	Genus	Bootstrap value
FG7DF1X10F780S	Bacteria	1	Actinobacteria	1	Actinobacteria	1	Actinobacteridae	1	Actinomycetales	1	Streptosporangineae	0.87	Nocardioseae	0.86	Streptomonospora	0.29
FG7DF1X10F9J3C	Bacteria	1	Actinobacteria	1	Actinobacteria	1	Actinobacteridae	1	Actinomycetales	1	Micrococineae	0.99	Dermabacteraceae	0.89	Brachyacterium	0.86
FG7DF1X10F5VJH	Bacteria	0.96	Actinobacteria	0.96	Actinobacteria	0.96	Rubrobacteridae	0.96	Scaldisprobaetales	0.72			Comexibacteraceae	0.57	Coneixibacter	0.57
FG7DF1X10F8T4S	Bacteria	1	Proteobacteria	1	Betaproteobacteria	1	Burkholderiales	1					Alcaligenaceae	1	Alcaligenes	0.49
FG7DF1X10F959A	Bacteria	1	Actinobacteria	0.99	Actinobacteria	0.99	Actinobacteridae	0.99	Actinomycetales	0.99	Streptosporangineae	0.95	Nocardioseae	0.94	Nocardiosis	0.44
FG7DF1X10F6EFD	Bacteria	1	Actinobacteria	1	Actinobacteria	1	Actinobacteridae	1	Actinomycetales	1	Streptosporangineae	0.83	Nocardioseae	0.57	Haloactinospora	0.39
FG7DF1X10GF46A	Bacteria	1	Actinobacteria	1	Actinobacteria	1	Actinobacteridae	1	Actinomycetales	1	Streptosporangineae	0.83	Nocardioseae	0.78	Streptomonospora	0.34
FG7DF1X10GBQ7I	Bacteria	1	Actinobacteria	1	Actinobacteria	1	Actinobacteridae	1	Actinomycetales	1	Micrococineae	1	Dermabacteraceae	0.93	Brachyacterium	0.91
FG7DF1X10GA74C	Bacteria	1	Actinobacteria	1	Actinobacteria	1	Actinobacteridae	1	Actinomycetales	1	Streptosporangineae	0.93	Nocardioseae	0.93	Streptomonospora	0.43
FG7DF1X10FBK0O	Bacteria	1	Actinobacteria	0.94	Actinobacteria	0.94	Acidimicrobiae	0.92	Acidimicrobiales	0.75	Acidimicrobiineae	0.75	Iamiaeae		Iamia	0.73

系統分類結果を元に菌叢構造をグラフ化

No	Taxon	Count
1	Nocardioseae	5669
2	Bacillaceae	1324
3	Actinomycetales	892
4	Glycomyces	527
5	Flavobacteriaceae	475
6	Yaniella	453
7	Brachyacterium	336
8	Bacilli	319
9	Micrococineae	278
10	Aequorivita	198
11	Bacillales	153
12	Gammaproteobacteria	149
13	Planifilum	148
14	Streptosporangineae	124
15	Brevibacterium	114
16	Unclassified_Bacteria	92
17	Pseudonocardioseae	91
18	Arthrobacter	80
19	Corynebacterium	70
20	Joostella	61



信頼性が高い(弊社デフォルトではブートストラップ確率が0.6以上)分類名を抜き出してカウント

*参考文献
(1)Wang Q *et al.*, (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73(16):5261-7

