

# ***READ ME FIRST!***

## **Installing Mascot for the first time?**

If you are installing Mascot for the first time, please read the notes starting on page 3 before following the installation instructions in Chapter 3 of the Installation and Setup Manual.

## **Upgrading from an earlier version of Mascot?**

If you already have a working version of Mascot, please read the notes starting on page 4. Note that complete, up-to-date copy of the Installation and Setup Manual can be found on the CD-ROM (*manual.pdf*).

## **New: Memory efficient mode for MudPIT searches**

MS/MS searches of any size can now be performed on 32 bit platforms.

## **New: Sequence Tags**

Sequence tags, including error tolerant tags, are fully implemented.

## **New: Reports are faster and more memory efficient**

Reports are much faster and more memory efficient because they now use compiled code from Mascot Parser.

## **New: Mascot Distiller support in Mascot Daemon**

Mascot Daemon can use Mascot Distiller to process binary files in batch or real-time monitor mode, (requires additional licence).

## **New: Mascot Daemon runs as service**

Mascot Daemon now runs as service and connects to a Mascot server asynchronously.

## **And lots more**

Support for high energy side chain cleavage ions (d, v, w); Web pages redesigned to eliminate frames; Full on-line help for Mascot Daemon; Nucleic acid translation now uses taxonomy to determine correct genetic code; Database maintenance Utility has simplified interface; ambiguous residues X, B and Z are explicitly iterated to find best match; etc.; etc.

## **Changes since last release**

Details of individual changes since the previous major release of Mascot are listed on page 6 of these notes.

## **Last minute additions to documentation**

See page 2

## *Last minute additions to documentation*

---

- If, for any reason, you need to uninstall and then re-install Mascot, you **must** reboot your system between the uninstall and the re-install.
- If you need to install Mascot from a remote drive, this cannot be a UNC path. You must first map the remote drive to a local drive letter
- Windows 2003 is fully supported by the installation program. It should not be necessary to perform any manual configuration of IIS 6.
- Under IIS 6 (Windows 2003), if a search is submitted from a browser and the connection is broken before the search is complete, the search will be killed. We are working to resolve this new feature of IIS 6. One workaround is to use Apache 2.x as your web server.
- The installation program now supports and configures Apache 2.x web server. If you configure Apache yourself, ensure that ForkForUnixApache in the Options section of mascot.dat is set to 1
- Perl 5.8 installation requires Microsoft Windows Installer 2.0 or later. Under Windows NT4 and 2000, you may need to update Windows Installer. To do so, change to the daemon directory on the Mascot CD, and execute the file InstMsiW.exe
- Some lines of text in the Setup and Installation manual may not display correctly in old versions of Adobe Acrobat. The solution is to upgrade Acrobat reader to a more recent version
- If you use Mascot Daemon to search data files from remote drives, you will need to change the service to log on as a named user. See the Mascot Daemon on-line help for further details.
- Several report scripts and utilities now use Mascot Parser. The latest version will be installed into the mascot\bin directory. If you already have an earlier version of Mascot Parser installed, you should update this to the same version (1.2.2) to avoid potential confusion.

# *Installing Mascot for the first time*

---

## **Perl**

If Perl is already installed, and the version is less than 5.6, it must be upgraded. Uninstall your old version of Perl, then install Perl 5.8 from the Mascot CD. The procedure is described on page 30 of the manual.

Under Windows NT4 and 2000, you may need to update Windows Installer. To do so, change to the daemon directory on the Mascot CD, and execute the file InstMsiW.exe

## **Hyper-threading**

Mascot now supports hyper-threading on Intel Pentium 4 Xeon processors. See page 41 of the manual for full details

## **To install Mascot Daemon**

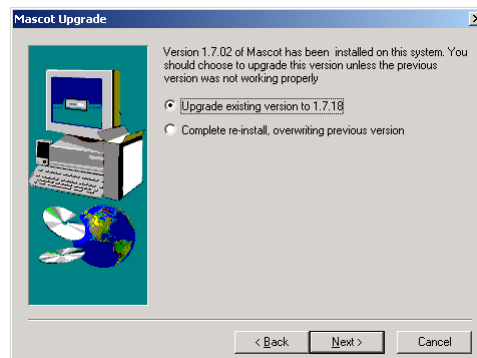
Refer to the file README.rtf in the daemon directory on the Mascot CD

