

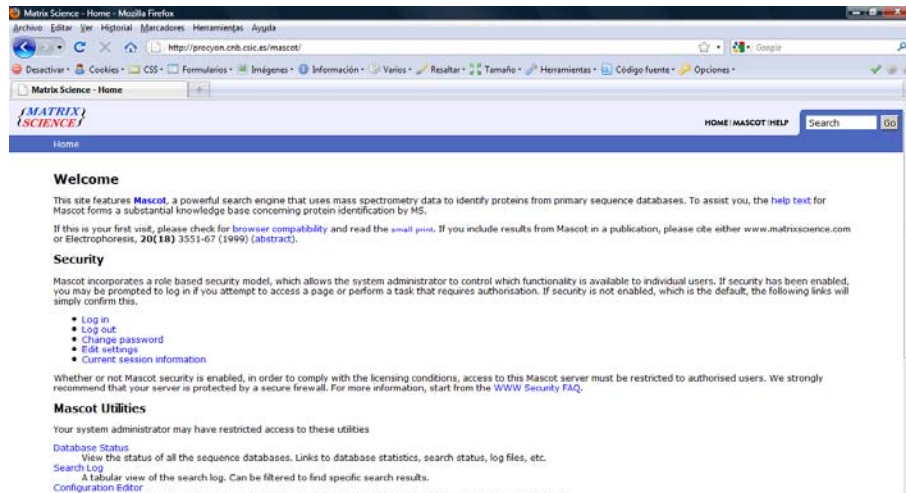
# How to install the PME6 sequence databases in MASCOT

## Copy the fasta files in your local MASCOT server

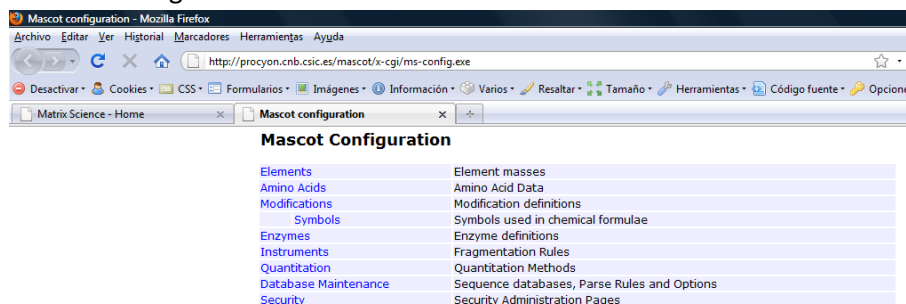
1. Go to sequence folder on your MASCOT server (ej: E:\MASCOT\sequence)  
(note that your local server installation can probably be located in other folder or unit)
2. Create one folder for each database: E:\MASCOT\sequence\PME6\_decoy
3. Go to each folder and create a subfolder called "current":  
E:\MASCOT\sequence\PME6\_decoy\current
4. Copy the fasta files in each folder: PME6\_decoy.fasta to  
E:\MASCOT\sequence\PME6\_decoy\current

## Install PME6 database in your local MASCOT server

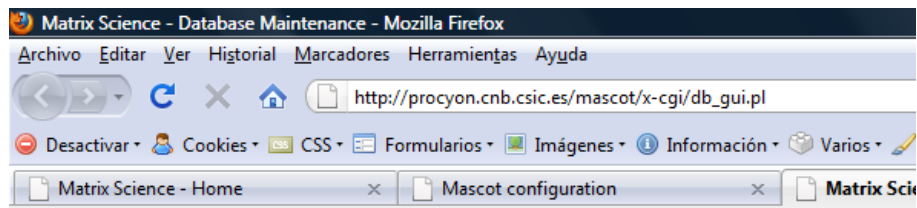
1. Go to the main web page in your local MASCOT server (ej: <http://procyon.cnb.csic.es/mascot/>)



2. Click on Configuration Editor



3. Click on Database Maintenance



### Mascot Database Maintenance: Edit Database Definitions

Help Window

Select: MSDB

Name: MSDB  Active  Inactive

Path: E:/MASCOT/sequence/MSDB/current/MSDB\_\*.fasta

AA  NA  Mem map  Mem lock

Threads: 8 Local ref file

Taxonomy: MSDB REF

Rule to parse accession string from Fasta file:  
Rule 17: ">\([^ ]\*\)"

Rule to parse description string from Fasta file:  
Rule 18: ">[^ ]\* \(.\*\)"

Rule to parse accession string from local reference file:  
Rule 19: ">[A-Z] [0-9];\([^ ]\*\) [ ]\*"

Source and parse rule for full text report (optional):

Host: localhost Port: 80

Path: C:/INETPUB/MASCOT/x-cgi/ms-getseq.exe MSDB #ACCESSION# a  
Rule 9: "\.\*.\*\(>[A-Z] [0-9];.\*\)"