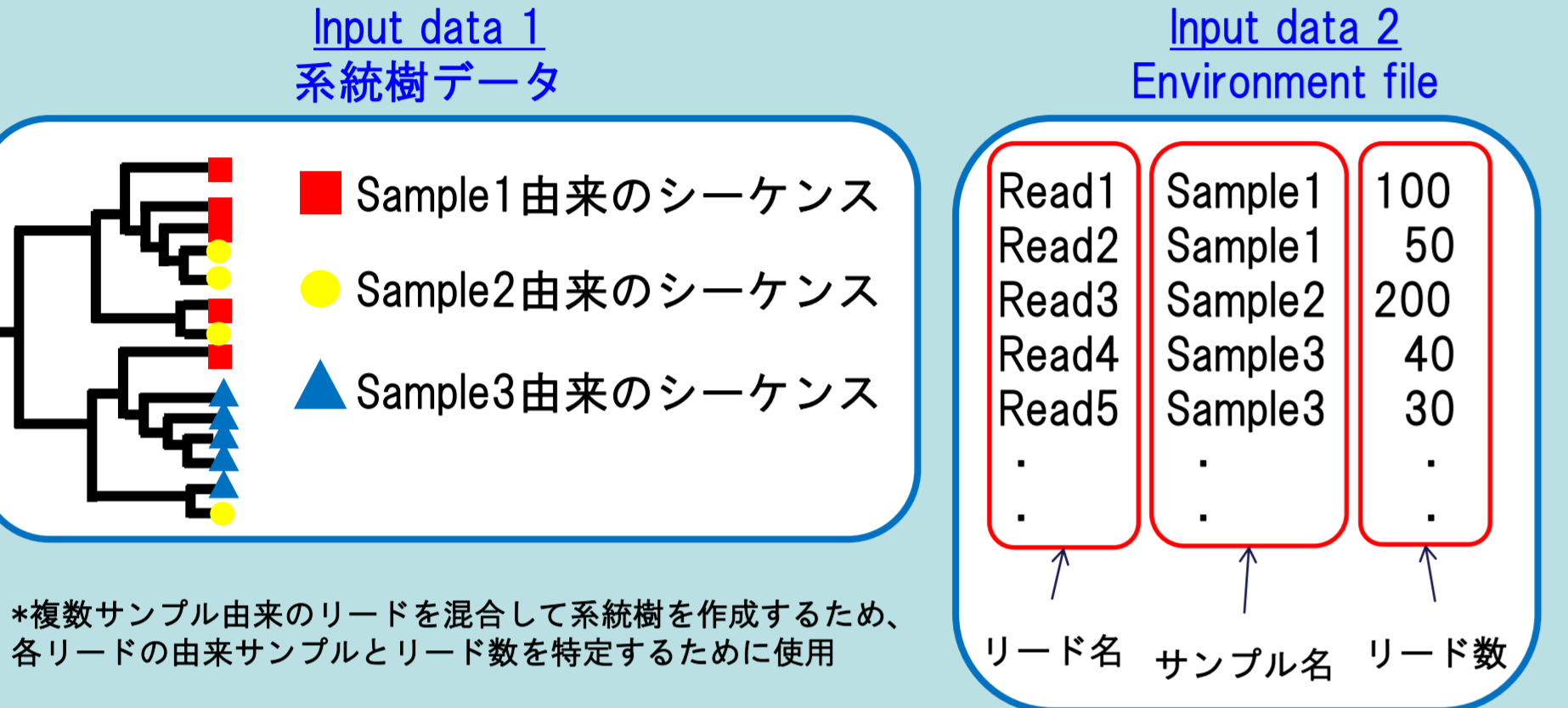
次世代シーケンサーを用いた解析のご紹介

<16S rRNAシーケンス解析②～菌叢比較解析(UniFrac⁽¹⁾解析)～>

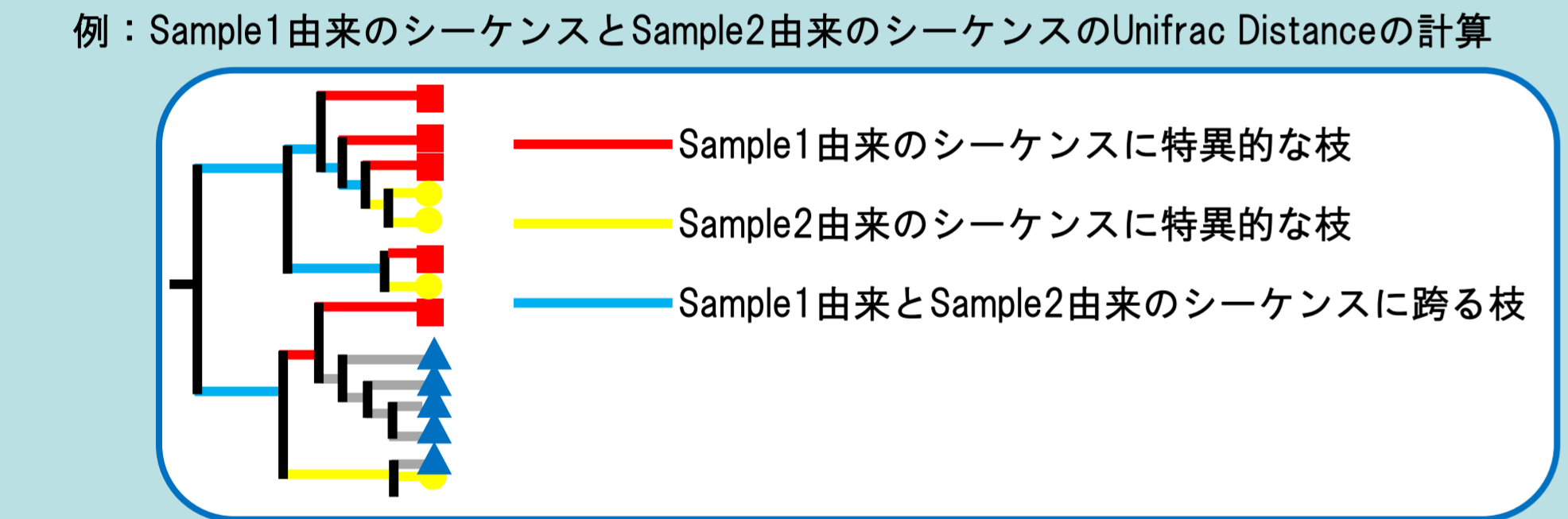
タカラバイオ株式会社 ドラゴンジェノミクスセンター

菌叢比較解析の流れ

1. 配列長をそろえたリードデータの作成
2. 相同配列のクラスタリングおよび代表配列の抽出
3. クラスタリング結果および代表配列を用いて、系統樹データ(Input data 1)およびEnvironment file*(Input data 2)を作成



4. Input data 1と2を用いてUniFrac distance matrixを作成



$$\text{UniFrac distance} = \frac{\text{各サンプルに固有の枝長[赤+黄]}}{\text{全ての枝長の和[赤+黄+水色]}}$$

全てのサンプル由来のシーケンス間 (赤 vs 黄、黄 vs 青、青 vs 赤) で UniFrac Distance (Weighted、Unweighted) を計算

UniFrac Distance Matrix

	■	●	▲
■	0	0.3	0.7
●	0.3	0	0.6
▲	0.7	0.6	0

Weighted UniFrac
: community structure についての解析 (リード数を考慮した解析)

Unweighted UniFrac
