Mammalian-PERFECT

addGeneNr

株式会社メイズ maze

Allows for easy differentiation between similar proteins based on the UnigeneID

A protein database for rapid gene identification

＜FEATURE＞

In use addGeneNr ,seppending the Unigene ID Appending the Unigene ID to
the beginning of each NCNInr description lets you distinguish at a glance proteins
that are alternative splicing variants from those originating from different gene families.

◆ **ADVANTAGE**

The NCBI nr database contains sequences for both alternative splicing variants of the same gene and those from different genetic families.
With a long list of homologous search results, it is impossible to tell meaningful differences at a glance.

```
1.  gi|7414205                    73        Score: 111    E
    unnamed protein product [
    gi|12843160                   Mass:            Score: 104    Ex
    unnamed protein product [Mus musculus]
    gi|200679                     Mass: 20671      Score: 83     Ex
    retinol-binding protein

2.  gi|206589                     Mass: 20331      Score: 110    Ex
    retinol-binding protein
    gi|132407                     Mass: 23547      Score: 104    Ex
    Plasma retinol-binding protein precursor (PRBP) (RBP)

3.  gi|62642230
    PREDICTED: retinol b
```

Score

Matched proteins

Both results are from the AHSP70 gene, but the differing UnigeneID tells us at a glance that these are not splicing variants, but come from different gene families.
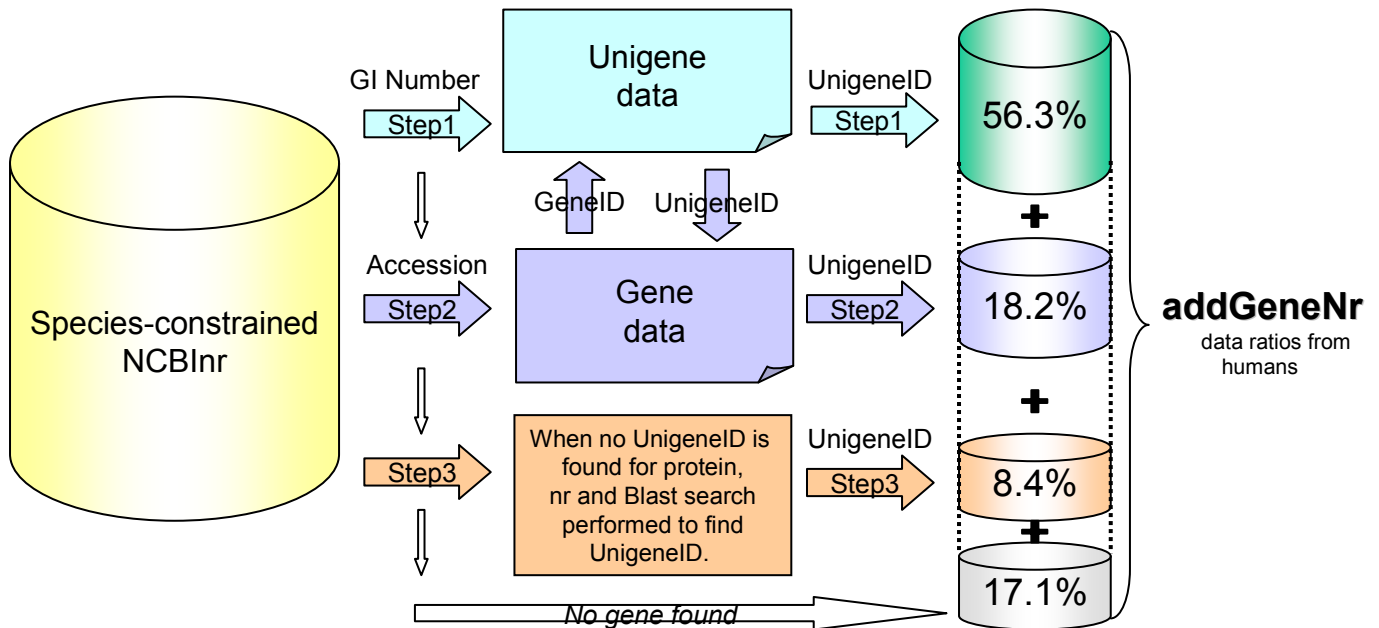
```
gi|16507237      Mass: 72446      Score: 1065   Queries matched: 19
/uid-Hs.52293 heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa) [Homo sapiens]gi|69001
☐ Check to include this hit in error tolerant search or archive report

Query  Observed  Mr(expt)  Mr(calc)  Delta  Miss Score  Expect   Rank Peptide
 461    729.49    728.48    727.89   0.59   0   31     1.1     1   K.IQQLVK.E
1771    615.08   1228.15   1228.31  -0.16   0   65   0.00033   1   R.VEIIAHDQGNR.I
1838    621.56   1241.10   1240.43   0.68   1   55   0.0033    1   K.MKETAEAYLGK.K
2222    659.53   1317.05   1316.41   0.64   0   65   0.00036   1   R.NELESYAYSLK.N
2646    700.14   1398.25   1397.61   0.64   0   32    0.62     1   K.ELEEIVQPIISK.L
3488    758.20   1514.37   1512.73   1.65   1   89   1.2e-006  1   R.AKFEELNMDLFR.S 3605
3664    769.75   1537.47   1536.79   0.69   0   65   0.00033   1   K.TFAPEEISAMVLTK.M
4083    796.01   1590.01   1588.76   1.26   1  108   1.8e-008  1   K.KSDIDEIVLVGGSTR.I
4195    803.29   1604.57   1604.80  -0.23   0   83   5.2e-006  1   K.TKPYIQVDIGGGQTK.T
4622    830.77   1659.52   1659.88  -0.36   0   49    0.014    1   R.IINEPTAAAIAYGLDK.R
4819    839.95   1677.89   1677.77   0.12   0   78   1.8e-005  1   K.NQLTSNPENTVFDAK.R
6042    908.99   1815.97   1816.06  -0.10   1   73   4.4e-005  1   R.IINEPTAAAIAYGLDKR.E
6070    910.69   1819.36   1818.89   0.48   0   33    0.52     2   K.LYGSAGPPPTGEEDTAEK.D
6217    919.53   1837.05   1836.99   0.05   0   55    0.003    1   K.SQIFSTASDNQPTVTIK.V
