

NCBI



Biology Review

Disease

Function

Structure

Mature Peptide

ProPeptide

mRNA

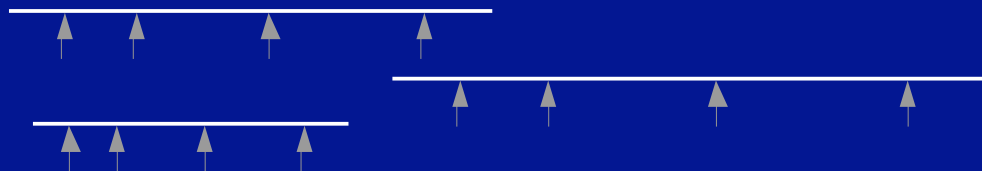
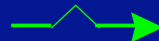
Transcript

Chromosome

Genetics

Genomes

Organisms



NCBI



DDBJ/EMBL/GenBank Model

■ Record (Locus)

- Commentary – Title, References, Comment
- Features –
 - location on the DNA sequence
 - information about what is there
 - most important is CDS -> protein
- DNA sequence

NCBI



Locus

LOCUS DMADHGC 7423 bp DNA linear INV 18-APR-2005

DEFINITION D. melanogaster alcohol dehydrogenase structural gene and flanks

FEATURES Location/Qualifiers

mRNA 5508..7050

CDS join(5544..5642,5708..6112,6183..6449)

/codon_start=1

/translation="MSFTLTNKNVIFVAGLGGIGLDTSKELLKRDLKNLVILDRIENP

AAIAELKAINPKVTVTFYPYDVTVPIAETTKLLKTIFAQLKTVDVVLINGAGILDDHQI

ORIGIN

```
1 catcctcgcc cgtttccacg ccgtcgtcct cctcatcatc ggcgagagct gattgcgtgg
61 tggtcagagg cgaaccagcg gtcttcgtgg agctgggacc cagatcaagg ctgctcaaca
```

NCBI



Accession

```
LOCUS          DMADHGC              7423 bp    DNA        linear     INV 18-APR-2005
DEFINITION     D. melanogaster alcohol dehydrogenase structural gene and flanks
ACCESSION     Z00030 J01066 M14802 V00197 V00198 V00199
FEATURES             Location/Qualifiers
     mRNA             5508..7050
     CDS             join(5544..5642,5708..6112,6183..6449)
                   /codon_start=1
                   /translation="MSFTLTNKNVIFVAGLGGIGLDTSKELLKRDLKNLVILDRIENP
                   AAIAELKAINPKVTVTFYPYDVTVPPIAETTKLLKTIFAQLKTVDVVLINGAGILDDHQI
ORIGIN
     1  catcctcgcc  cgtttcacg  ccgtcgtcct  cctcatcatc  ggcgagagct  gattgcgtgg
    61  tggtcagagg  cgaaccagcg  gtcttcgtgg  agctgggacc  cagatcaagg  ctgctcaaca
```

NCBI



GI numbers

LOCUS DMADHGC 7423 bp DNA linear INV 18-APR-2005
DEFINITION D. melanogaster alcohol dehydrogenase structural gene and flanks
ACCESSION Z00030 J01066 M14802 V00197 V00198 V00199
GI 7558

FEATURES Location/Qualifiers
mRNA 5508..7050
CDS join(5544..5642,5708..6112,6183..6449)
/codon_start=1
/db_xref="GI:599727"
/translation="MSFTLTNKNVIFVAGLGGIGLDTSKELLKRDLKNLVILDRIENP
AAIAELKAINPKVTVTFYPYDVTVPVIAETTKLLKTIFAQLKTVDVVLINGAGILDDHQI

ORIGIN

1 catcctcgcc cgtttcacg ccgtcgtcct cctcatcatc ggcgagagct gattgcgtgg
61 tggtcagagg cgaaccagcg gtcttcgtgg agctgggacc cagatcaagg ctgctcaaca

NCBI



NIDs and PIDs

LOCUS DMADHGC 7423 bp DNA linear INV 18-APR-2005

DEFINITION D. melanogaster alcohol dehydrogenase structural gene and flanks

ACCESSION Z00030 J01066 M14802 V00197 V00198 V00199

NID g7558

FEATURES Location/Qualifiers

mRNA 5508..7050

CDS join(5544..5642,5708..6112,6183..6449)

/codon_start=1

pid="g599727"

/translation="MSFTLTNKNVIFVAGLGGIGLDTSKELLKRDLKNLVILDRIENP
AAIAELKAINPKVTVTFYPYDVTVPVIAETTKLLKTIFAQLKTVDVLINGAGILDDHQI

ORIGIN

1 catcctcgcc cgtttcacg ccgtcgtcct cctcatcatc ggcgagagct gattgcgtgg
61 tggtcagagg cgaaccagcg gtcttcgtgg agctgggacc cagatcaagg ctgctcaaca

NCBI



Accession.version and /protein_id

```
LOCUS          DMADHGC              7423 bp    DNA        linear     INV 18-APR-2005
DEFINITION     D. melanogaster alcohol dehydrogenase structural gene and flanks
ACCESSION     Z00030 J01066 M14802 V00197 V00198 V00199
VERSION       Z00030.1  GI:7558
FEATURES             Location/Qualifiers
     mRNA             5508..7050
     CDS               join(5544..5642,5708..6112,6183..6449)
                     /codon_start=1
                     /protein_id="CAA77330.1"
                     /db_xref="GI:599727"
                     /translation="MSFTLTNKNVIFVAGLGGIGLDTSKELLKRDLKNLVILDRIENP
AAIAELKAINPKVTVTFYPYDVTVPVIAETTKLLKTIFAQLKTVDVVLINGAGILDDHQI

ORIGIN
1  catcctcgcc  cgtttccacg  ccgtcgtcct  cctcatcatc  gccgagagct  gattgcgtgg
61  tggtcagagg  cgaaccagcg  gtcttcgtgg  agctgggacc  cagatcaagg  ctgctcaaca
```

NCBI



Summary

- ID should not be a name
- Clear definition of what ID identifies
- Must have practical utility to community
- Must be easily unique across assigning groups
- Ongoing struggle in compound objects and objects in wide use with different audiences.