: community member についての解析 (リード数を考慮しない解析)



各解析手法の特徴

PCoA

古典的多次元尺度構成法または主座標分析と呼ばれる多変量解析の1手法。
各サンプルを点で表現し、個体間の親近性データを2次元あるいは3次元空間に、**類似したものを近くに、そうでないものは遠くに配置**する方法(納品物イメージ参照)。

Cluster Samples

UPGMA (Unweighted Pair Group Method with Arithmetic mean:非加重結合)法による階層的クラスタリング解析。
UniFrac distanceの値が小さい(= サンプル間の菌叢の違いが少ない)組み合わせから順にクラスタリングされていく(納品物イメージ参照)。

納品物イメージ

Turqueらのデータ⁽²⁾ (環境汚染の程度が異なる2つの環境における海綿共生細菌叢と海水中の細菌叢とを比較) をテストデータとして使用し、解析を行った (Accession No: GU058339-GU058920)

背景

海綿動物の環境適応機構の一部は体内に共生する細菌に関係すると考えられている。また、海綿動物に共生する古細菌は、アンモニア酸化作用を有していることが示唆されており、富栄養化など環境汚染が進んだ海域においてアンモニアの代謝にかかわっていると考えられている。そこで筆者らは、環境汚染が進んだ海域に生息する海綿共生細菌群と環境汚染が比較的進んでいない海域に生育する海綿共生細菌群、さらに両海域で共生を行わず存在している細菌群の菌叢比較を行った。