7170    968.10   1934.19   1934.16   0.03   0   54    0.004    1   K.DNHLLGTFDLTGIPPAPR.G
7519    988.75   1975.49   1975.07   0.41   0   73   4e-005    1   K.IEWLESHQDADIEDFK.A
8895   1083.67   2165.31   2165.27   0.05   0   34    0.35     1   R.IEIESFYEGEDFSETLTR.A
8972   1088.83   2175.64   2176.25  -0.61   1   33    0.49     1   K.LYGSAGPPPTGEEDTAEKDEL.-

32.  gi|24234686      Mass: 53631      Score: 646   Queries matched: 13
/uid-Hs.180414 heat shock 70kDa protein 8 isoform 2 [Homo sapiens]gi|11526573|dbj|BAB18615.1| heat
☐ Check to include this hit in error tolerant search or archive report

Query  Observed  Mr(expt)  Mr(calc)  Delta  Miss Score  Expect   Rank Peptide
1771    615.08   1228.15   1228.31  -0.16   0   65   0.00033   1   K.VEIIANDQGNR.T
1817    619.51   1237.00   1235.45   1.55   0   66   0.00026   1   R.MVNHFIAEFK.R
1899    627.99   1253.96   1252.48   1.48   1   60   0.0011    1   K.MKEIAEAYLGK.T
1903    628.30   1254.58   1253.36   1.22   0   50    0.011    1   R.FEELNADLFR.G
2738    707.21   1412.41   1410.55   1.86   1   35    0.3      1   R.RFDDAVVQSDMK.H
3270    742.62   1483.22   1481.65   1.57   0   99   1.3e-007  1   K.SQIHDIVLVGGSTR.I
3301    744.81   1487.60   1487.57   0.04   0   43    0.052    1   R.TTPSYVAFTDTER.L
4421    817.33   1632.65   1632.83  -0.18   0   32    0.75     1   K.SFYPEEVSSMVLTK.M
4622    830.77   1659.52   1659.88  -0.36   0   49    0.014    1   R.IINEPTAAAIAYGLDK.K
4703    834.30   1666.58   1665.82   0.75   0   75   3.2e-005  1   K.NQVAMNPTNTVFDAK.R
4940    846.82   1691.63   1691.67  -0.03   0   49    0.014    1   K.STAGDTHLGGEDFDNR.M 4944
9633   1130.97   2259.92   2260.46  -0.53   0   32    0.51     1   K.SINPDEAVAYGAAVQAAILSGDK.S

Proteins matching the same set of peptides:
gi|62896815      Mass: 53613      Score: 646   Queries matched: 13
/uid=Hs.180414 heat shock 70kDa protein 8 isoform 2 variant [Homo sapiens]
```

※Above: MatrixScience's Mascot search engine

Try out !!

Free Download http://pronem.maze.co.jp/

addGeneNr DATABASE CONSTRUCTION

GI number and UnigeneID are found for each sequence in the NCBInr database and the UnigeneID appended to the front of the nr description

Step1 · GI Number → Unigene data → UnigeneID · Step1 → 56.3%

GeneID · UnigeneID

Step2 · Accession → Gene data → UnigeneID · Step2 → 18.2%

Step3 · When no UnigeneID is found for protein, nr and Blast search performed to find UnigeneID. → UnigeneID · Step3 → 8.4%

No gene found → 17.1%

addGeneNr data ratios from humans

◆ HOW TO SETUP

It can set up in house server of the protein search engine, can utilize. Setup is possible with simple job.

◆ UPDETA & MORE INFORMATON

addGeneNr is free for use on an ongoing basis.

http://pronem.maze.co.jp/en/

• we have other types of a highly useful proteomics database

Maze, Inc.

TS Building, #101 3-20-2 Hatagaya, Shibuya-ku, Tokyo
Tel:813-3378-7430 Fax:813-3378-7247
http://www.maze.co.jp/en
sales@maze.co.jp

We are Bioinformatics Service Provider.
maze
株式会社メイズ