|  |  |  |
| --- | --- | --- |
| TABLE | main \_data |  |
| PK | **acc\_num** | **VARCHAR(10)** |
|  | chrom\_loc | VARCHAR(20) |
|  | gene\_id | VARCHAR(10) |
|  | prot\_name | VARCHAR(50) |
|  | locus\_span | VARCHAR(15) |
|  | gene\_span | VARCHAR(15) |
|  | cds\_span | VARCHAR(15) |
|  | start\_cod | CHAR(8) |

|  |  |  |
| --- | --- | --- |
| TABLE | seq\_data |  |
| PK | **acc\_num** | **VARCHAR(10)** |
|  | cds\_map | TEXT |
|  | locus\_seq | MEDIUMTEXT |

**Figure 2. Physical model describing a MySQL relational database containing Genbank data fields describing protein coding genes.**

CDS = coding DNA sequence; acc\_no = accession number;

chrom\_loc = chromosomal location; gene\_id = name of gene feature; locus\_seq = Genbank record DNA sequence;

gene\_span = gene sequence span; cds\_span = CDS span;

cds\_map = list of CDS elements comprising CDS span;

prot\_name; PK = Primary Key; FK = Foreign Key

Sequence spans are expressed as a range of base numbers.