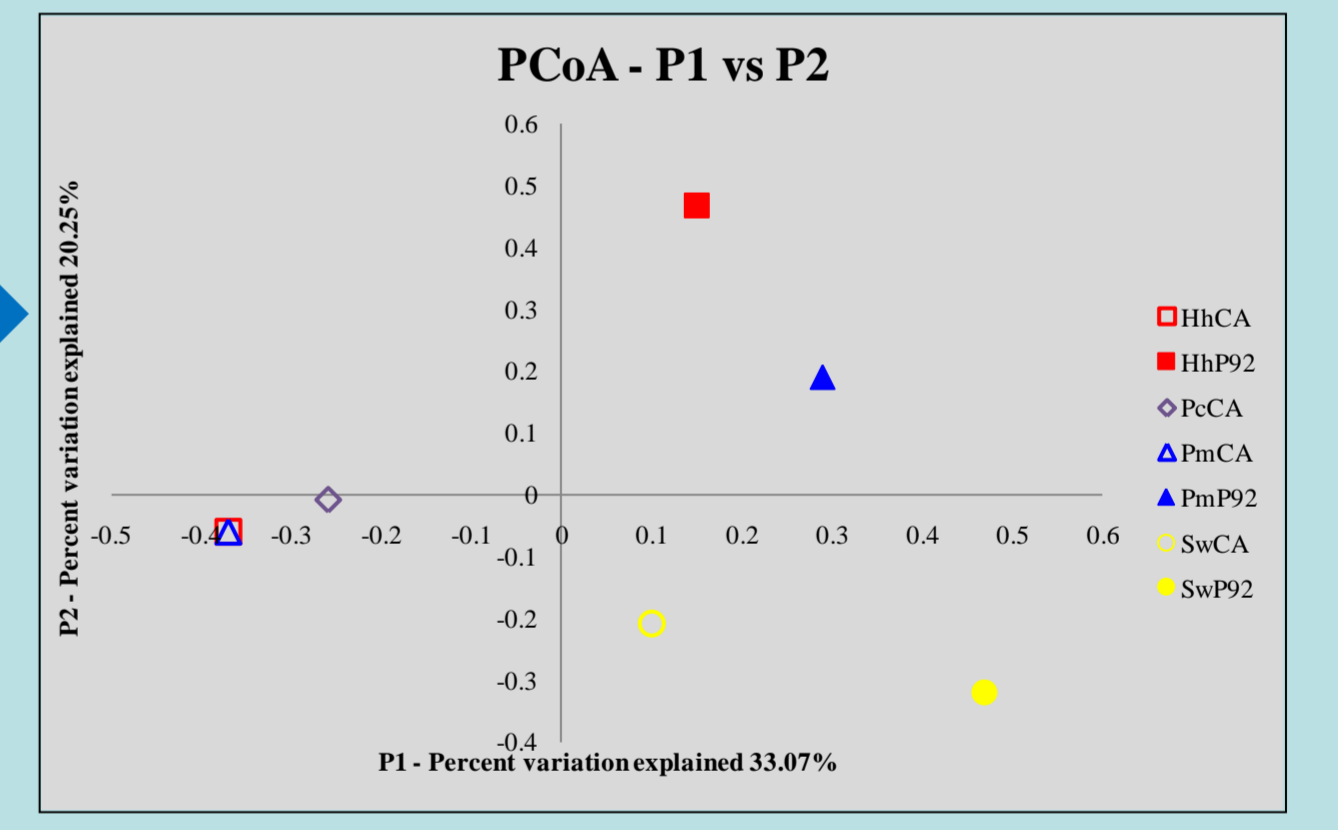
解析結果

○ UniFrac distance matrix (Unweighted)

