

(a)

菌株1.fasta

```

>gene1
GCTAGCA...
>gene3
TTCAATG...
>gene5_rev
ACACTACA...
>gene4
GGCTAGG...
>gene6
CCACATTC...

```

菌株2.fasta

```

>gene3
TTCAATG...
>gene4_rev
AGGGTAGA...
>gene6
CCACATTC...
>gene7_rev
TGTGCAT...
>gene9
CGGCAA...

```

菌株3.fasta

```

>gene2
AGGTTAA...
>gene3
TTCAATG...
>gene6_rev
TCGGATTT...
>gene7_rev
TGTGCAT...
>gene8
CTCTGAGA...

```

(b)

菌株1.fasta

```

>gene1
GCTAGCA...
>gene3
TTCAATG...
>gene5_rev
ACACTACA...
>gene4
GGCTAGG...
>gene6
CCACATTC...

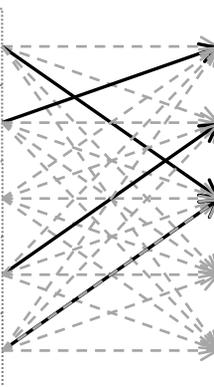
```

菌株2.fasta

```

>gene3
TTCAATG...
>gene4_rev
AGGGTAGA...
>gene6
CCACATTC...
>gene7_rev
TGTGCAT...
>gene9
CGGCAA...

```



(c)

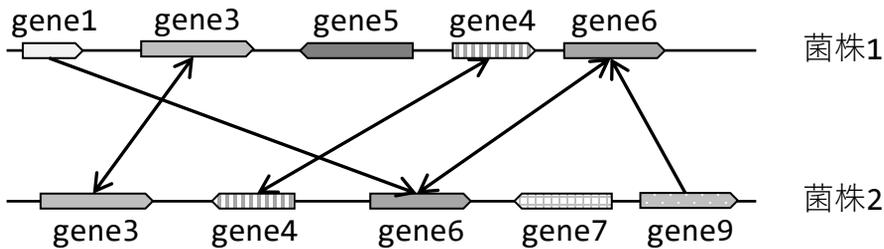


図2.2 ゲノム (菌株) 間での同一遺伝子の対応付け (オーソログの同定)

(a) (multi-)FASTA形式ファイルの例。大なり記号 (>) から始まる行のことをdescription行といい、その次の行から配列情報が記載される。配列情報は複数行になってもよい。ファイルの拡張子は、.fasta以外に.faでもよい。また、配列情報は塩基配列でもアミノ酸配列でもよいので、拡張子の段階でどちらかを見分けられるように、前者の場合は.fna、そして後者の場合は.faaが用いられることもある。(b)BLASTを用いた配列類似性検索例。ここでは一方向 (菌株1 → 菌株2) のみの検索結果の概念図を示している。(c) 双方向のヒット結果のみを示した概念図。両端に矢印がある3遺伝子がオーソログ判定結果となる。