# Upgrading Mascot

---

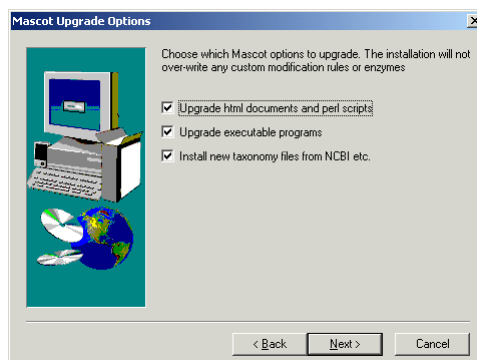
## To upgrade Mascot

1. If your version of Perl is less than 5.6, it must be upgraded. Uninstall Perl, then install Perl 5.8 from the Mascot CD. The procedure is described on page 30 of the manual. Under Windows NT4 and 2000, you may need to update Windows Installer. To do so, change to the daemon directory on the Mascot CD, and execute the file InstMsiW.exe
2. Ensure that any other users are aware that they will not be able to perform searches while the upgrade is taking place.
3. Check that you are logged in with administrator privileges, put the Mascot CD in the drive, and run the Setup.exe program
4. Check the details on the screen that shows the information about system for any errors. If anything on this screen is incorrect, please refer to the troubleshooting section of the manual, beginning page 42.
5. In the following screen:



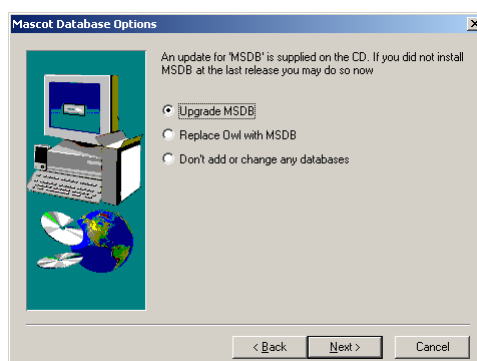
choose to upgrade unless the previous version was not working properly, or if you want to install Mascot onto a new drive, or install it into a new Web Server. Note that it is not possible to run both the old and the new version of Mascot on the same system.

6. Press 'Next' to display the following screen:



You should leave all options selected unless you understand the system very well, and have customised some perl scripts or html pages. Any additional enzymes that you have added will be unaltered. A new mod\_file will be installed, but your existing mod\_file will be backed up. You will need to manually add any custom modifications that you previously added to the new mod\_file after the installation has been completed.

7. Press 'Next' to display the following screen:



Check that the installation directory for MSDB is correct, and choose the appropriate option for sequence database setup.

8. Press Next and wait for the installation to complete.

**NOTE:** It is not always possible for the installation program to update your own additions and changes in mascot.dat. It is always a good idea to compare your current version of mascot.dat with the latest template file, not.mascot.dat, to judge whether further changes should be made. Mascot.dat entries are fully described in Chapter 6 of the Setup and Installation manual.