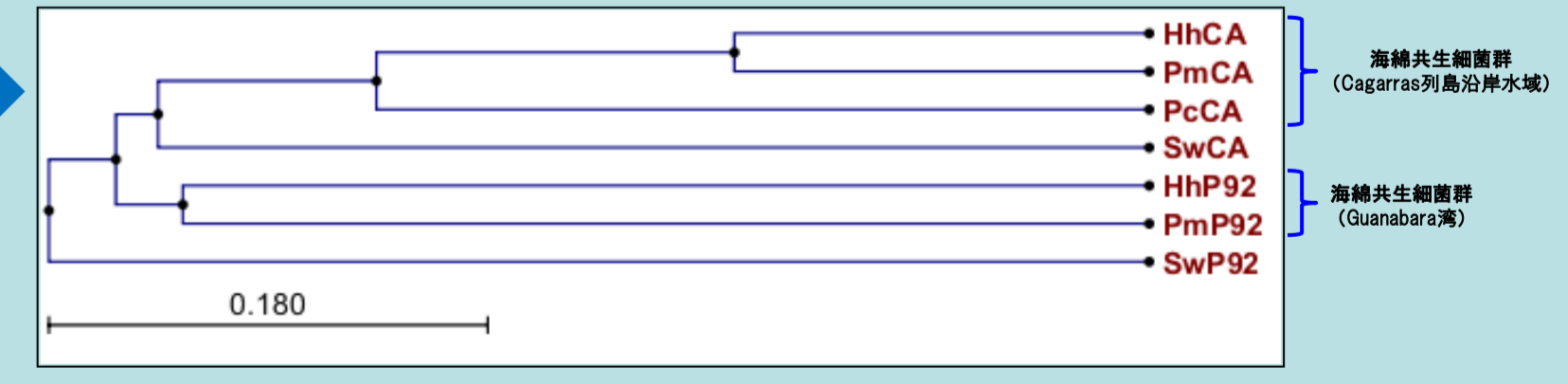
	HhCA	HhP92	PcCA	PmCA	PmP92	SwCA	SwP92
HhCA	0	0.88	0.74	0.39	0.90	0.86	0.96
HhP92	0.88	0	0.89	0.88	0.82	0.93	0.93
PcCA	0.74	0.89	0	0.75	0.91	0.90	0.97
PmCA	0.39	0.88	0.75	0	0.90	0.86	0.96
PmP92	0.90	0.82	0.91	0.90	0	0.81	0.87
SwCA	0.86	0.93	0.90	0.86	0.81	0	0.89
SwP92	0.96	0.93	0.97	0.96	0.87	0.89	0

- HhCA : Cagarras列島沿岸水域*の海綿動物*Hymeniacidon heliophilai*に共生する細菌群
 - HhP92 : Guanabara湾付近*の海綿動物*Hymeniacidon heliophilai*に共生する細菌群
 - PcCA : Cagarras列島沿岸水域の海綿動物*Petromica citrinata*に共生する細菌群
 - PmCA : Cagarras列島沿岸水域の海綿動物*Paraleucilla magna*に共生する細菌群
 - PmP92 : Guanabara湾付近の海綿動物*Paraleucilla magna*に共生する細菌群
 - SwCA : Cagarras列島沿岸水域に存在するfree-livingの細菌群
 - SwP92 : Guanabara湾に存在するfree-livingの細菌群
- *注1, Cagarras列島沿岸水域: 環境汚染が進んでいない海域
*注2, Guanabara湾付近: 環境汚染が進んでいる海域

○ PCoAの結果をScatter Plotで表現した結果



○ Cluster Samplesを行った結果



これらの解析から示唆されること

- ・Cagarras列島沿岸水域(環境汚染が進んでいない)の菌叢とGuanabara湾(環境汚染が進んでいない)の菌叢は異なっている
 - ・同じ環境に生息する海綿共生細菌叢(HhCAとPmCAとPcCA、HhP92とPmP92)は互いに類似しており、free-livingの細菌叢とは異なっている
- ➡ 本解析においても、Turqueらと同様の結論が導かれた

*参考文献

(1) Hamady M *et al.*, (2010) Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME Journal* 4(1): 17-27
(2) Turque AS *et al.*, (2010) Environmental shaping of sponge associated archaeal communities. *PLoS One*. 30;5(12):e15774.