4. We have to add a new database definition for each database, so click on New definition button and fill de fields as follows:
  - o Name: PME6\_decoy
  - o Path: E:\MASCOT\sequence\PME6\_decoy\current\\*.fasta
  - o Local ref file: No selected
  - o Taxonomy: --- None ---
  - o Rule to parse accession string from Fasta file: You have to select a rule like this: ">\([^ ]\*\)"
    - If you don't have this rule in your rule list:
      - Click on "Edit parse rules" button
      - Click on "Add new rule" button
      - Paste ">\([^ ]\*\)" in the last rule field, the one that is empty.
      - Click on "Return to database definitions" button.
  - o Rule to parse description string from Fasta file: You have to select a rule like this: ">[^ ]\* \(.\*\)"
    - If you don't have this rule in your rule list, proceed as in previous section to add this new rule.
  - o Rule to parse accession string from local reference file: ---no local reference file---
  - o Host: [www.expasy.org](http://www.expasy.org)

- Port:80
- Path: /cgi-bin/get-sprot-raw.pl?#ACCESSION#
- Rule: "\(.\*\)" (proceed as in previous sections if you don't have this rule in your rule list).

### Mascot Database Maintenance: Edit Database Definitions

Help Window

---

Select

Name  Active  Inactive

Path

AA  NA  Mem map  Mem lock

Threads  Local ref file

Taxonomy

Rule to parse accession string from Fasta file:

Rule to parse description string from Fasta file:

Rule to parse accession string from local reference file:

Source and parse rule for full text report (optional):

Host  Port

Path

**No changes are written to mascot.dat until you choose:**

5. Click on "Test this definition" to test the parse rules. You have to obtain a result like this:

## Mascot Database Maintenance

### Testing Database Definition PME6\_decoy

Testing entries at beginning and end of E:/MASCOT/sequence/PME6\_decoy/current/PME6\_decoy.fasta:

Accession	Description
P00883	Fructose-bisphosphate aldolase A OS=Oryctolagus cuniculus GN=ALDOA PE=1 SV=2
P00489	Glycogen phosphorylase, muscle form OS=Oryctolagus cuniculus GN=PYGM PE=1 SV=3
P02666	Beta-casein OS=Bos taurus GN=CSN2 PE=1 SV=2
P31946	14-3-3 protein beta/alpha OS=Homo sapiens GN=YWHAB PE=1 SV=3
P62258	14-3-3 protein epsilon OS=Homo sapiens GN=YWHAE PE=1 SV=1
rev_Q6WRX3	Reverse sequence, was Protein zyg-11 homolog A OS=Homo sapiens GN=ZYG11A PE=2 SV=2
rev_Q9C0D3	Reverse sequence, was Protein zyg-11 homolog B OS=Homo sapiens GN=ZYG11B PE=1 SV=2
rev_Q15942	Reverse sequence, was Zyxin OS=Homo sapiens GN=ZYX PE=1 SV=1
rev_O43149	Reverse sequence, was Zinc finger ZZ-type and EF-hand domain-containing protein 1 OS=Homo sapiens GN=ZZEF1 PE=1 SV=6
rev_Q8IYH5	Reverse sequence, was ZZ-type zinc finger-containing protein 3 OS=Homo sapiens GN=ZZZ3 PE=1 SV=1

### Example of full text report:

```
ID ALDOA_RABBIT Reviewed: 364 AA.
AC P00883; Q28671;
DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DT 23-JAN-2007, sequence version 2.
DT 15-JUN-2010, entry version 104.
DE RecName: Full=Fructose-bisphosphate aldolase A;
DE EC=4.1.2.13;
DE AltName: Full=Muscle-type aldolase;
GN Name=ALDOA;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
```

- Click on "Return to database definitions". If you have obtained some error, or the Accession column has captured any other data, please review the parse rules again. It is important that all laboratories capture the Accession of each protein as you can see in previous figure (P00883, P00489, ...).
- Click on "APPLY" button to save changes.
- You will obtain this screen if all is OK:

## Mascot Database Maintenance

Changes have been saved to	<a href="#">mascot.dat</a>
The previous contents of mascot.dat are in	<a href="#">mascot.dat.160</a>
To make further changes, <b>do not</b> press the browser back button, Follow	<a href="#">This link</a>
Display	<a href="#">Database Status</a>

- Click on Database Status to see how the database is being installed automatically by MASCOT:

---

```
Name = PME6_decoy Family = E:/MASCOT/sequence/PME6_decoy/current/*.fasta
Filename = PME6_decoy.fasta Pathname = E:/MASCOT/sequence/PME6_decoy/current/PME6_decoy.fasta
Status = In use Statistics
State Time = Mon Aug 09 13:40:44 # searches = 0
Mem mapped = YES Request to mem map = YES Request unmap = NO Mem locked = YES
Number of threads = 8 Current = YES
```

---