## To upgrade Mascot Daemon:

Refer to the file README.rtf in the daemon directory on the Mascot CD

## Changes since Version 1.9.05

ID	Description
88	Add support for sequence tag. This is a peptide mass together a pair of fragment ion masses which enclose a short stretch of interpreted sequence. E.g. 1890 tag(1004, LSADTG, 1548)
506	Validate entries in Mascot Daemon data import options dialog (e.g.sequest params)
1466	New option in mascot.dat – UnixDirPerm. Provides fine grained control over permissions of result file directories
1471	Support high energy side chain cleavage ions (w, d, and v)
1473	Memory efficient mode for large searches: In Memory efficient mode, queries are searched in batches. Memory usage is minimised by storing intermediate results on disk. Also, in the case of a cluster, results are continuously trickled back to the master node. In Memory efficient mode, there are no results in the Summary block of the results file, so it is not possible to create a protein summary report. Memory efficient mode is principally designed to work around our inability to search MudPIT data on a 32 bit platform. It also eliminates the communication bottleneck at the end of a large search on a cluster.
1474	Fix false PMF false positives caused by abnormal amino acid abundances. E.g. poly K
1475	Clean Windows install not creating MSDB/old MSDB/incoming
1476	Windows install making unnecessary changes to MSDB Reference parse rule in mascot.dat
1478	Mascot Daemon loses external process strings when resuming a paused task
1479	In the database status page, numbers for CPUs not shown correctly for Linux
1480	Make error 133 non fatal (peptide mass too small)
1481	If there are many databases, Database Maintenance utility form becomes unwieldy. It would be better to display the definitions for just one database at a time, selected via a drop-down list
339	Allow B and Z in seq, comp and tag statements. seq(n-[BZ]X) should be 'translated' to seq(n-[NDBEQZ][A-Z])
477	Add error tolerant sequence tag that can match to a peptide even when there is an unsuspected modification or point mutation by allowing either the ion masses or the peptide mass to "float".
530	If sequence contains an X, MS/MS search now iterates through all 20 residues looking for best match. Iterates N and D for B and E and Q for Z.
639	Don't lock MSDB if less than 1GB RAM
679	The genetic code is not the same for all species. Mascot now uses taxonomy to determine the correct table. See <a href="http://www.ncbi.nlm.nih.gov/htbin-post/Taxonomy/wprintgc?mode=c">http://www.ncbi.nlm.nih.gov/htbin-post/Taxonomy/wprintgc?mode=c</a>
778	Option to show title lines and molecular weights for ALL matches in a master results report, even those which are a single peptide. Needs to be an option because it could slow down large reports by making hundreds of calls to ms-getseq.exe
799	Need error message if taxonomy specified as coming from Swiss-Prot DAT file, but there is only a Fasta file
997	Titles in the enzymes file should allow spaces
998	Mascot Daemon real time monitor mode to support wildcards in directory names as well as in file names
1038	Fix Mascot Daemon problem: When exiting, a window pops up and asks for MS Jet OLE-Db Initialisation Information
1070	In cluster mode, if fail to copy file onto node, this should be treated as an error rather than carrying on blindly
1138	Should preserve case on user defined parameter names (with leading underscore)

ID	Description
1167	owl.dmp has a missing tab in the line: 6724   Ibacus ciliatus
1220	Registry problems if Mascot Daemon installed by Administrator then run as non-privileged user
1234	Add options to sort unassigned list in peptide summary by (i)ascending query number, (ii) descending score, (iii) descending TIC
1261	In cluster mode, if run out of disk space on node when distributing compressed files, the master computer doesn't get the error code back
1296	Mascot Daemon should connect asynchronously to Mascot. Mainly to improve reliability for very long searches
1305	IgnoreIonsScoreBelow should take effect in the result reports, not during the search
1336	Memory allocation handler needs to be implemented to catch failures in memory allocation, raise an error and quit.
1381	Default ICAT mods can now be specified in mascot.dat
1382	Need to speed up MSDB compression
1399	Need to update db_gui.pl and browser.pl Javascript for latest generation browsers
1415	Fix problem with weak PMF false positives in G rich protein
1419	Mascot Daemon needs reasonable time delay before re-submitting a search when error is "too many searches"
1438	Enable Mascot Daemon to connect to secure server (https://)
1448	Fix problem in database maintenance utility – if there were multiple sets of Fasta and Dat / Ref files in the current directory, utility did not pair them off correctly
1449	Master_results.pl should use the electron mass from the results file (where available)
1452	Unigene results – occasional missing accessions in yellow popups.
1458	Database maintenance utility was commenting out Processors section of mascot.dat
1461	Need fatal error if taxonomy files missing
1462	In Mascot Daemon, Wiff sample name was not carried over to title field of follow-up searches
1485	Help for SwissProt said cannot use AC with local dat file ... not true
1486	Protein view could try to create GenBank feature table report for a PMF
1487	Mascot Daemon: In Bioworks 3.1, it seems that extract_msn.exe works in place of lq_dta.exe. So, need to remove the name check
1488	Parsing errors if an MGF file had an ions block with no fragment ion data.
1494	With mixed pmf and ms-ms spectra, could get discrepancies in variable modifications between peptides block and summary block in result file
1495	Problems with substring matching for modifications which begin ICAT_light or ICAT_heavy. Never get to the longer strings if they are later in mod_file
1496	Incorrect BRE tests successfully in database maintenance utility. For example, rules using \w or \d would succeed because script is using ERE
1499	Mascot Wizard could create an input file that crashed Mascot
1501	Hit numbering could go wrong in result report if RequireBoldRed set to 1
1502	A file of 3147 bytes of 'null' characters crashes nph-mascot:
1503	Combinations of variable terminus mods were reported in masses section of results file
1511	Limit of 16 cpu for a Windows cluster setup increased to 1024
1515	In protein view, "Links to retrieve other entries containing this sequence from NCBI Entrez" were broken.
1516	A very large repeat search under AIX could stall, right at the beginning
1517	Linux regex library could crash when matching long strings. The original fix was to truncate the string at 200 characters. However, IPI database now has accession strings longer than 200 characters, so there was no way to get the description. Now properly fixed
1518	Corrupt ions ions query could cause crash

