

第一次作业

(1)

第四列为所示特征序列在目标序列/染色体中的起始位点，第五列为该特征序列的终止位点。如果该序列是一个exon，则其长度应该为第五列减去第四列加上1。

(2)

```
IV      ensembl CDS      1510859 1511461 .      -      0      gene_id "YDR537C";  
gene_version "1"; transcript_id "YDR537C"; transcript_version "1"; exon_number  
"1"; gene_source "ensembl"; gene_biotype "protein_coding"; transcript_name  
"YDR537C"; transcript_source "ensembl"; transcript_biotype "protein_coding";  
protein_id "YDR537C"; protein_version "1";
```

```
IV      ensembl CDS      1510902 1511627 .      +      0      gene_id "YDR538W";  
gene_version "1"; transcript_id "YDR538W"; transcript_version "1"; exon_number  
"1"; gene_name "PAD1"; gene_source "ensembl"; gene_biotype "protein_coding";  
transcript_name "PAD1"; transcript_source "ensembl"; transcript_biotype  
"protein_coding"; protein_id "YDR538W"; protein_version "1";
```

```
IV      ensembl CDS      1512094 1513602 .      +      0      gene_id "YDR539W";  
gene_version "1"; transcript_id "YDR539W"; transcript_version "1"; exon_number  
"1"; gene_name "FDC1"; gene_source "ensembl"; gene_biotype "protein_coding";  
transcript_name "FDC1"; transcript_source "ensembl"; transcript_biotype  
"protein_coding"; protein_id "YDR539W"; protein_version "1";
```

```
IV      ensembl CDS      1517139 1517675 .      -      0      gene_id "YDR540C";  
gene_version "1"; transcript_id "YDR540C"; transcript_version "1"; exon_number  
"1"; gene_name "IRC4"; gene_source "ensembl"; gene_biotype "protein_coding";  
transcript_name "IRC4"; transcript_source "ensembl"; transcript_biotype  
"protein_coding"; protein_id "YDR540C"; protein_version "1";
```

```
IV      ensembl CDS      1519667 1520698 .      -      0      gene_id "YDR541C";  
gene_version "1"; transcript_id "YDR541C"; transcript_version "1"; exon_number  
"1"; gene_source "ensembl"; gene_biotype "protein_coding"; transcript_name  
"YDR541C"; transcript_source "ensembl"; transcript_biotype "protein_coding";  
protein_id "YDR541C"; protein_version "1";
```

```
IV      ensembl CDS      1523249 1523608 .      +      0      gene_id "YDR542W";  
gene_version "1"; transcript_id "YDR542W"; transcript_version "1"; exon_number  
"1"; gene_name "PAU10"; gene_source "ensembl"; gene_biotype "protein_coding";  
transcript_name "PAU10"; transcript_source "ensembl"; transcript_biotype  
"protein_coding"; protein_id "YDR542W"; protein_version "1";
```

```
IV      ensembl CDS      1524637 1524933 .      -      0      gene_id "YDR543C";  
gene_version "1"; transcript_id "YDR543C"; transcript_version "1"; exon_number  
"1"; gene_source "ensembl"; gene_biotype "protein_coding"; transcript_name  
"YDR543C"; transcript_source "ensembl"; transcript_biotype "protein_coding";  
protein_id "YDR543C"; protein_version "1";
```

```
IV      ensembl CDS      1525098 1525523 .      -      0      gene_id "YDR544C";
gene_version "1"; transcript_id "YDR544C"; transcript_version "1"; exon_number
"1"; gene_source "ensembl"; gene_biotype "protein_coding"; transcript_name
"YDR544C"; transcript_source "ensembl"; transcript_biotype "protein_coding";
protein_id "YDR544C"; protein_version "1";

IV      ensembl CDS      1530866 1531342 .      -      0      gene_id "YDR545C-
A"; gene_version "1"; transcript_id "YDR545C-A"; transcript_version "1";
exon_number "1"; gene_source "ensembl"; gene_biotype "protein_coding";
transcript_name "YDR545C-A"; transcript_source "ensembl"; transcript_biotype
"protein_coding"; protein_id "YDR545C-A"; protein_version "1";

IV      ensembl CDS      1526321 1531708 .      +      0      gene_id "YDR545W";
gene_version "1"; transcript_id "YDR545W"; transcript_version "1"; exon_number
"1"; gene_name "YRF1-1"; gene_source "ensembl"; gene_biotype "protein_coding";
transcript_name "YRF1-1"; transcript_source "ensembl"; transcript_biotype
"protein_coding"; protein_id "YDR545W"; protein_version "1";
```

(3)

```
853 start_codon
853 stop_codon
886 gene
886 transcript
895 CDS
933 exon
```