ID	Description
1519	Need to disable SEG (protein mass) for scoring PMF mixtures. Otherwise, can get false positives
1522	Error tolerant search - can select which peptides to repeat the search with. Need to enforce search of all peptides because reports assume a one to one mapping when checking to see if the score is greater than the previous score.
1523	Individual accessions for PMF mixture components in full protein summary should be hyperlinks to protein view
1526	Mascot Daemon gave incorrect error if follower task owned by remote user
1527	Help text for sequence setup doesn't say anywhere that the taxonomy files should be put into the taxonomy directory and should be uncompressed
1533	Email messages are no longer sent to the Mascot administrator when the server is overloaded, ("too many searches")
1534	Error tolerant search - If you have &_showallfromerrortolerant=1 on the URL for master_results.pl, and then click on the protein view, the &_showallfromerrortolerant=1 state should be passed on
1535	Intranet home page should have prominent warning about being behind a firewall
1536	Add side chain cleavage series dvw to peptide_view.pl
1537	Require 3 new fragmentation rules for d, v, and w ions 22 # d and d'series 24 # v series 25 # w and w'series
1538	Monitor was not deleting *.nam files
1539	Mascot Daemon not doing follow up if MS/MS search had no results
1540	Database maintenance utility could fail to get current directory, making paths in WWW wrong (e.g. /ms-getseq.exe instead of /usr/local/mascot/x-cgi/ms-getseq.exe).
1541	Master results needs to decompress result file for error tolerant search parent
1542	Unrecognised INSTRUMENT should give an error message and halt
1543	WWW parse rule 24 is wrong. It can match to Swiss-Prot line FT DISULFID 581 590
1545	The number of queries in the summary block is not updated to reflect queries rejected because they were too big in mass.
1547	Can no longer compress EST_others on some Windows systems – requires too much memory
1548	Licence by physical cpu rather than logical cpu. Particularly important for Linux with hyper-threaded cpu
1549	lcq_dta_shell.pl fails with Netscape 7. Error is 'cannot copy temporary file'. Problem seems to be that this browser returns uploaded file name rather than path.
1550	Selecting 221 variable modifications in the search form caused crash due to overflow!
1551	Scripts need to look for msparser in mascot/bin directory
1565	Mascot Daemon with European country setting gives Invalid Property Value (Run time error 380) with lcq_dta
1567	Add expectation values to reports
1568	Most common error conditions should be intercepted before search starts.
1572	Add z series back to ECD in fragmentation rules
1587	Installation needs check Perl version number and error if not 5.6 or 5.8
1591	When a user specifies b- y- in ions() statement nph-mascot.exe still finds matches with any available ion series, although for scoring purposes it calculates all constants correctly.
1594	Installer - needs a new way of uncompressing MSDB because it is too big to fit on the CD when compressed with the InstallShield compress